# DESI and the missing-observation problem

Davide Bianchi

ICG Portsmouth

Sexten 02/07/2018

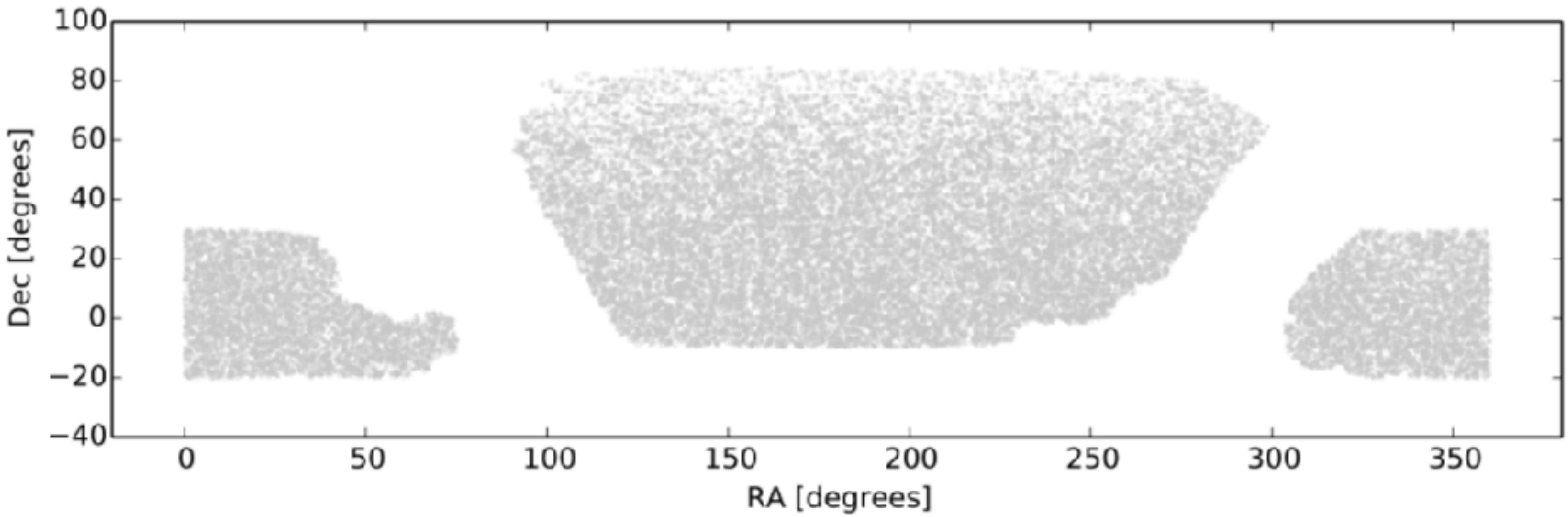
# Based on:

- **DB** & W. J. Percival - arXiv:1703.02070
- W. J. Percival & **DB** - arXiv:1703.02071
- **DB** et al. - arXiv:1805.00951

# The Dark Energy Spectroscopic Instrument (DESI) galaxy survey


Mayall telescope (4m) in kitt Peak, Arizona

# DESI footprint




$$\text{total area} \sim 14,000\ \text{deg}^2$$

$$\text{total number of targets} \sim 30 \times 10^6$$

# DESI targets

| Galaxy type | Redshift range | Bands used | Targets per deg$^2$ | Exposures per deg$^2$ | Good $z$'s per deg$^2$ | Baseline sample |
|---|---|---|---|---|---|---|
| LRG | 0.4–1.0 | $r$,$z$,$W1$ | 350 | 580 | 285 | 4.0 M |
| ELG | 0.6–1.6 | $g$,$r$,$z$ | 2400 | 1870 | 1220 | 17.1 M |
| QSO (tracers) | $< 2.1$ | $g$,$r$,$z$,$W1$,$W2$ | 170 | 170 | 120 | 1.7 M |
| QSO (Ly-$\alpha$) | $> 2.1$ | $g$,$r$,$z$,$W1$,$W2$ | 90 | 250 | 50 | 0.7 M |
| **Total in dark time** | | | **3010** | **2870** | **1675** | **23.6 M** |
| BGS | 0.05–0.4 | $r$ | 700 | 700 | 700 | 9.8 M |
| **Total in bright time** | | | **700** | **700** | **700** | **9.8 M** |

DESI collaboration arXiv:1611.00036

ELG = emission line galaxies

LRG = luminous red galaxies

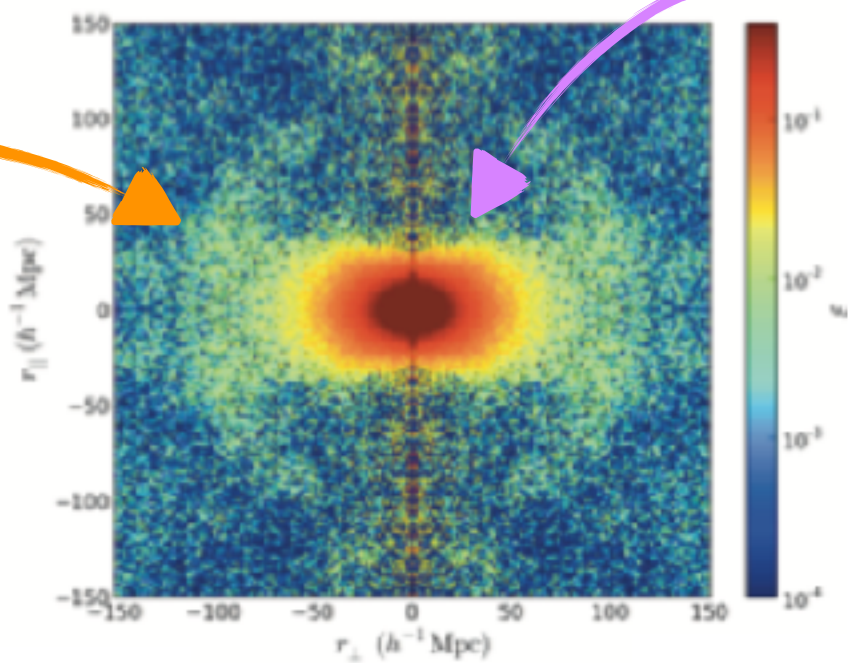
QSO = quasars

BGS = bright galaxy survey

# Why do we want to measure the 3D position of 30M galaxies?

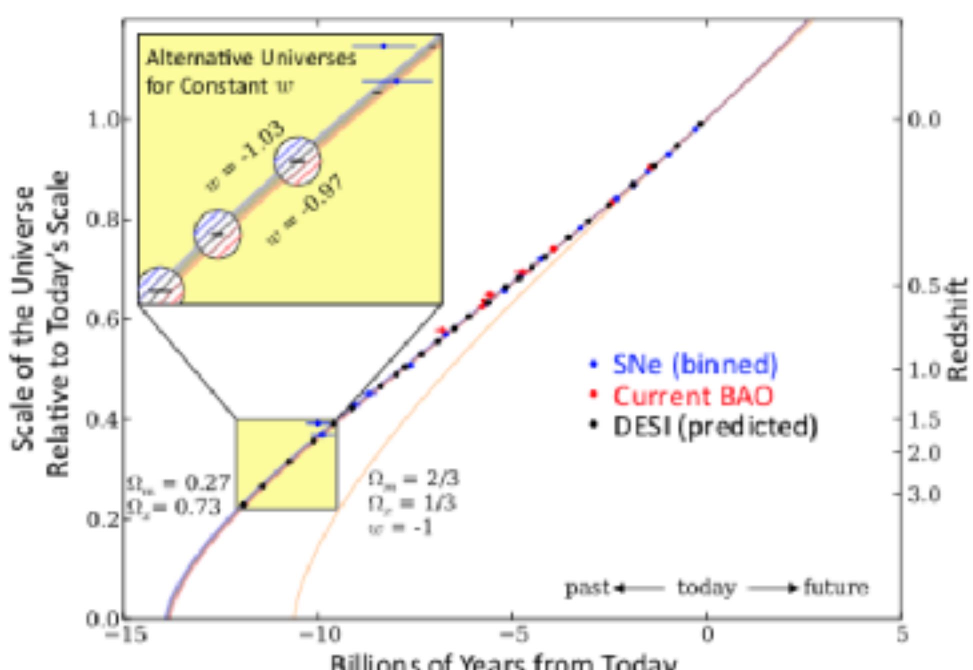
2-point correlation function

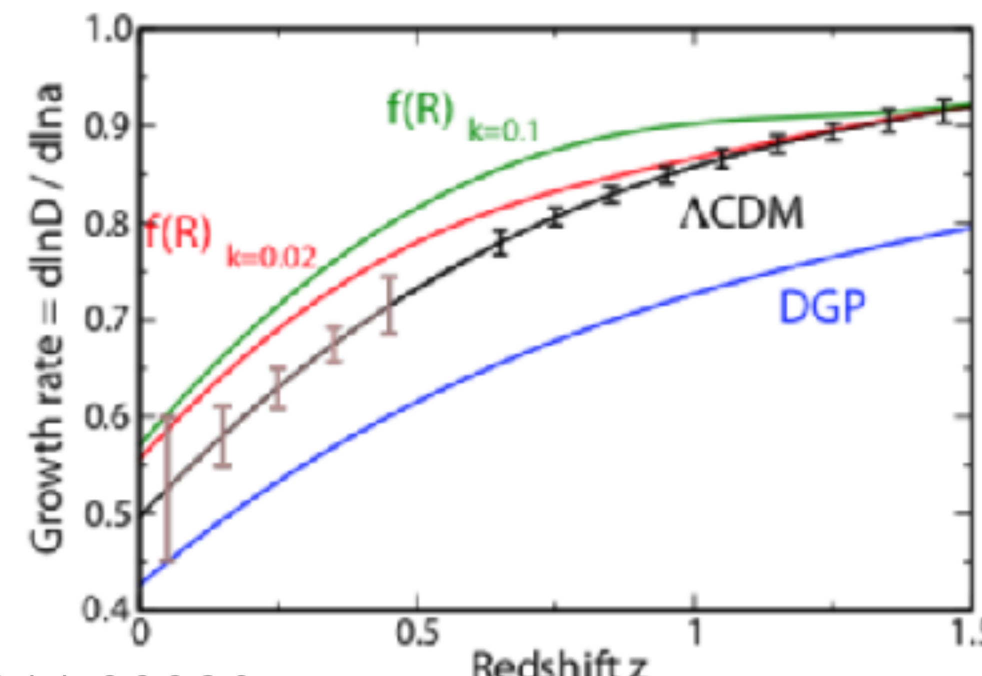
baryon acoustic oscillation (BAO) peak




BOSS DR11

redshift space distortion (RSD) squashing




DESI collaboration arXiv:1611.00036

# Parent catalogue of potential targets

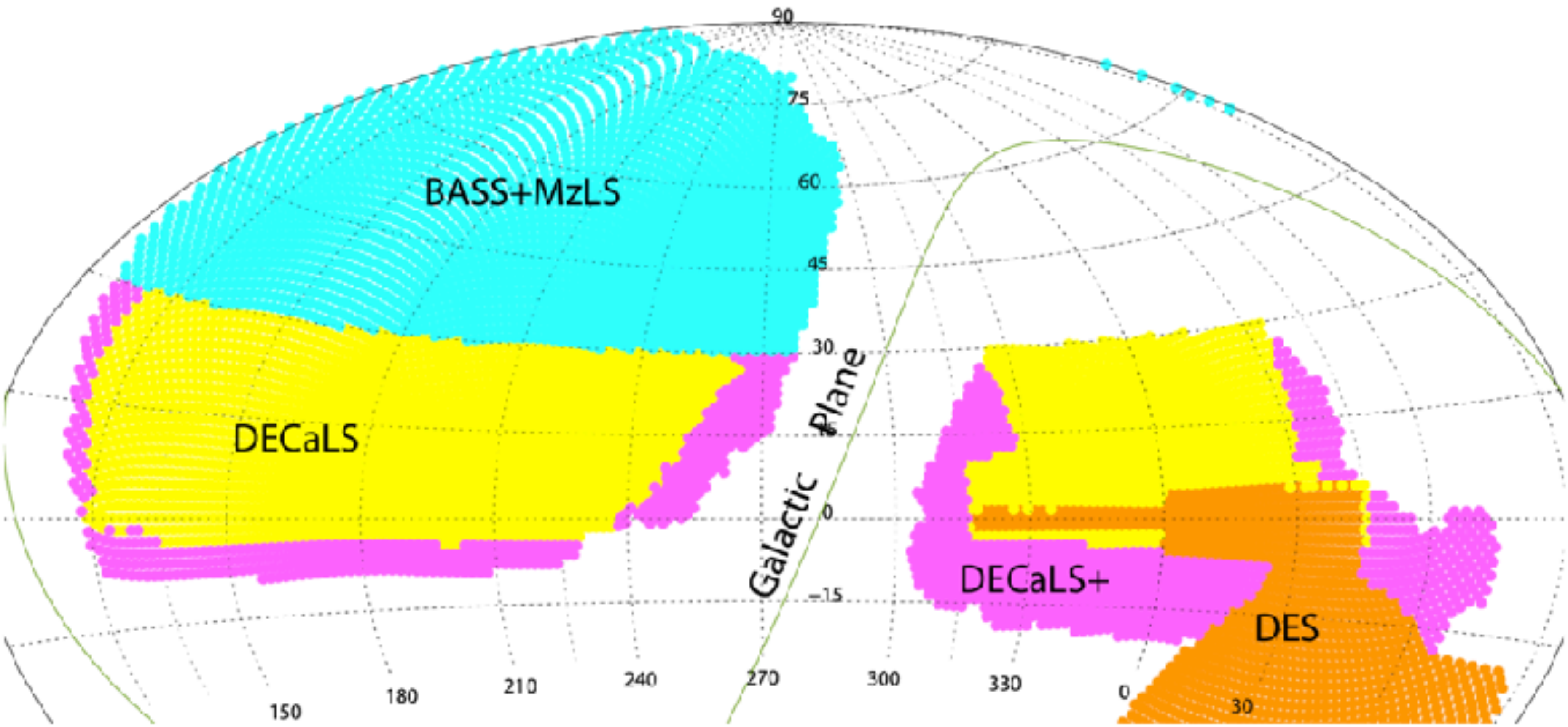
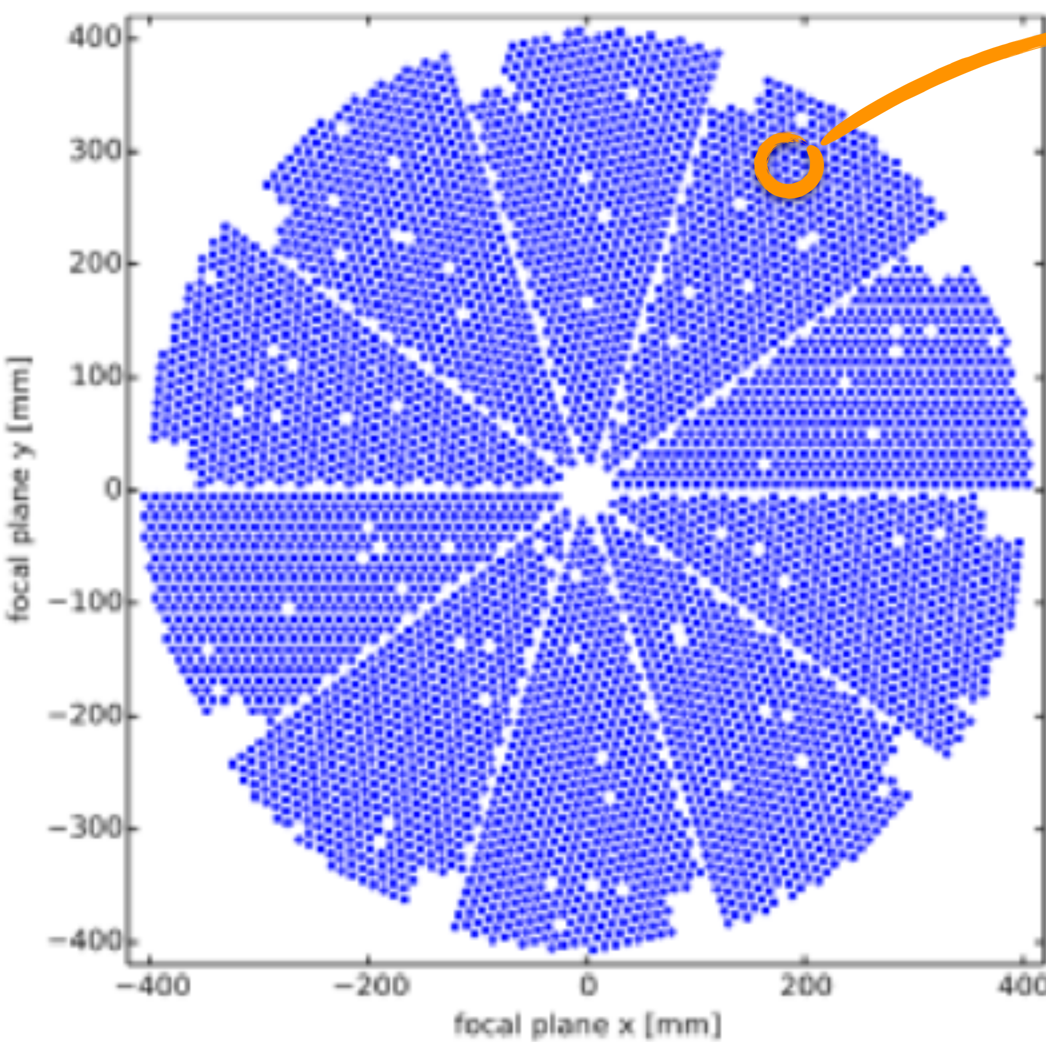


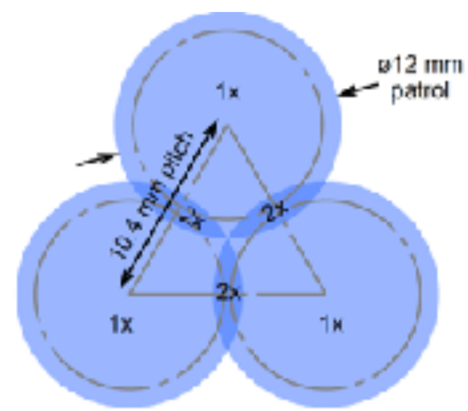

**Figure 3.19:** The primary imaging surveys that will result in targeting data for the DESI project. The footprint at DEC $\leq +34°$ will be covered using the Dark Energy Camera (DECam) on the Blanco 4m telescope at Cerro Tololo Inter-American Observatory. The Dark Energy Camera Legacy Survey (DECaLS, in yellow), the Dark Energy Survey (DES, in orange), and the extended DECaLS in the North Galactic Cap (DECaLS+, in purple on left) are underway. A proposal for the remaining extended DECaLS in the South Galactic Cap (DECaLS+, in purple on right) will be submitted. Imaging of the North Galactic Cap region at DEC $\geq +34°$ (cyan) will be covered with the 90Prime camera at the Bok 2.3-m telescope in $g-$ and $r-$bands (BASS: the Beijing-Arizona Sky Survey) and with the upgraded MOSAIC-3 camera on the Mayall 4m telescope in $z$-band (MzLS: the MOSAIC z-band Legacy Survey). Both the Bok and Mayall telescopes are located on Kitt Peak National Observatory.

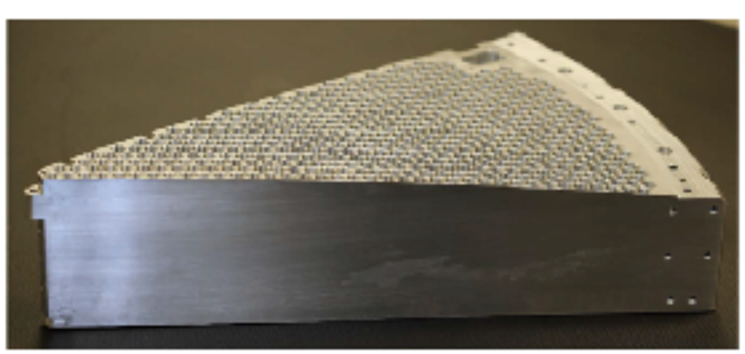
DESI collaboration 2016

# Fiber assignment and missing observations

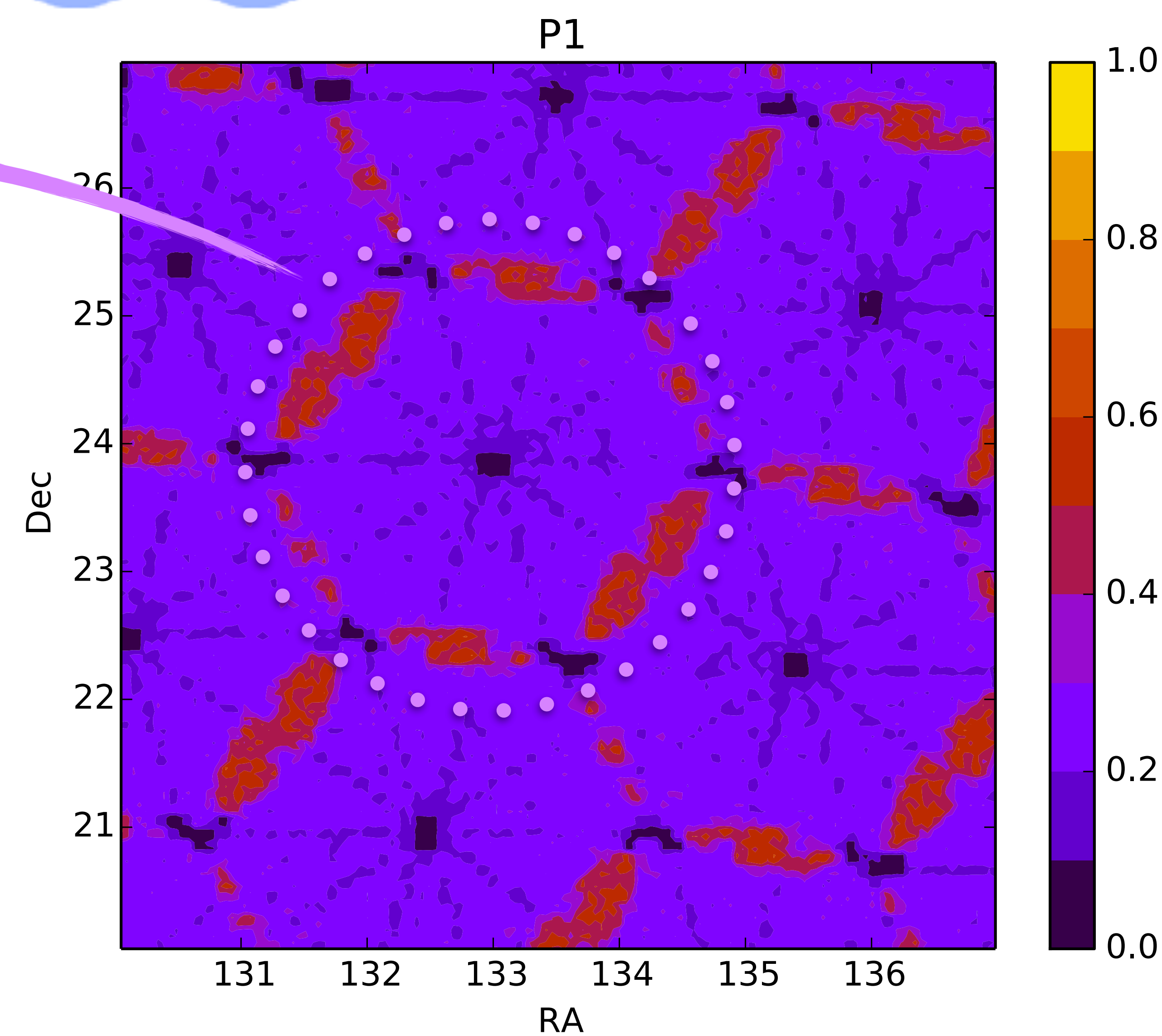
5,000 fibers    focal plane area = 7.5 $deg^2$



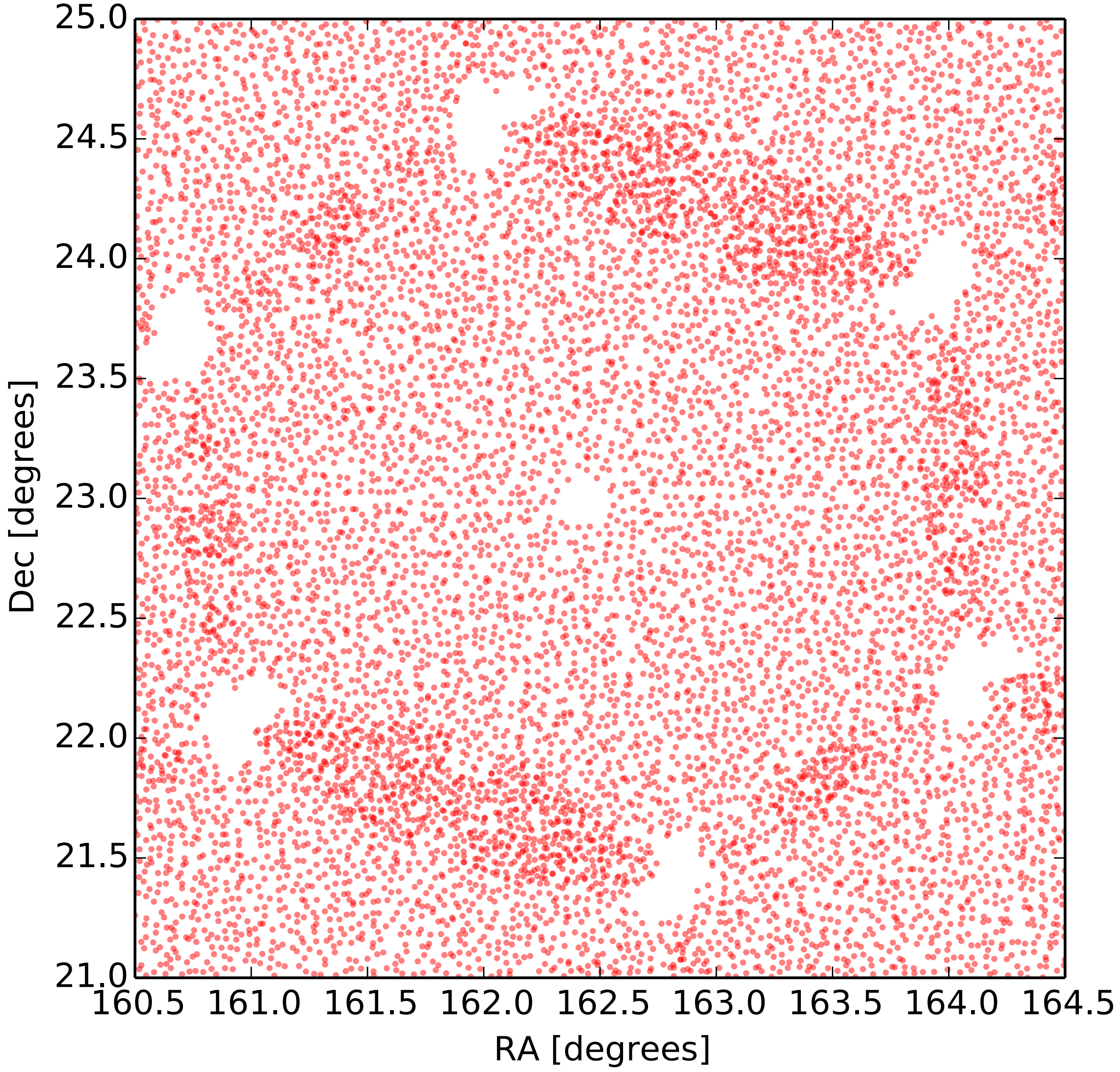



completeness after 1 pass of the DESI instrument



# DESI observing strategy

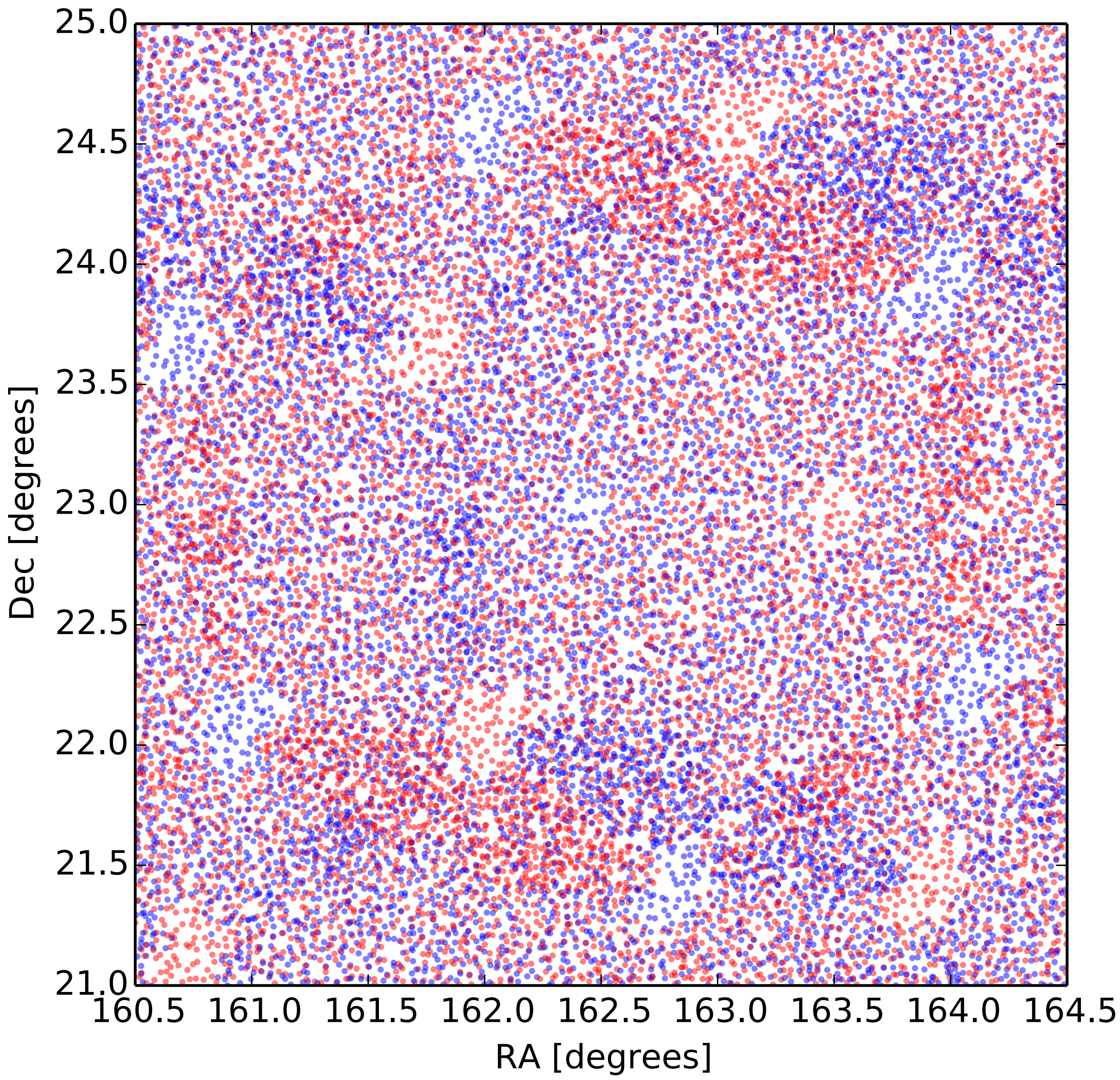
*pass 1*



# DESI observing strategy

*pass 1*

*pass 2*

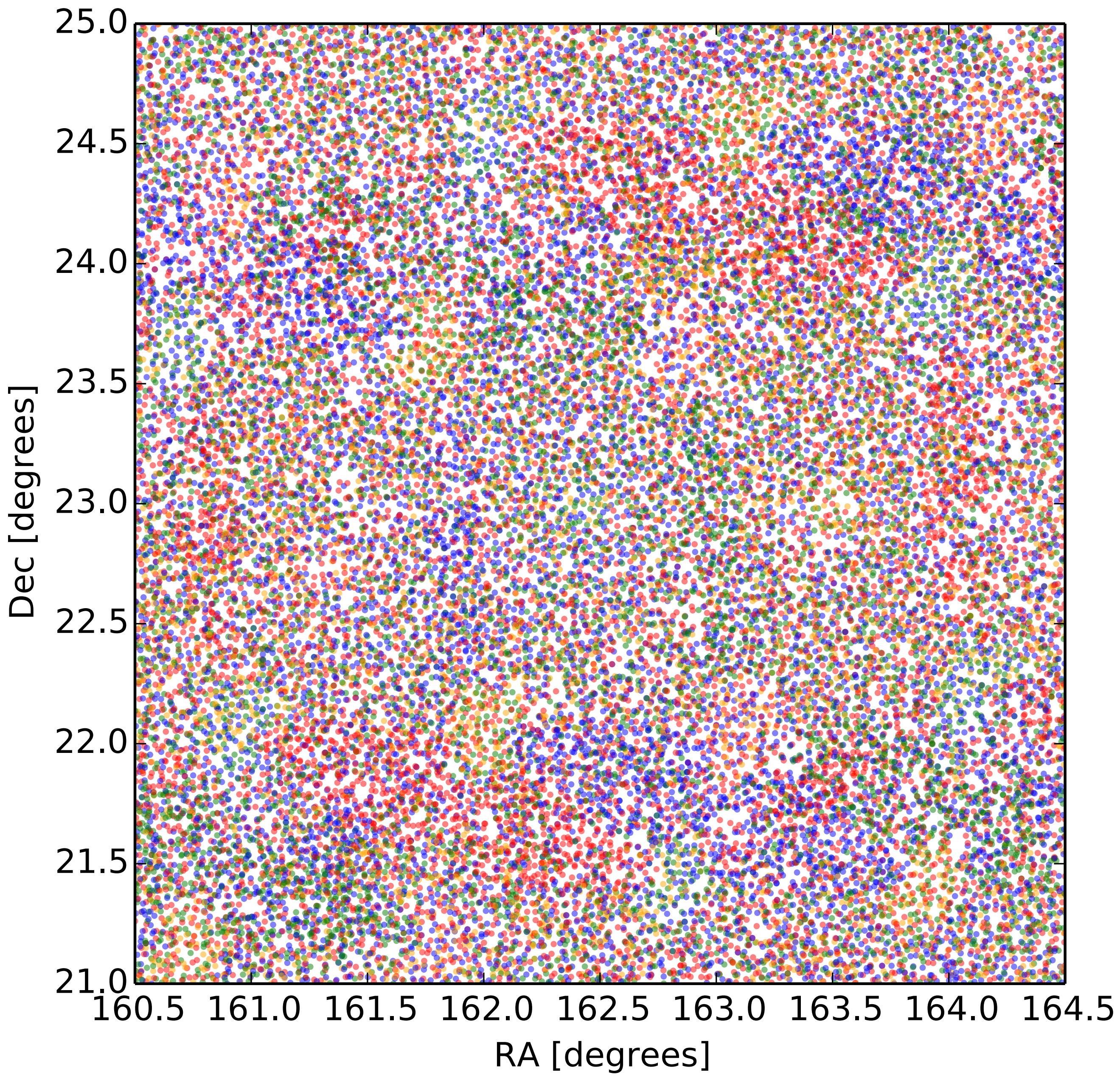


# DESI observing strategy

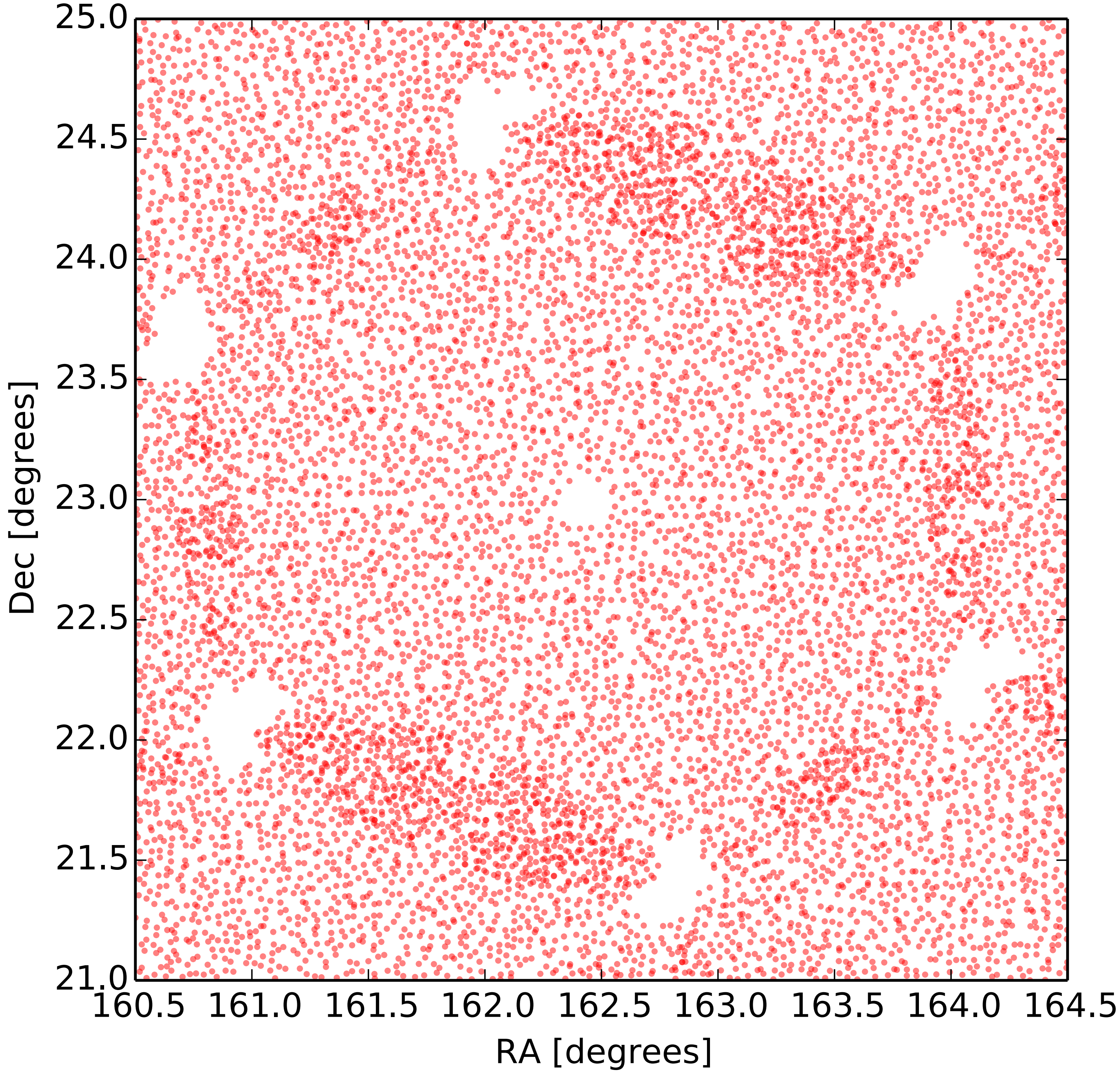
*pass 1*

*pass 2*

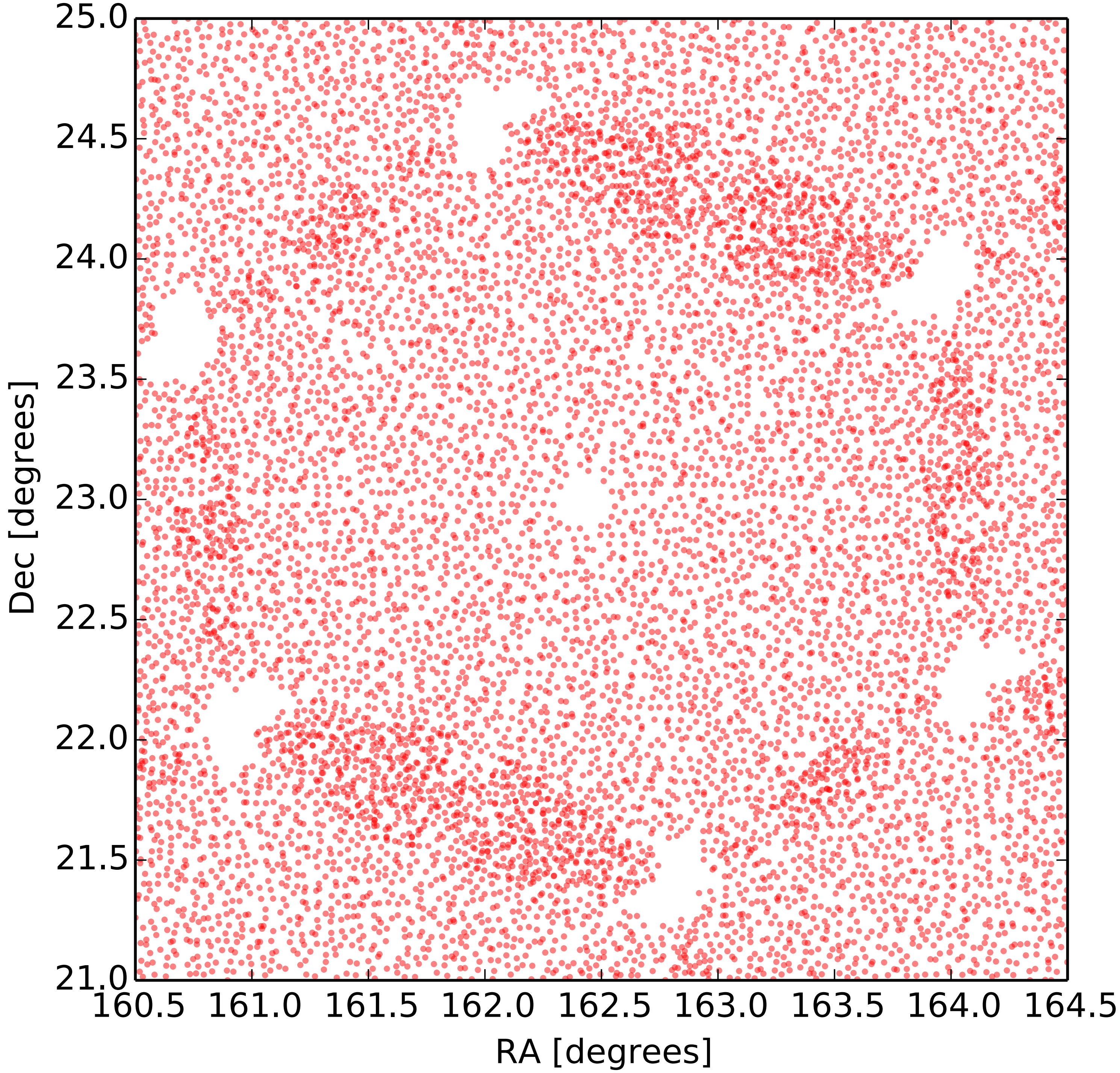
*etc…*



# Targeting realisation 1 (pass 1)



# Targeting realisation 2 (pass 1)



# 2 point correlation function

galaxy-galaxy pairs

galaxy-random pairs

$$\xi(\vec{s}) = \frac{DD(\vec{s})}{RR(\vec{s})} - 2\frac{DR(\vec{s})}{RR(\vec{s})} + 1$$

Landy & Szalay 1993

random-random pairs

# 2 point correlation function

galaxy-galaxy pairs

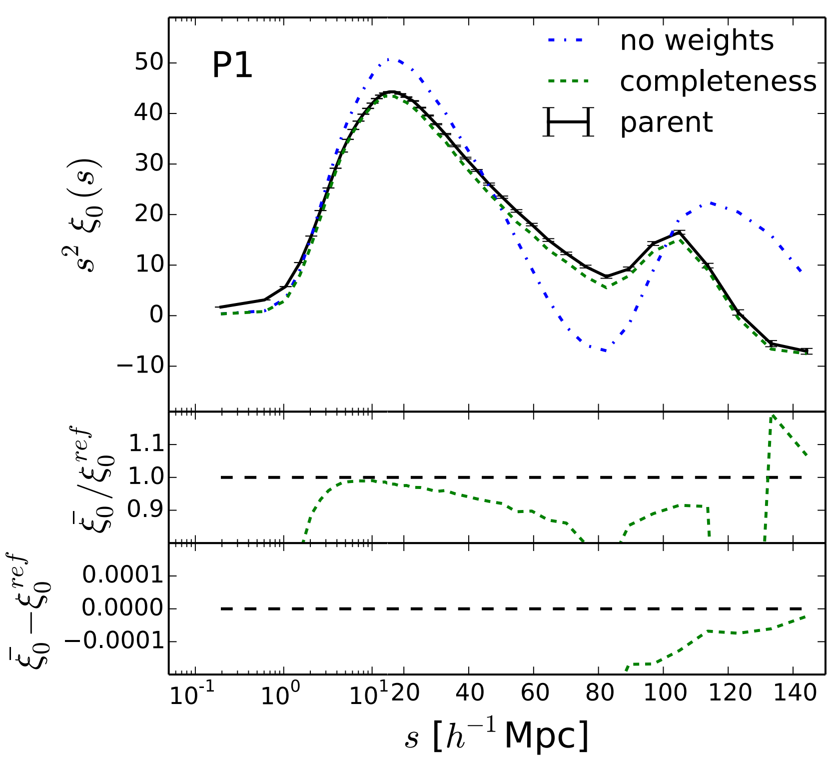
galaxy-random pairs

$$\xi(\vec{s}) = \frac{DD(\vec{s})}{RR(\vec{s})} - 2\frac{DR(\vec{s})}{RR(\vec{s})} + 1$$

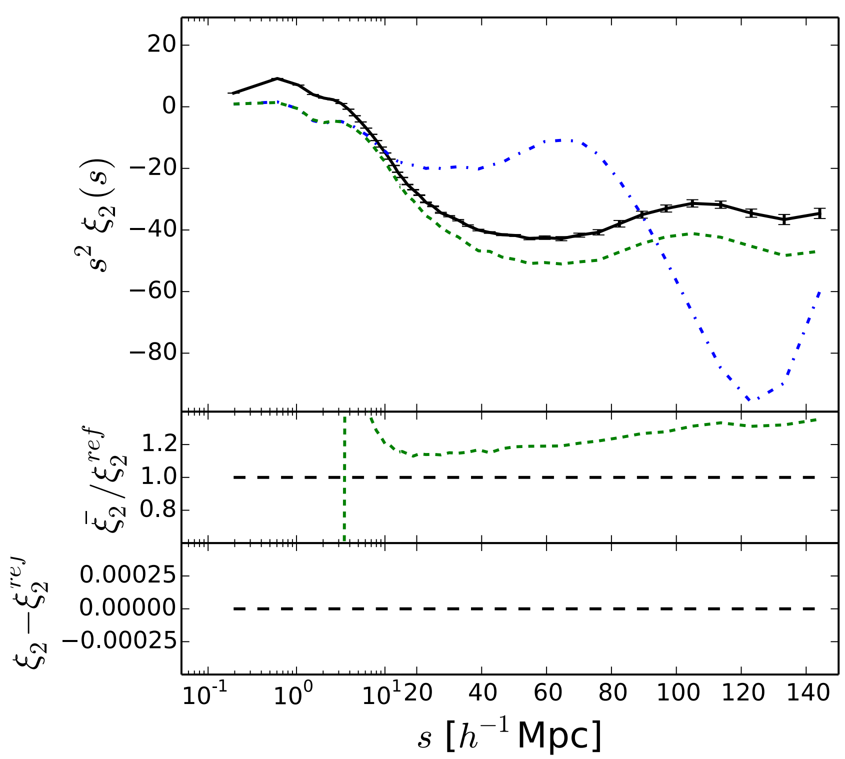
Landy & Szalay 1993

random-random pairs

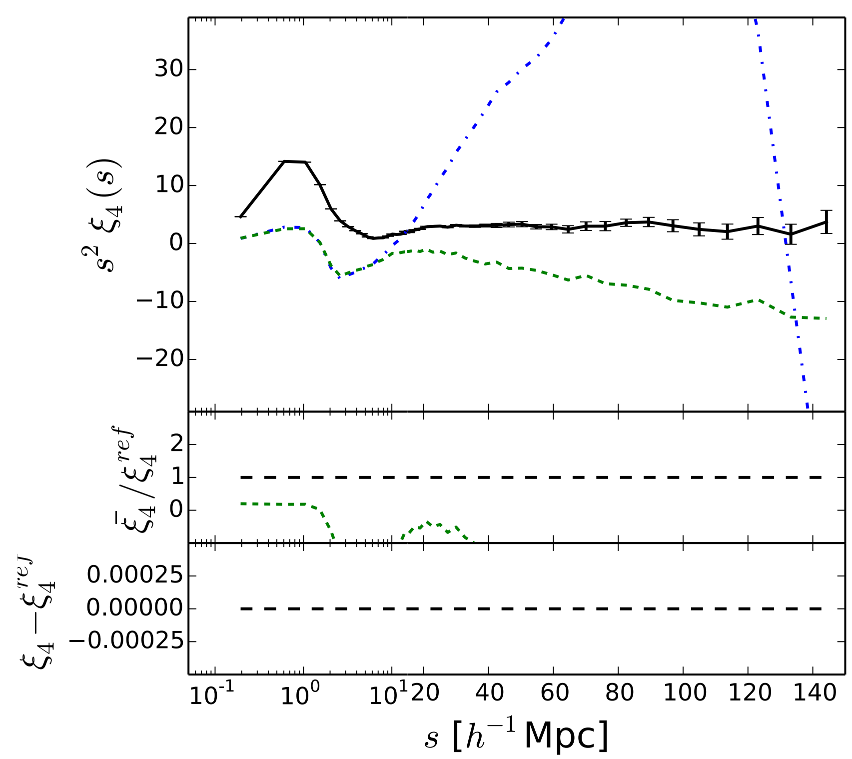
## Monopole




## Quadrupole




## Hexadecapole



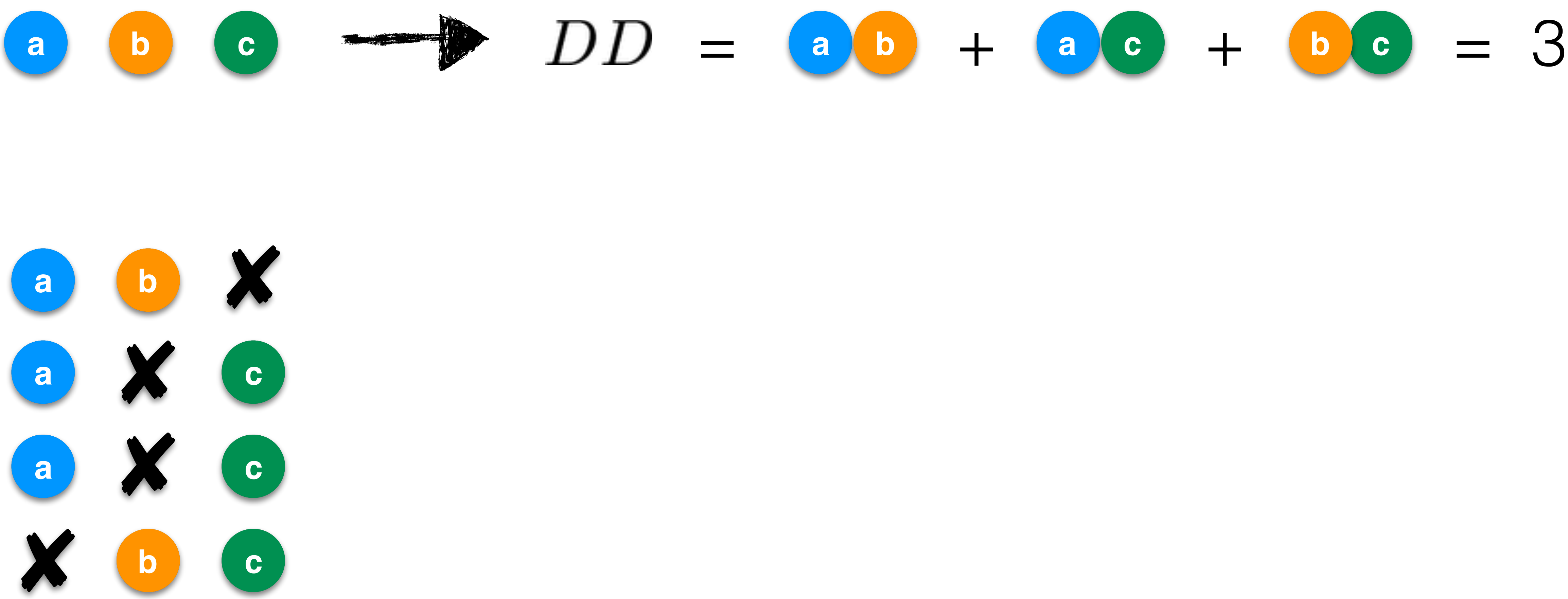
# Countermeasures for missing observations

- Agular upweighting - Hawkins et al. 2003
- Nearest neighbour - Anderson et al. 2012
- Correlation function decomposition - Guo et al. 2012
- Distribution function & top hat function - Hahn et al .2016
- Target sampling rate - De la Torre et al. 2013, Pezzotta et al. 2017
- Mode subtraction (configuration space) - Burden et al. 2017
- Mode subtraction (Fourier space) - Pinol et al. 2017, Hand et al. 2017
- Pairwise inverse probability (PIP) - Bianchi & Percival 2017, Percival & Bianchi 2017

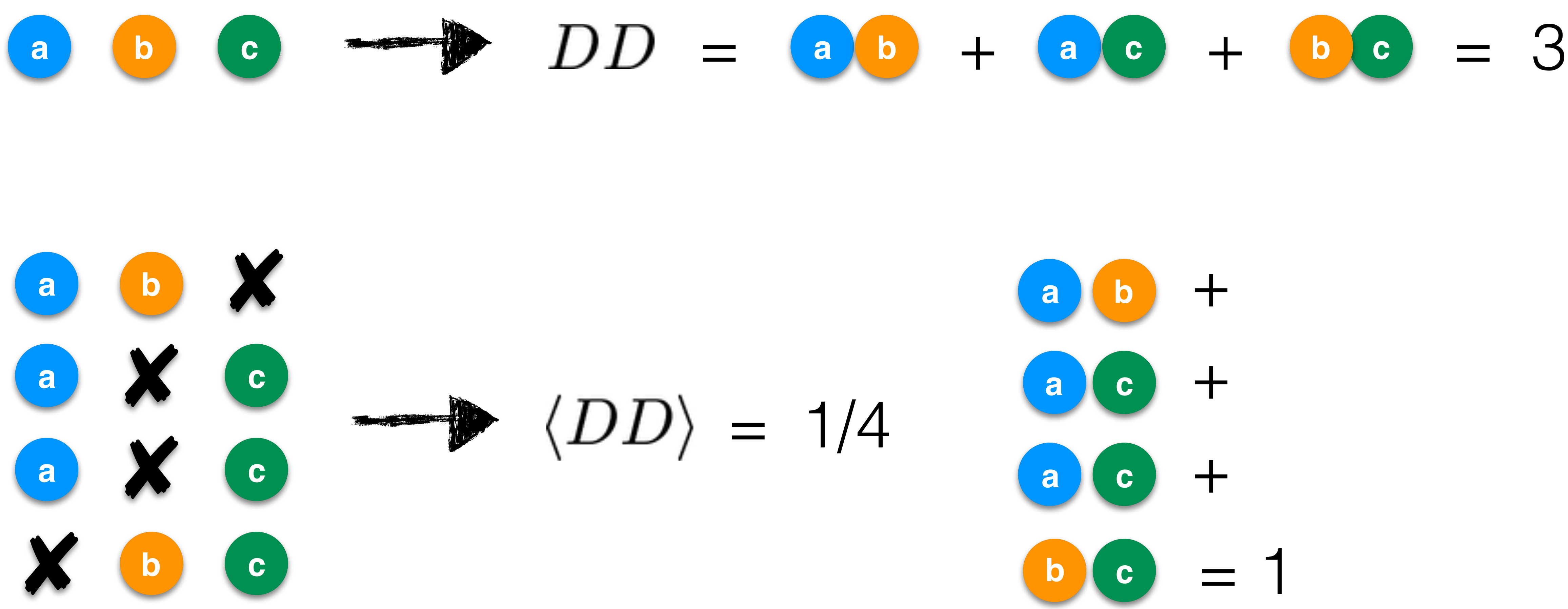
# Pairwise-Inverse-Probability (PIP) weights: basic idea

# Pairwise-Inverse-Probability (PIP) weights: basic idea

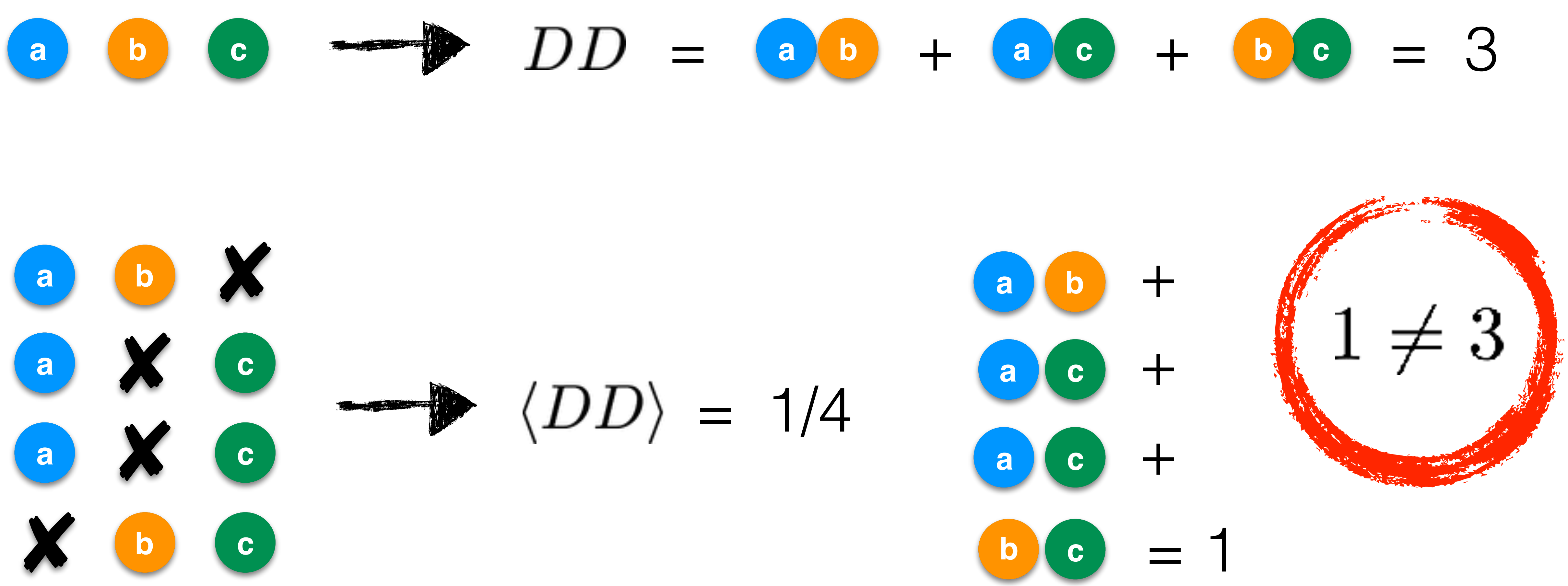
# Pairwise-Inverse-Probability (PIP) weights: basic idea



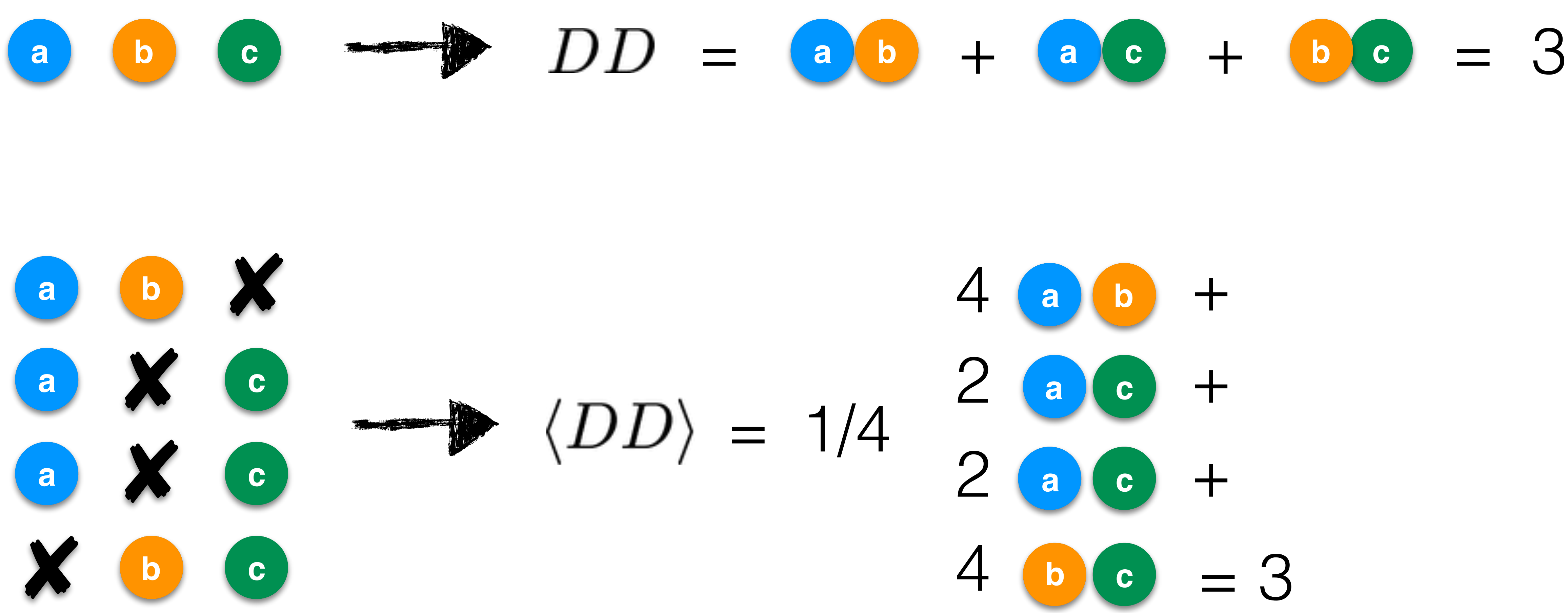
# Pairwise-Inverse-Probability (PIP) weights: basic idea



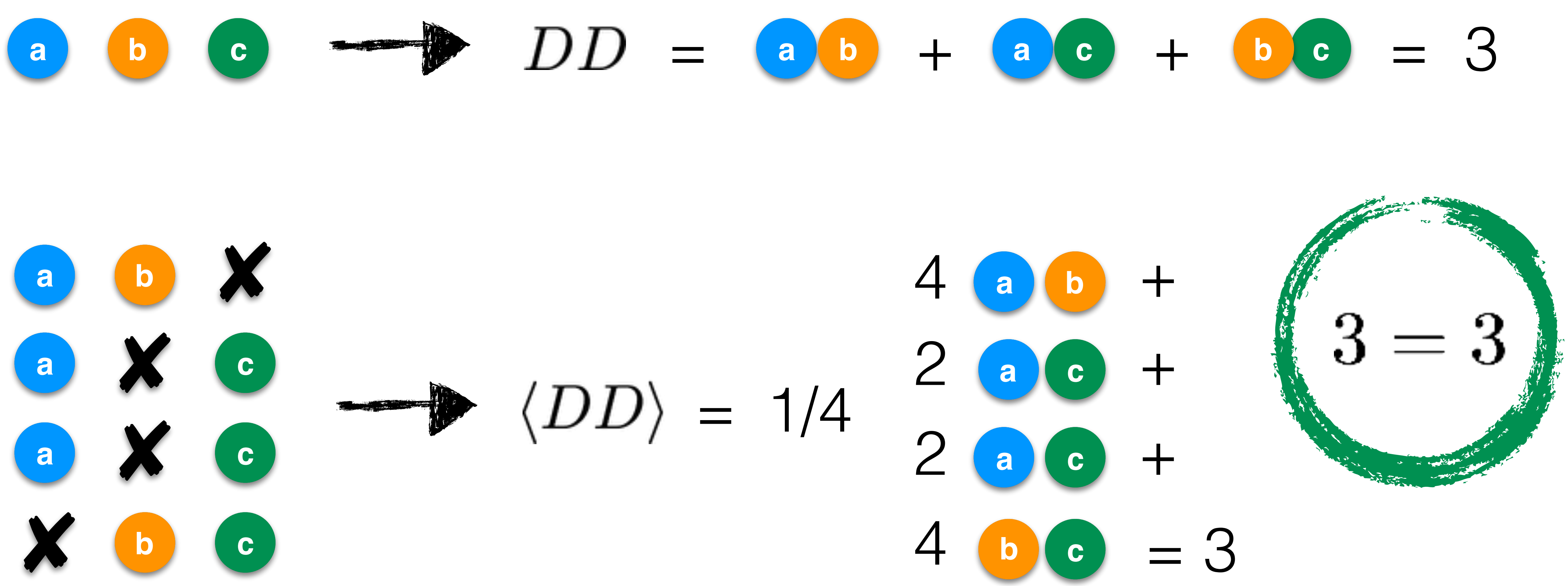
# Pairwise-Inverse-Probability (PIP) weights: basic idea



# Pairwise-Inverse-Probability (PIP) weights: basic idea



# Pairwise-Inverse-Probability (PIP) weights: basic idea



# Pairwise-Inverse-Probability (PIP) weights: correlation-function estimator

$$\xi(\vec{s}) = \frac{DD(\vec{s})}{RR(\vec{s})} - 2\frac{DR(\vec{s})}{RR(\vec{s})} + 1$$

Landy & Szalay 1993

# Pairwise-Inverse-Probability (PIP) weights: correlation-function estimator

$$\xi(\vec{s}) = \frac{DD(\vec{s})}{RR(\vec{s})} - 2\frac{DR(\vec{s})}{RR(\vec{s})} + 1$$

Landy & Szalay 1993

$$DD(\vec{s}) = \sum_{\vec{x}_m - \vec{x}_n \approx \vec{s}} w_{mn} \frac{DD_a^{(p)}(\theta)}{DD_a(\theta)}$$

$$DD_a(\theta) = \sum_{\vec{u}_m \cdot \vec{u}_n \approx \cos(\theta)} w_{mn}$$

$$w_{mn} = \frac{1}{p_{mn}} = \frac{1}{p_m p_n (1 + c_{mn})}$$

# Pairwise-Inverse-Probability (PIP) weights: correlation-function estimator

$$\xi(\vec{s}) = \frac{DD(\vec{s})}{RR(\vec{s})} - 2\frac{DR(\vec{s})}{RR(\vec{s})} + 1$$

Landy & Szalay 1993

$$DD(\vec{s}) = \sum_{\vec{x}_m - \vec{x}_n \approx \vec{s}} w_{mn} \frac{DD_a^{(p)}(\theta)}{DD_a(\theta)}$$

$$DD_a(\theta) = \sum_{\vec{u}_m \cdot \vec{u}_n \approx \cos(\theta)} w_{mn}$$

$$w_{mn} = \frac{1}{p_{mn}} = \frac{1}{p_m p_n (1 + c_{mn})}$$

$$DR(\vec{s}) = \sum_{\vec{x}_m - \vec{y}_n \approx \vec{s}} w_m \frac{DR_a^{(p)}(\theta)}{DR_a(\theta)}$$

$$DR_a(\theta) = \sum_{\vec{u}_m \cdot \vec{v}_n \approx \cos(\theta)} w_m$$

$$w_m = \frac{1}{p_m}$$

# Practical implementation: bitwise weights

$$N_{bits} = 8$$

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| m | ✓ | ✓ | ✗ | ✓ | ✗ | ✗ | ✓ | ✗ |
| | | | | | | | | |

# Practical implementation: bitwise weights

$$N_{bits} = 8$$

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| m | ✓ | ✓ | ✗ | ✓ | ✗ | ✗ | ✓ | ✗ |
| | | | | | | | | |

$$w_m^{(b)} = \{1, 1, 0, 1, 0, 0, 1, 0\}$$

$$\text{popcnt}\left[w_m^{(b)}\right] = 4$$

$$p_m = \frac{4}{N_{bits}} = \frac{1}{2}$$

$$w_m = \frac{1}{p_m} = 2$$

# Practical implementation: bitwise weights

$$N_{bits} = 8$$

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| m | ✓ | ✓ | ✗ | ✓ | ✗ | ✗ | ✓ | ✗ |
| n | ✗ | ✓ | ✓ | ✗ | ✓ | ✓ | ✓ | ✓ |

$$w_m^{(b)} = \{1, 1, 0, 1, 0, 0, 1, 0\}$$

$$\text{popcnt}\left[w_m^{(b)}\right] = 4$$

$$p_m = \frac{4}{N_{bits}} = \frac{1}{2}$$

$$w_m = \frac{1}{p_m} = 2$$

$$w_n^{(b)} = \{0, 1, 1, 0, 1, 1, 1, 1\}$$

$$\text{popcnt}\left[w_n^{(b)}\right] = 6$$

$$p_m = \frac{6}{N_{bits}} = \frac{3}{4}$$

$$w_n = \frac{1}{p_n} = \frac{4}{3}$$

# Practical implementation: bitwise weights

$$N_{bits} = 8$$

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| m | ✓ | ✓ | ✗ | ✓ | ✗ | ✗ | ✓ | ✗ |
| n | ✗ | ✓ | ✓ | ✗ | ✓ | ✓ | ✓ | ✓ |

$$w_m^{(b)} = \{1, 1, 0, 1, 0, 0, 1, 0\}$$

$$\text{popcnt}\left[w_m^{(b)}\right] = 4$$

$$p_m = \frac{4}{N_{bits}} = \frac{1}{2}$$

$$w_m = \frac{1}{p_m} = 2$$

$$w_n^{(b)} = \{0, 1, 1, 0, 1, 1, 1, 1\}$$

$$\text{popcnt}\left[w_n^{(b)}\right] = 6$$

$$p_m = \frac{6}{N_{bits}} = \frac{3}{4}$$

$$w_n = \frac{1}{p_n} = \frac{4}{3}$$

$$w_{mn} = w_m w_n = \frac{8}{3}$$

# Practical implementation: bitwise weights

$N_{bits} = 8$

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| m | ✓ | ✓ | ✗ | ✓ | ✗ | ✗ | ✓ | ✗ |
| n | ✗ | ✓ | ✓ | ✗ | ✓ | ✓ | ✓ | ✓ |

$$w_m^{(b)} = \{1, 1, 0, 1, 0, 0, 1, 0\}$$

$$w_n^{(b)} = \{0, 1, 1, 0, 1, 1, 1, 1\}$$

# Practical implementation: bitwise weights

$$N_{bits} = 8$$

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| m | ✓ | ✓ | ✗ | ✓ | ✗ | ✗ | ✓ | ✗ |
| n | ✗ | ✓ | ✓ | ✗ | ✓ | ✓ | ✓ | ✓ |

$$w_m^{(b)} = \{1, 1, 0, 1, 0, 0, 1, 0\}$$

$$w_n^{(b)} = \{0, 1, 1, 0, 1, 1, 1, 1\}$$

$$w_{mn}^{(b)} \equiv w_m^{(b)} \text{ and } w_n^{(b)} = \{0, 1, 0, 0, 0, 0, 1, 0\}$$

$$\text{popcnt}\left[w_{mn}^{(b)}\right] = 2 \qquad p_{mn} = \frac{2}{N_{bits}} = \frac{1}{4}$$

# Practical implementation: bitwise weights

$$N_{bits} = 8$$

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| m | ✓ | ✓ | ✗ | ✓ | ✗ | ✗ | ✓ | ✗ |
| n | ✗ | ✓ | ✓ | ✗ | ✓ | ✓ | ✓ | ✓ |

$$w_m^{(b)} = \{1, 1, 0, 1, 0, 0, 1, 0\}$$

$$w_n^{(b)} = \{0, 1, 1, 0, 1, 1, 1, 1\}$$

$$w_{mn}^{(b)} \equiv w_m^{(b)} \text{ and } w_n^{(b)} = \{0, 1, 0, 0, 0, 0, 1, 0\}$$

$$\text{popcnt}\left[w_{mn}^{(b)}\right] = 2 \qquad p_{mn} = \frac{2}{N_{bits}} = \frac{1}{4}$$

$$w_{mn} = \frac{1}{p_{mn}} = 4$$

# Practical implementation: summary

The pair weights are computed (as usual) while doing pair counts via a simple function of individual weights

*Standard approach*

$$w_{mn} = w_m w_n$$

*Bitwise weights*

$$w_{mn} = \frac{N_{bits}}{\mathrm{popcnt}\left[w_m^{(b)} \text{ and } w_n^{(b)}\right]}$$

# “Why angular upweighting?”

- Natural way to **minimise the variance**, see W. J. Percival & **DB** - arXiv:1703.02071
- Allows to relax the assumption that **no pair has zero probability of being observed** (**not true**, e.g., for BOSS and DESI) to the weaker assumption that **all the zero-probability pairs have the same clustering properties of the observed pairs** (**true**, e.g., for BOSS and DESI)

# Algorithm

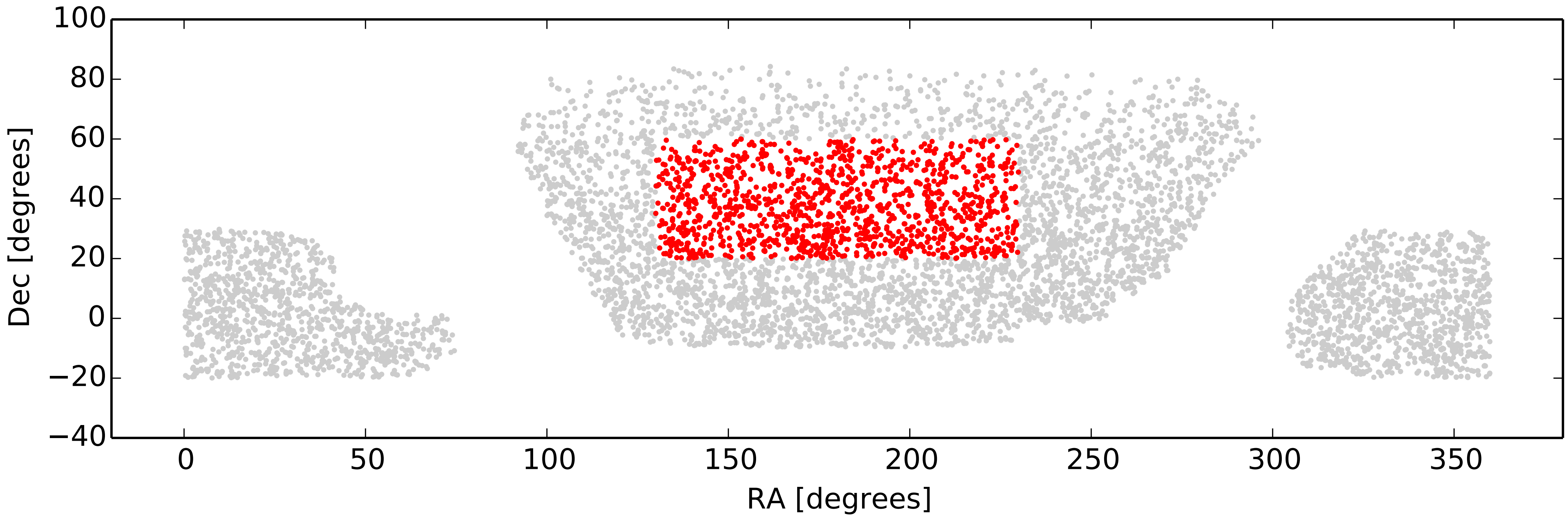
1. Evaluate PIP weights by running the targeting algorithm many times (N~1000?)

2. Compute the angular pair counts DD and DR of the full parent sample

3. Compute the angular pair counts DD and DR of the observed sample using PIP weights

4. Compute the 3D correlation function of the observed sample via PIP weights + ang. upweighting (i.e. ratios between 2 and 3)

# Comparison to mocks:
# DESI footprint vs our sample

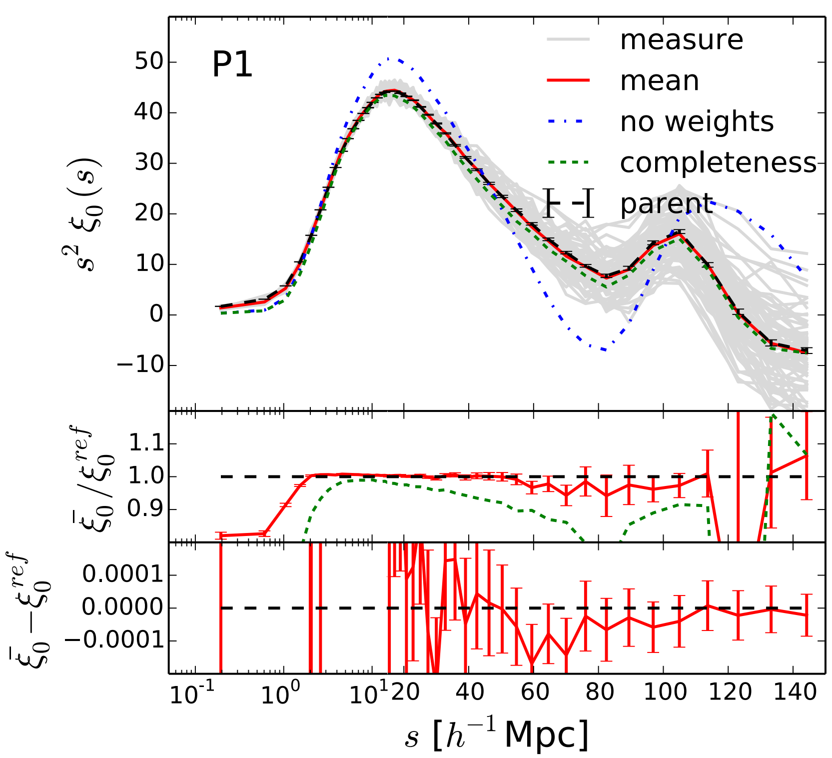



25 QPM mocks, see A. Burden, N. Padmanabhan Robert N. Cahn, M. J. White, L. Samushia 2016

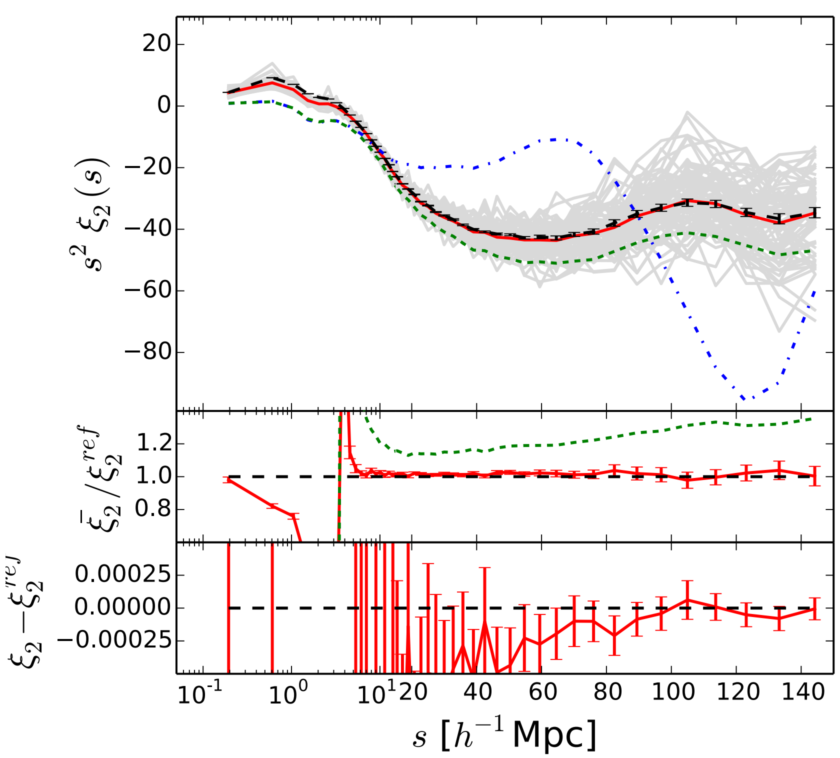
# Comparison to DESI mocks: pass 1

completeness = 25%

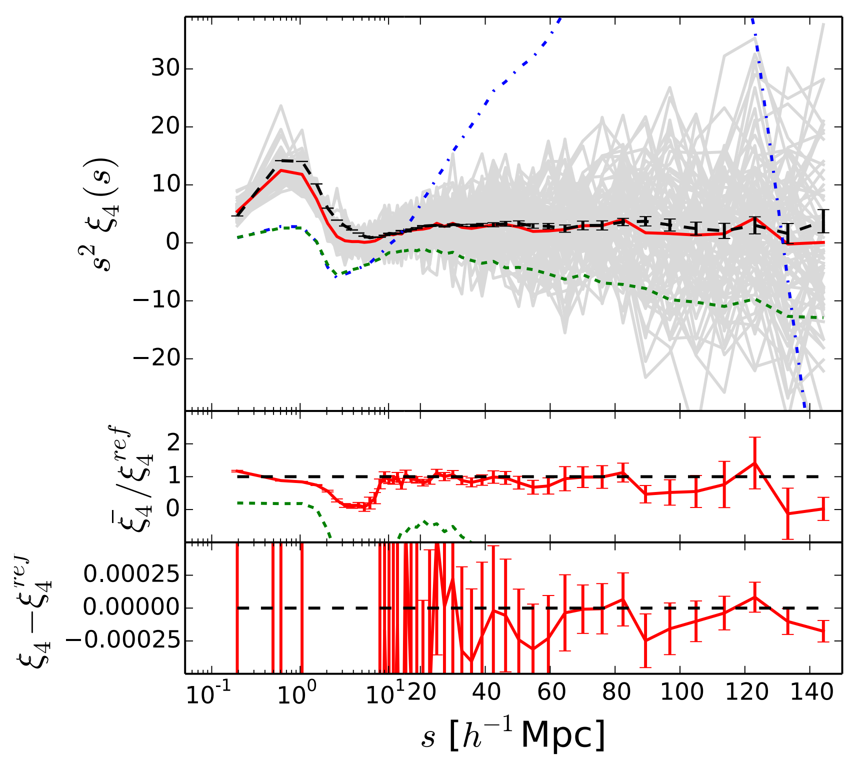
## Monopole




## Quadrupole



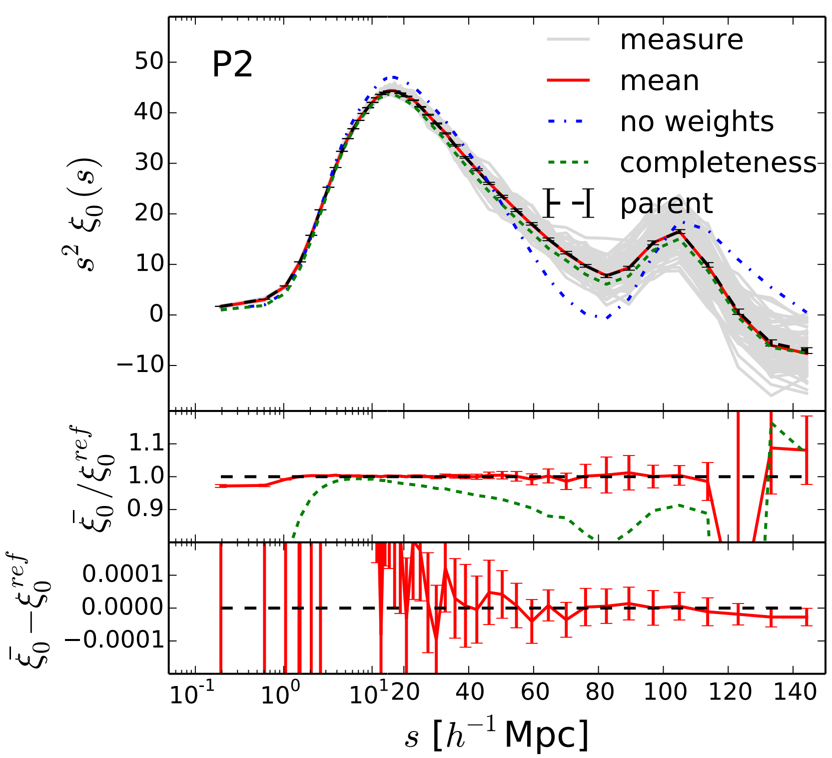

## Hexadecapole



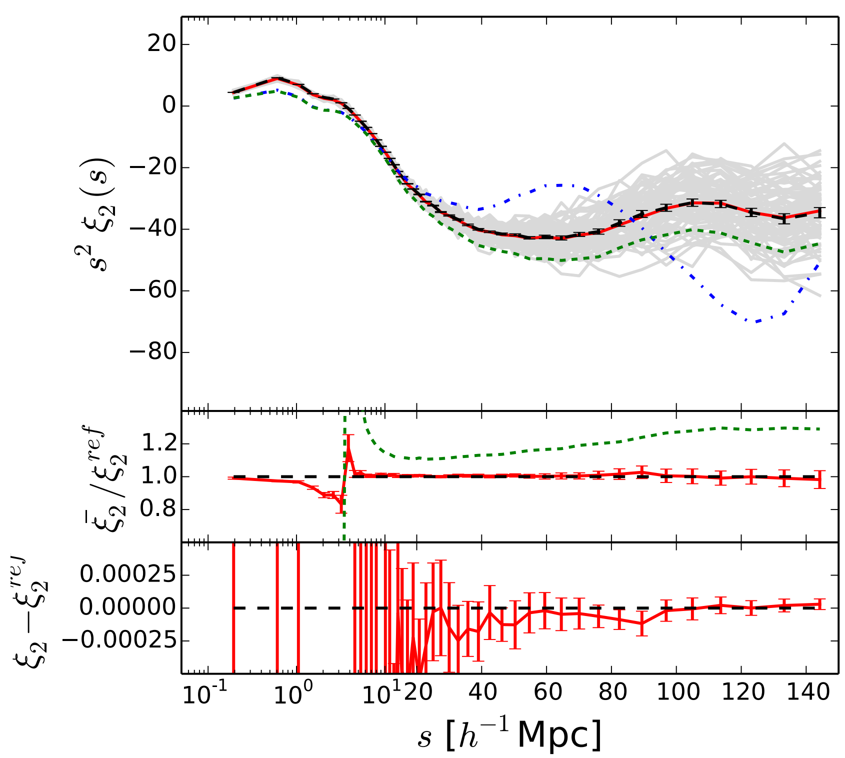
# Comparison to DESI mocks: pass 2

completeness = 48%

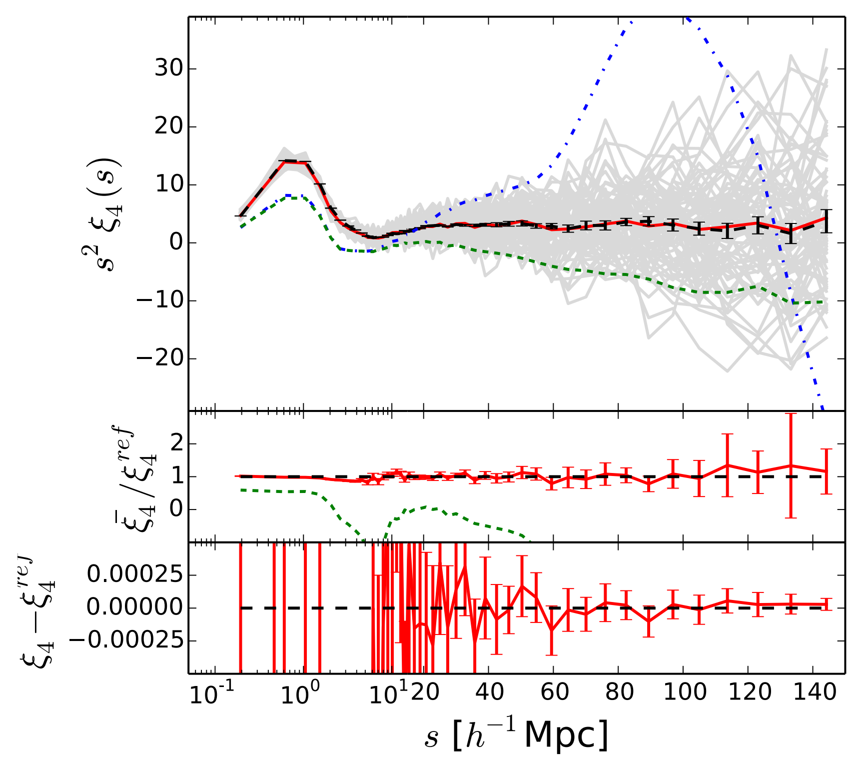
## Monopole




## Quadrupole



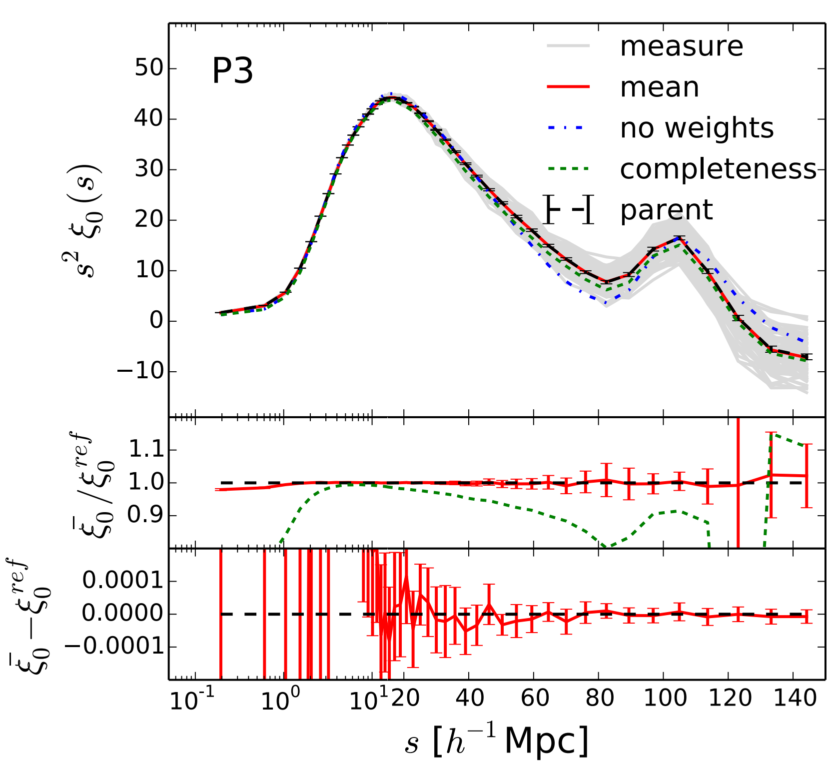

## Hexadecapole



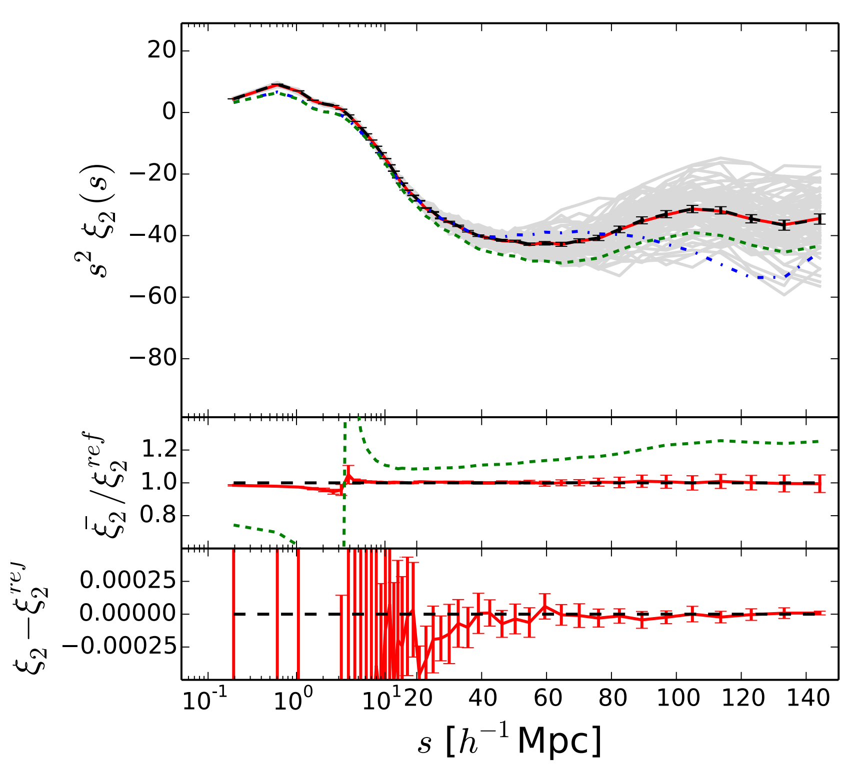
# Comparison to DESI mocks: pass 3

completeness = 67%

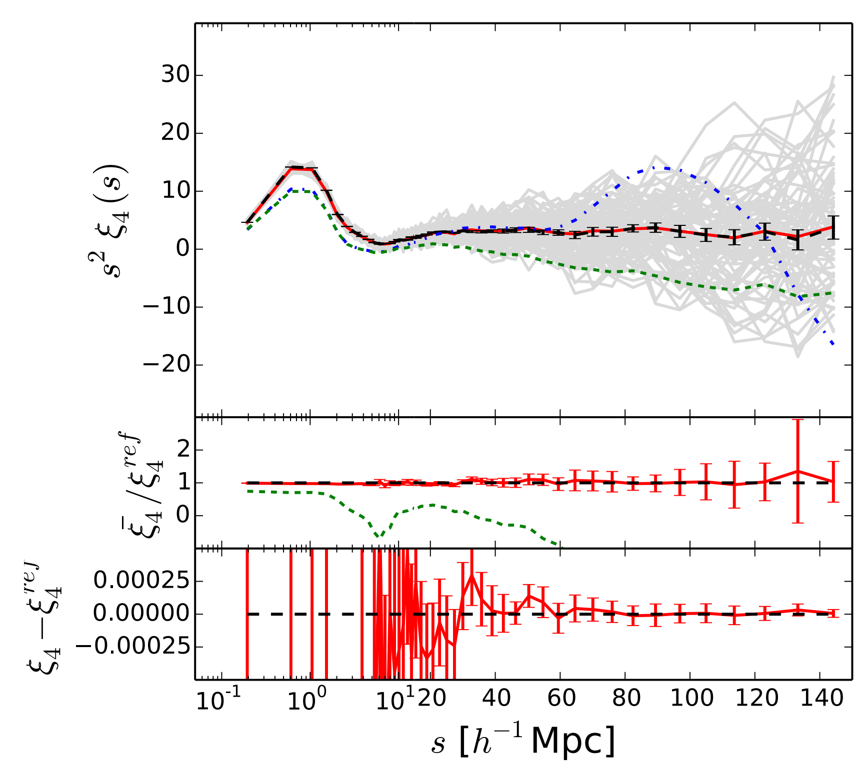
## Monopole




## Quadrupole



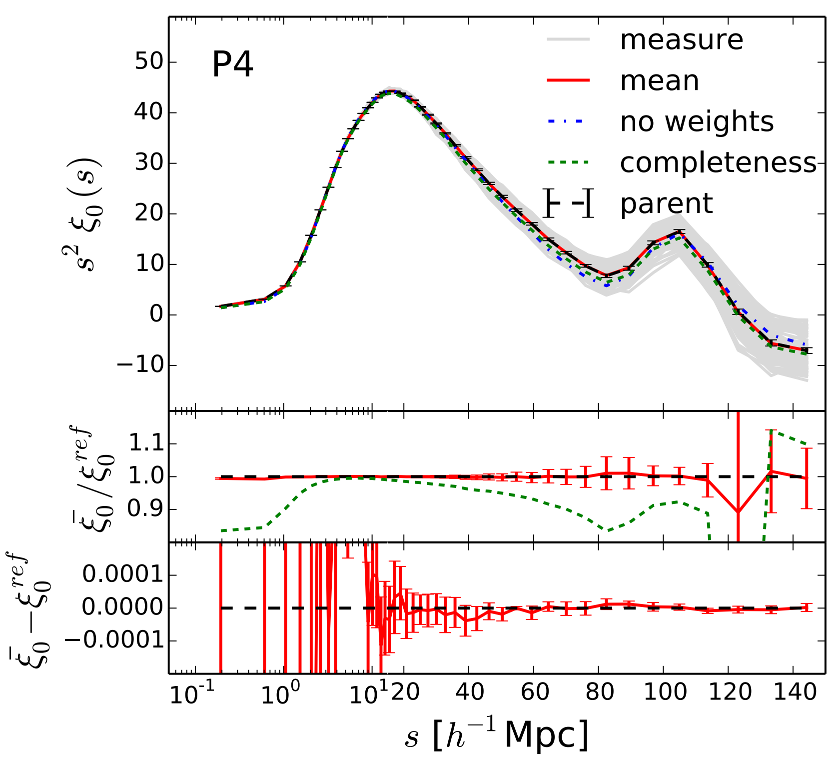

## Hexadecapole



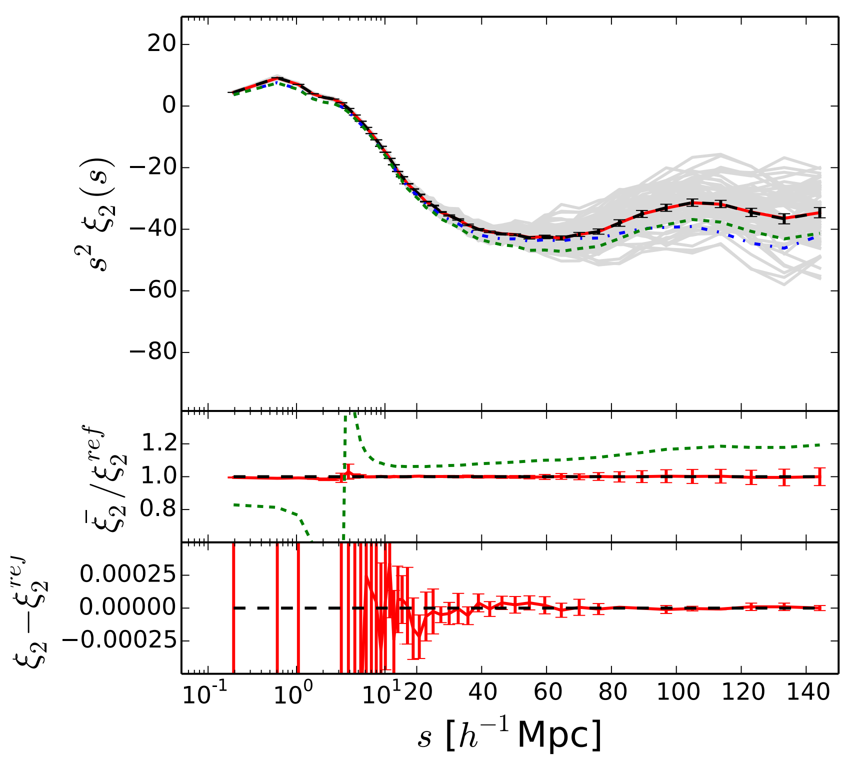
# Comparison to DESI mocks: pass 4

completeness = 81%

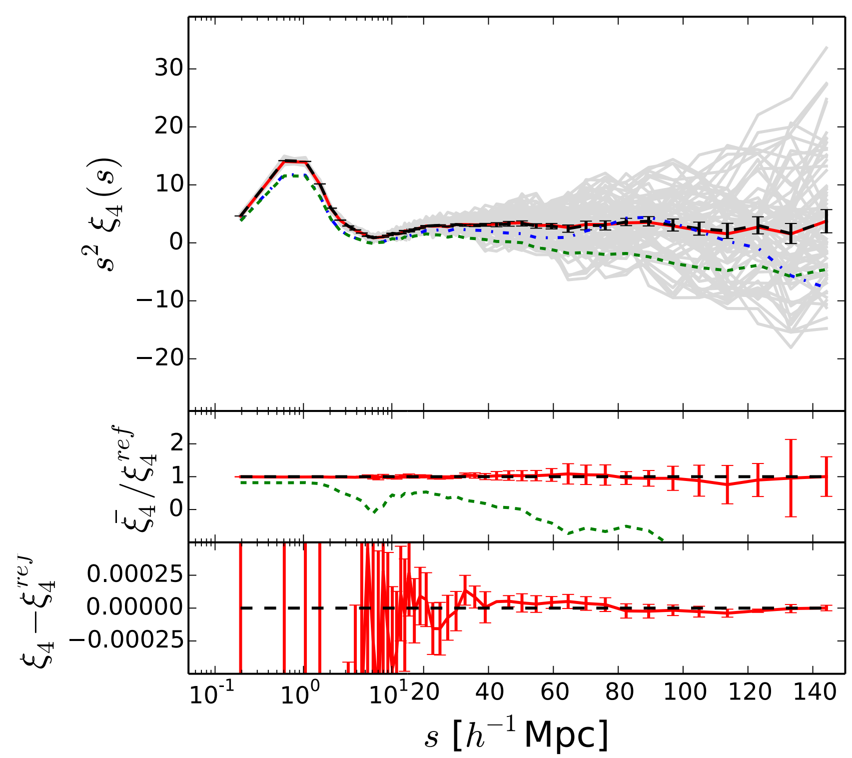
## Monopole




## Quadrupole




## Hexadecapole

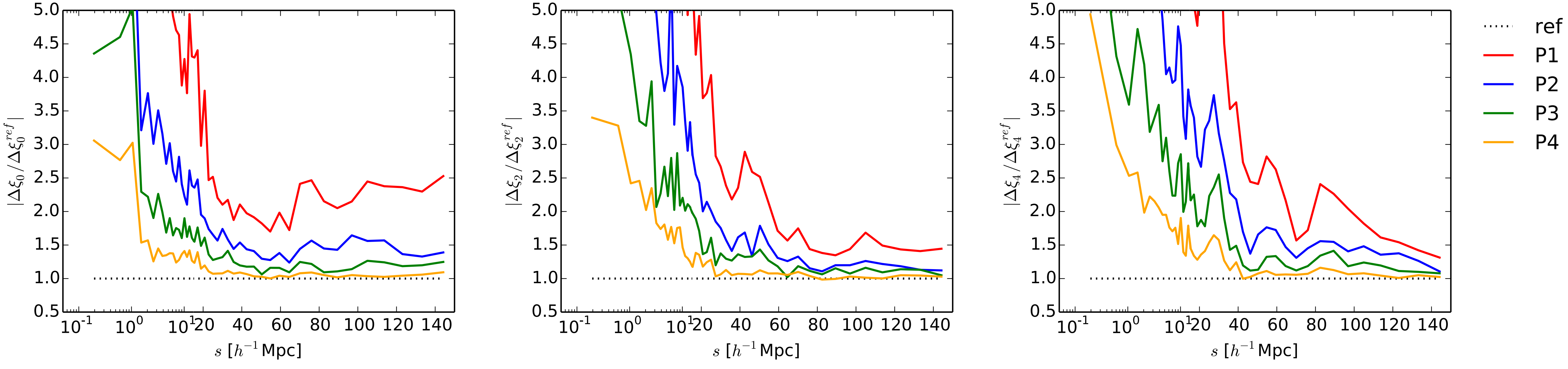


# Error vs cosmic variance

## Monopole

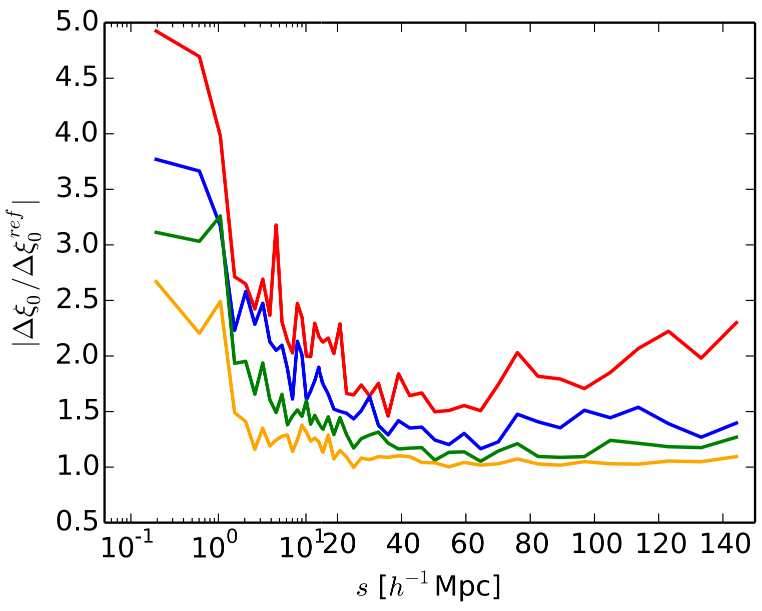
## Quadrupole

## Hexadecapole

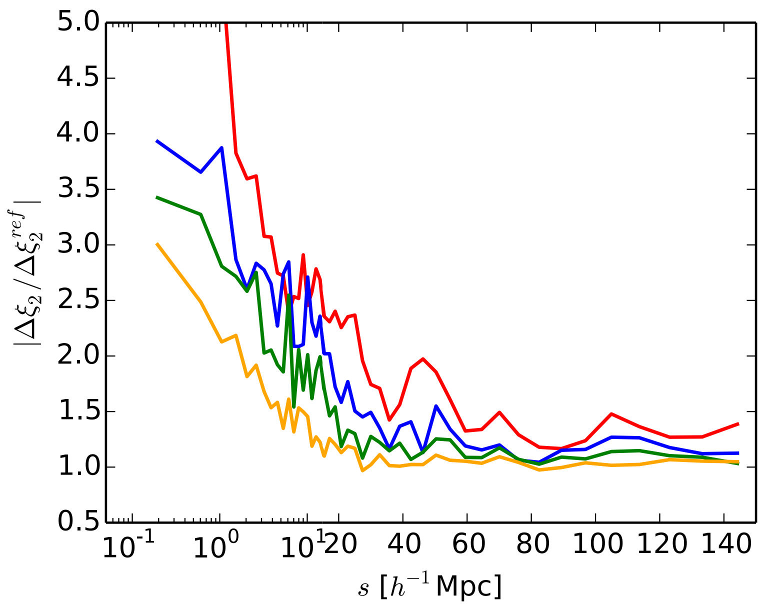


# Error vs uniform dilution

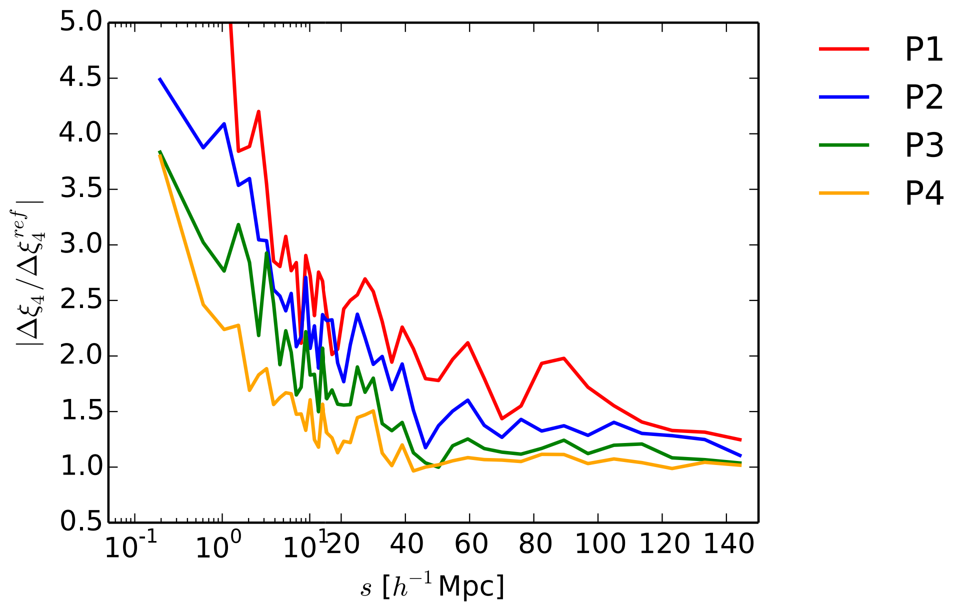
## Monopole




## Quadrupole




## Hexadecapole



# Summary

- We developed a theoretically-motivated weighting scheme capable to recover unbiased estimates of the galaxy clustering for a large class of targeting algorithm

- We provided a practical implementation of the algorithm

- We showed that this approach works for the DESI survey

(Main) Things to do:

- Fourier space
- Reconstruction