Euclid STAR Prize 2020 Timeline

Photometric Redshifts for Euclid

Stéphane Paltani

Lead of Photometric-Redshift Organization Unit

University of Geneva

Swiss Euclid Days February 4-5, 2020

UNIVERSITÉ DE GENÈVE

FACULTÉ DES SCIENCES

Département d'astronomie

Photometric-Redshift OU

- OU-PHZ is a very important Swiss responsibility in Euclid linked to Euclid's main science
  - About 50 contributors
- Main task is to determine redshifts for weak-lensing tomography
  - Redshifts need to be determined from photometry alone (no spectroscopy)
- OU-PHZ needs also to
  - Identify stars (and AGN) for PSF determination (and contamination)
  - Provide unbiased colors from 500 to 1000 nm for shape reconstruction
  - Compute galaxy physical parameters (Mass, SFR, SFH, reddening,...)

Photometric Redshifts

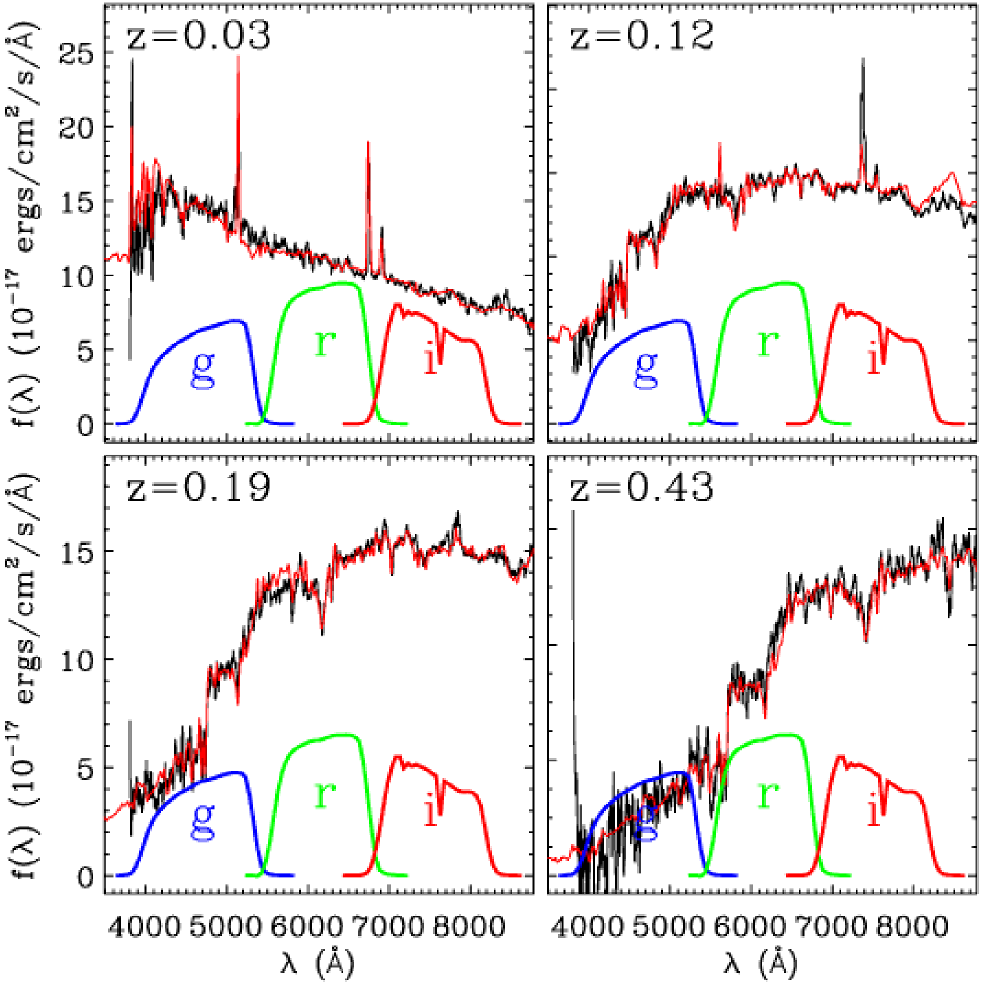
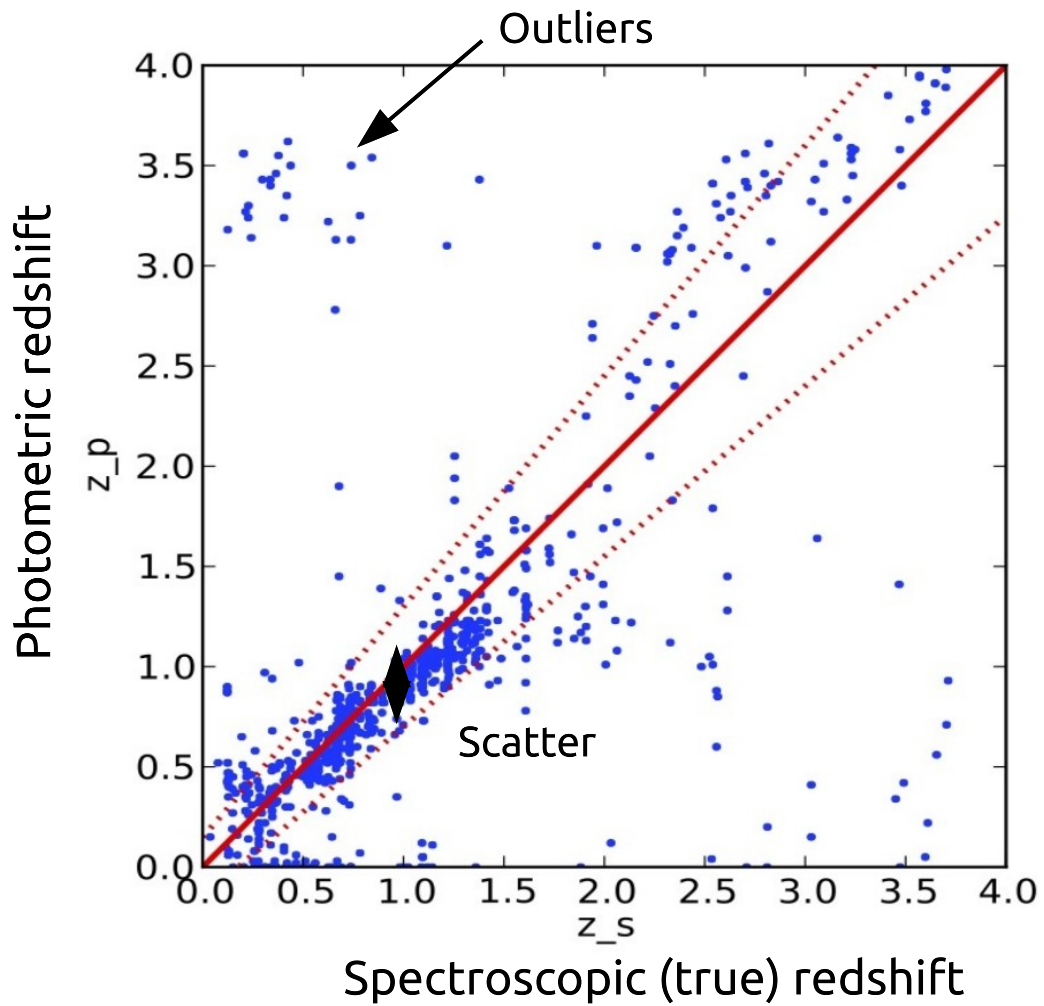


Photo-z vs True Redshifts



Euclid requirements :

Precision (scatter):
$\sigma_z < 0.05(1+z)$

Outlier fraction : < 10 %
beyond 0.15(1+z)

Accuracy (bias):
knowledge of $\bar{z}$ in any tomographic bin better than 0.002(1+z)

Produce PDF of photometric redshifts

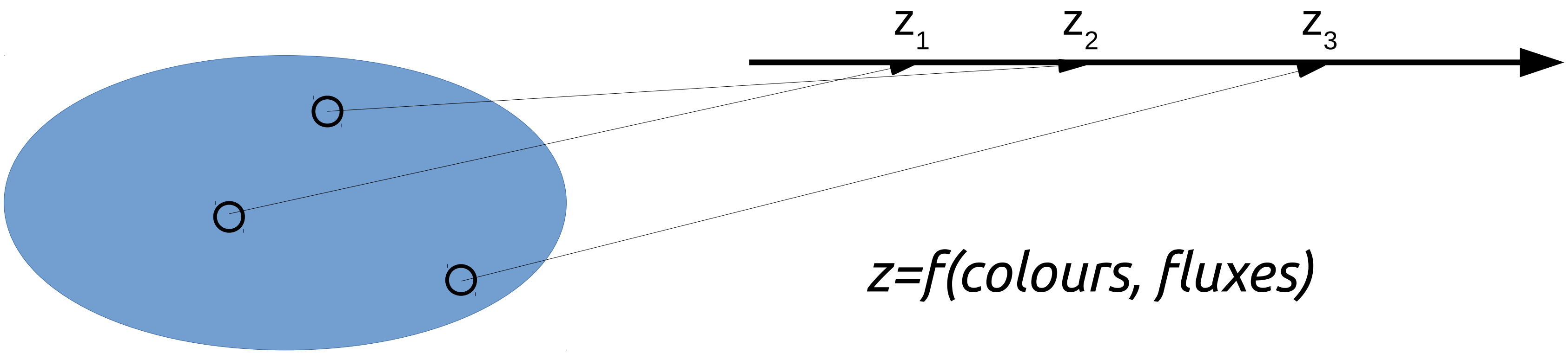
Challenges of Euclid Photo-z's

- Inhomogeneous photometry (different surveys, large sky area)
- Some regions of the sky with high Galactic reddening
- Very demanding requirement on photo-z accuracy (bias)
  - But moderate requirements on photo-z precision
- Huge number of objects makes the problem computationally hard
- Need for a fully statistical framework to propagate all uncertainties correctly into the cosmological parameters

Photometric-redshift Algorithms

Colour/Flux(+...) space

Redshift space



z=f(colours, fluxes)

Mapping *f* can be constructed based on prior astrophysical knowledge :

- Template-fitting: Hyper-Z, Le Phare, BpZ, Phosphoros,...

Or it can be discovered using known (spectroscopic) redshifts:

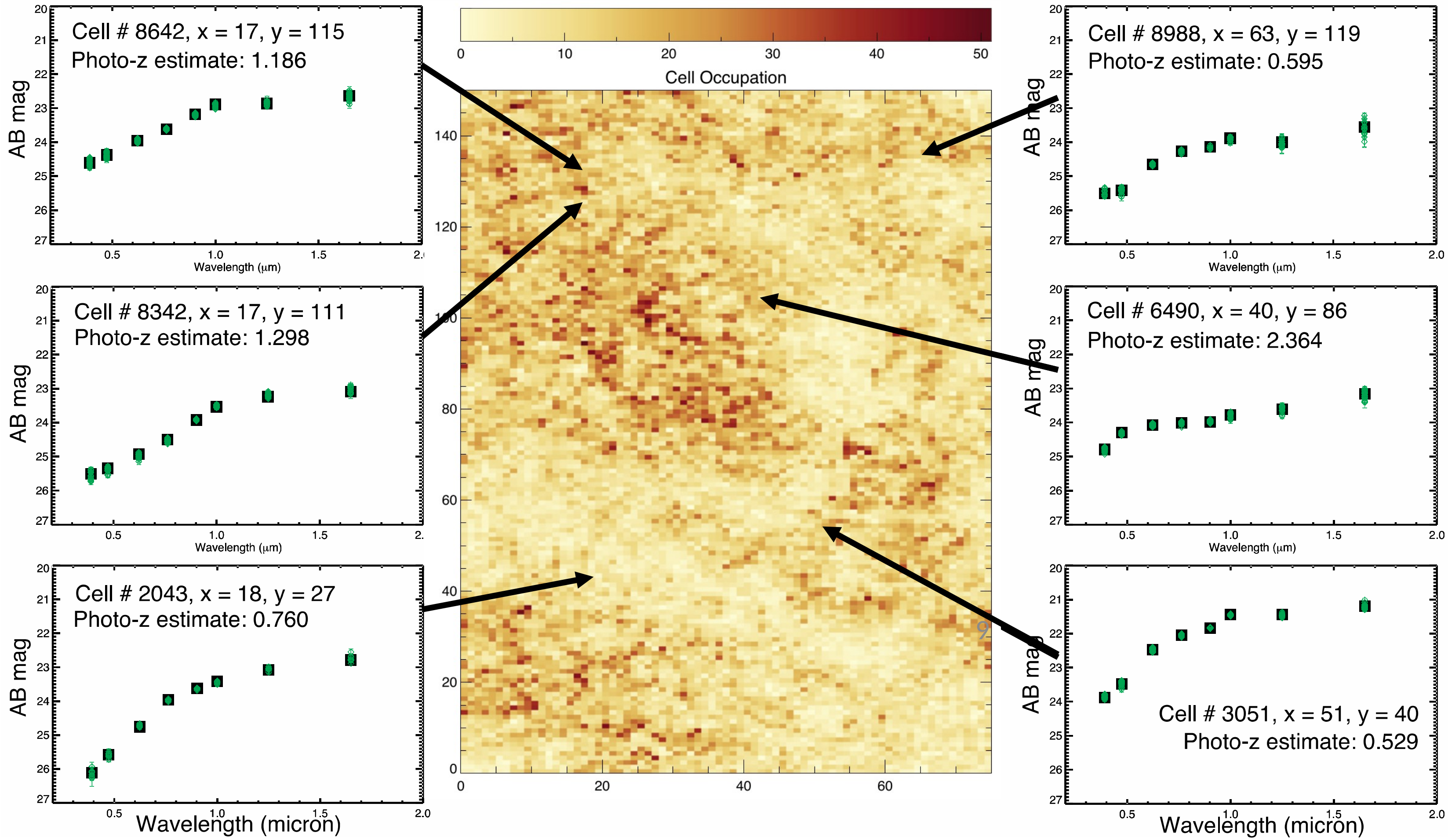
- Machine-learning: Nearest neighbours, Perceptron, Support vector regression, Random Forest, Adaboost, Gaussian Processes, …

- See Guillaume’s presentation

Template-fitting vs Machine-learning

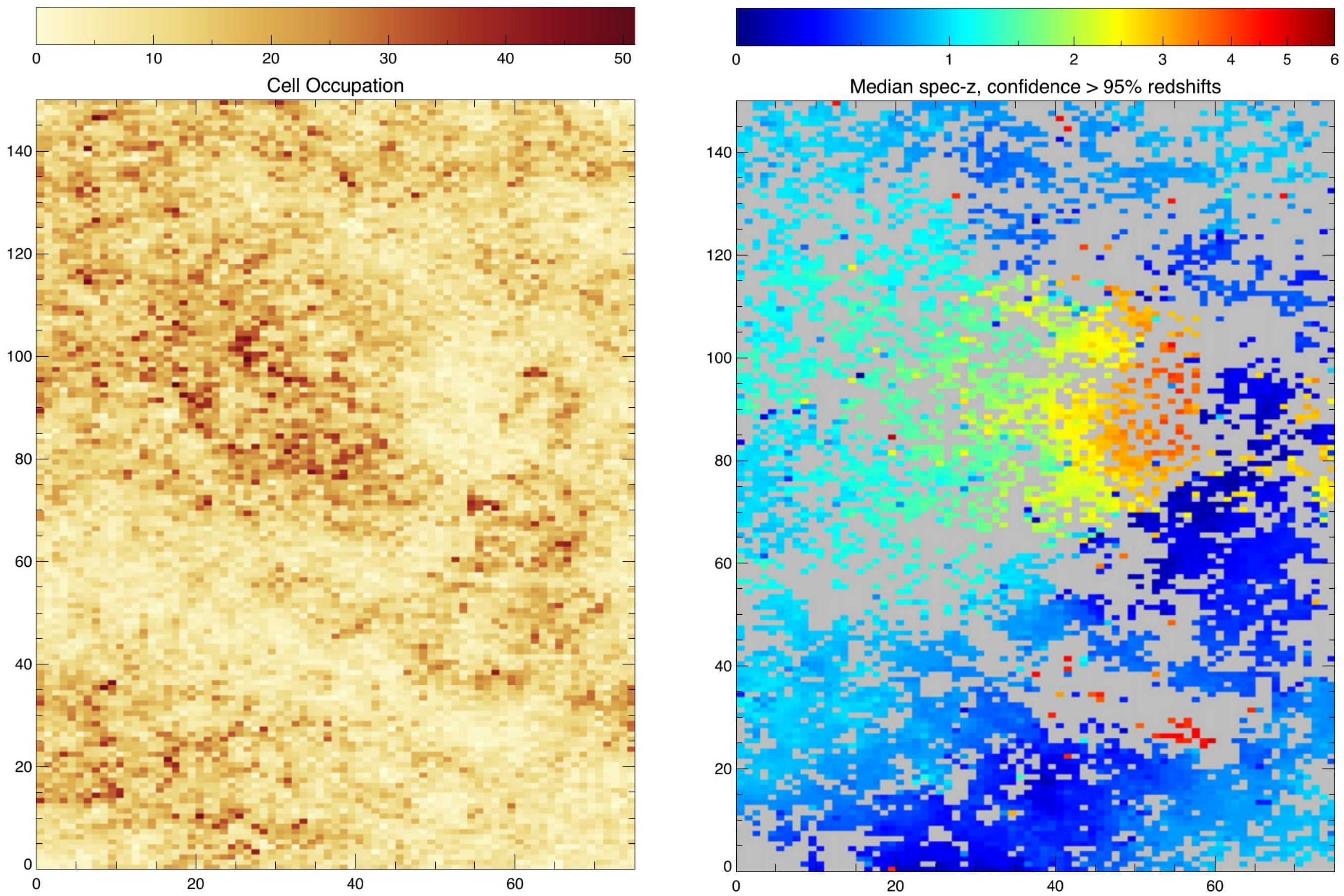
	Template-fitting	Machine-learning
Probabilistic	✓	✗
Does not require any understanding of physics	✗(✓)	✓(✗)
Does not require complete training sample	✓	✗
Implement priors naturally	✓	✗
Can be free of hidden priors	✓	✗
Can use heteroscedatic data (but useful?)	✗	✓
Little sensitive to extrapolations	✓	✗
Computationally efficient	✗	✓
Easy to prevent over/underfitting	✗	✓
Optimal for the test sample	✗	☺

Self-Organizing Map of Galaxy Colour Space



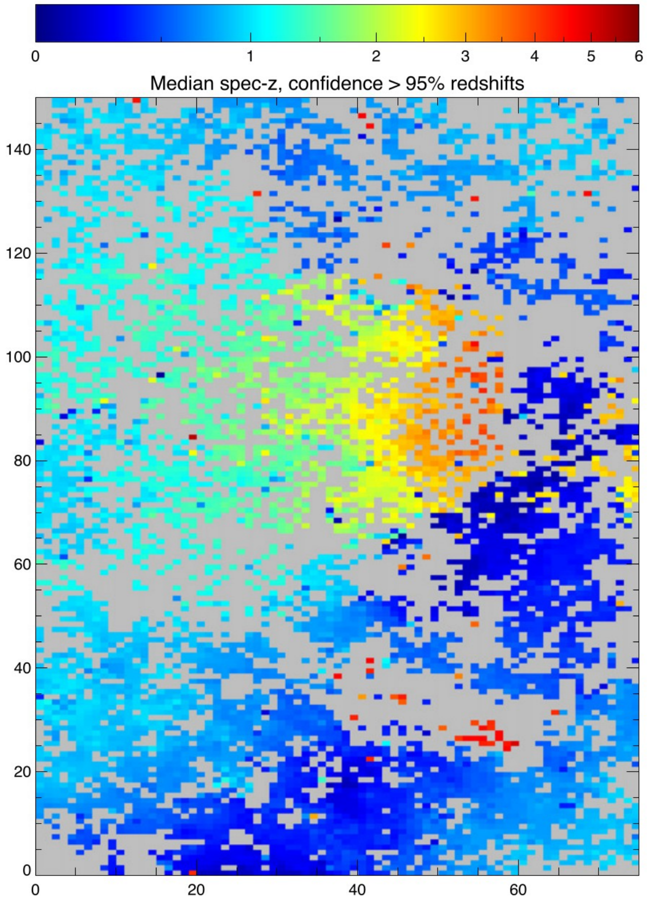
Completeness of the Training Sample

- A self-organizing map can be used to find out where are the missing objects in color space
- Large areas of the color space lack any spec-z coverage, some parts in high-density areas



C3R2 Program

- Complete Calibration of the Color-Redshift Relation
- 50 nights on Keck, 200h on VLT
- Goal is to determine <z> in each cell



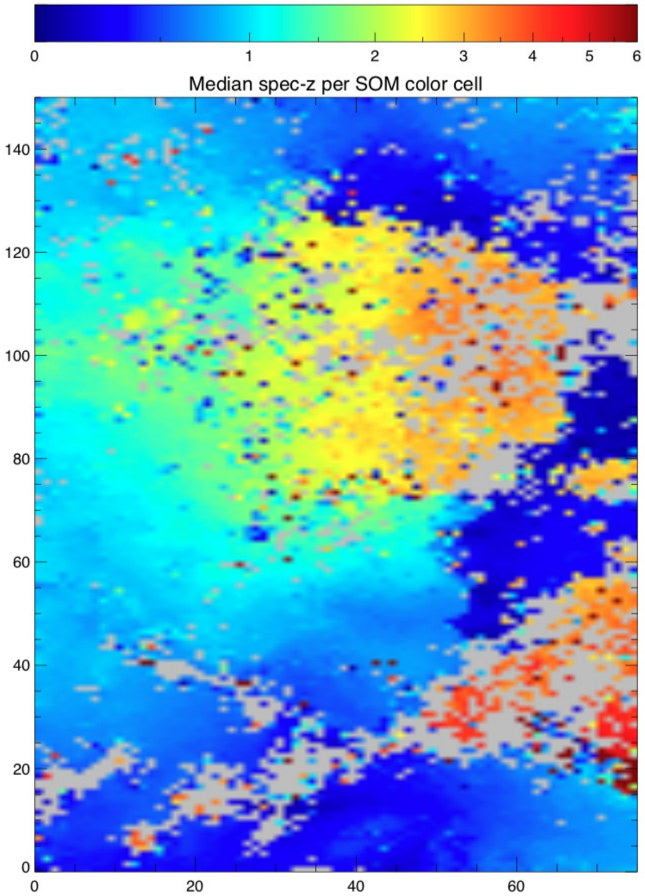
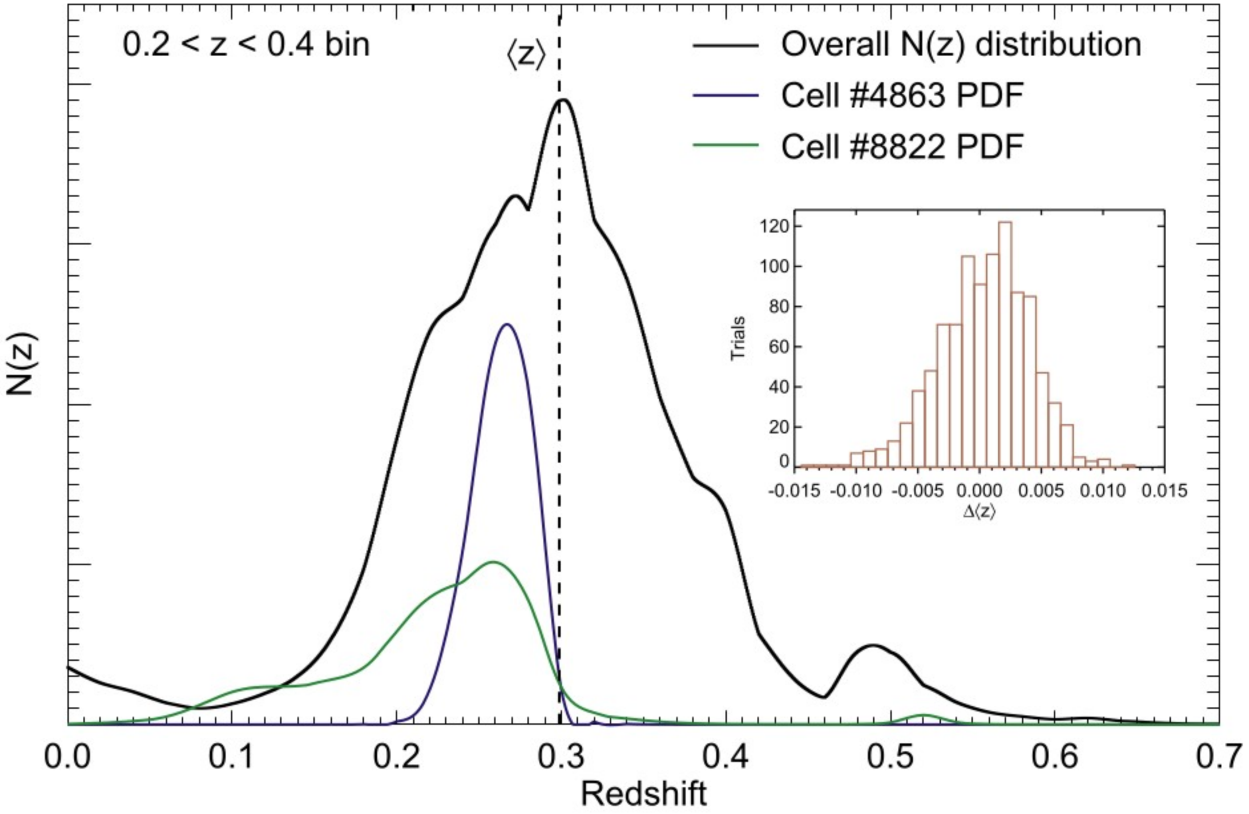


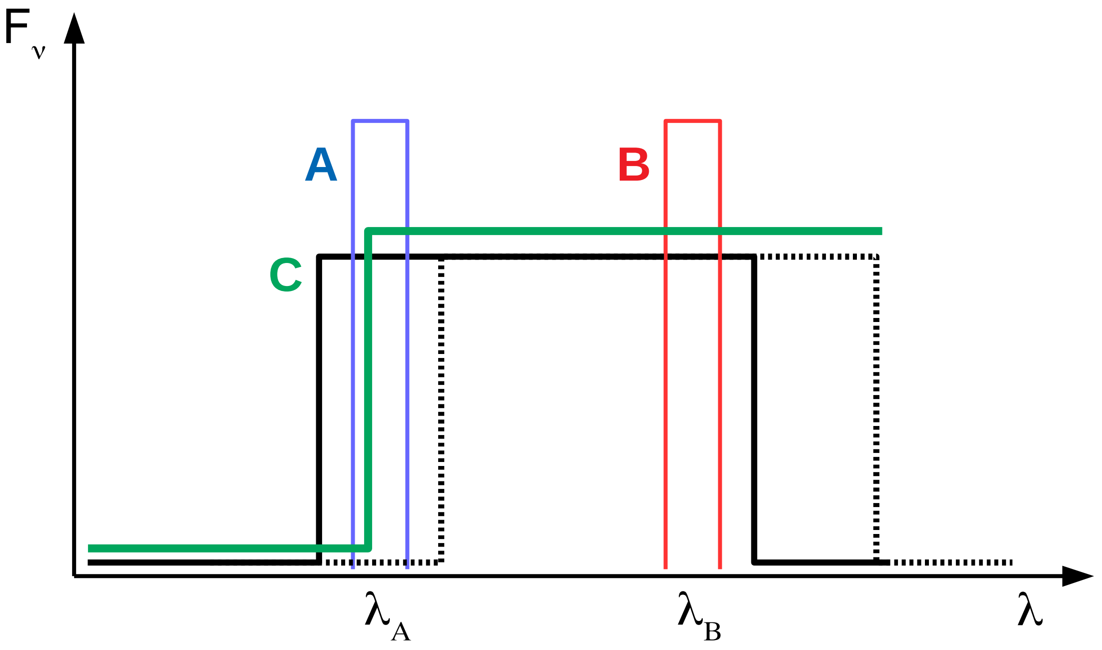
Photo-z Calibration using the SOM

$$< z >= \frac{1}{\sum_i n_i} \sum_i n_i < z_i >$$



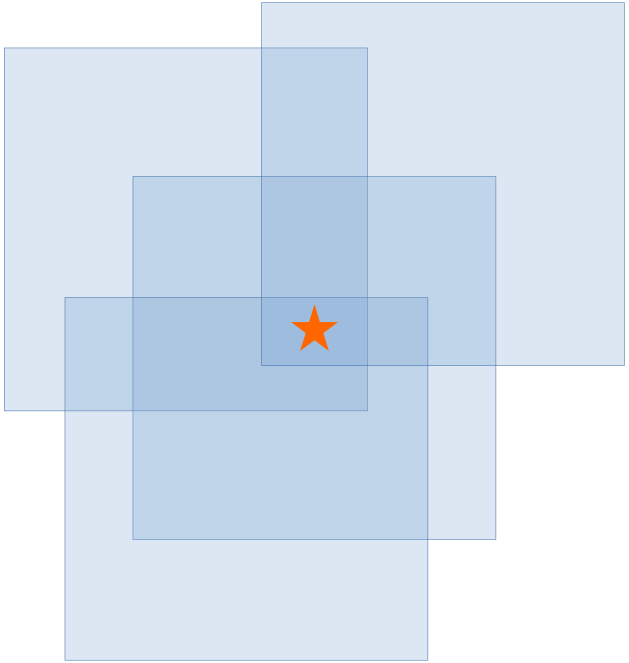
Inhomogeneous Photometric Survey

- Difference in transmissions lead to inhomogeneous colors
- This can immediately introduce a bias in the photo-z if the effect is not taken into account
- Usually addressed using a colour-term correction
  - But depends on SED

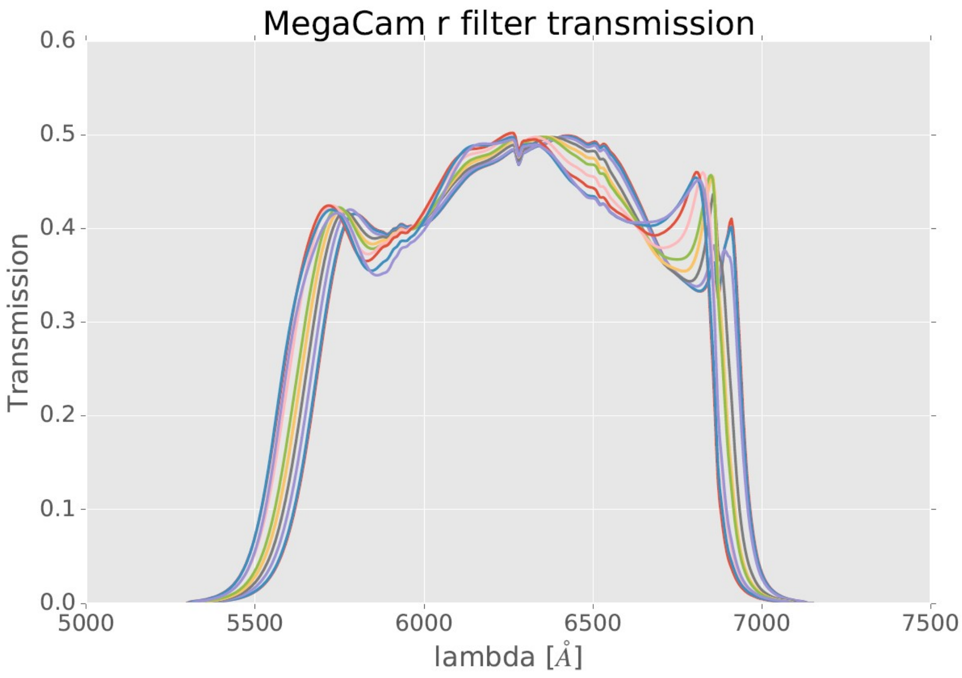


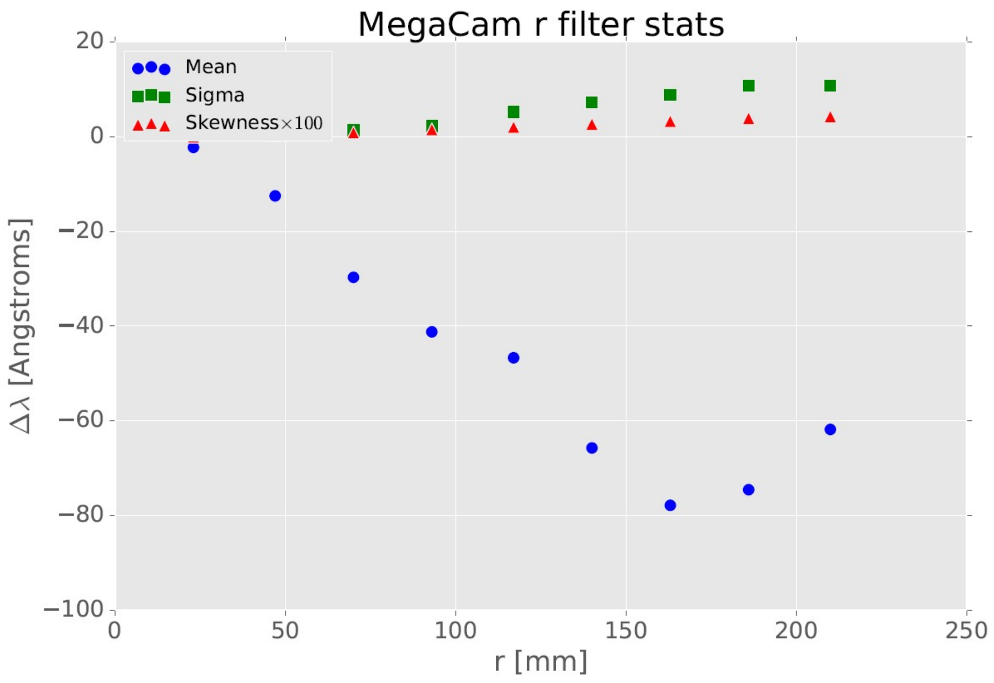
Variable Transmissions

- Variations in transmissions in a single survey can occur for many reasons:
  - Inhomogeneous filters and CCD QEs, dependent on the location in the FOV
  - Filter or CCD aging, dependent on time
  - Galactic reddening, dependent on sky position
- Some effects are mitigated by having several pointings, but not all, and not clear to which amount



MegaCam *r* Filter Transmission(s)

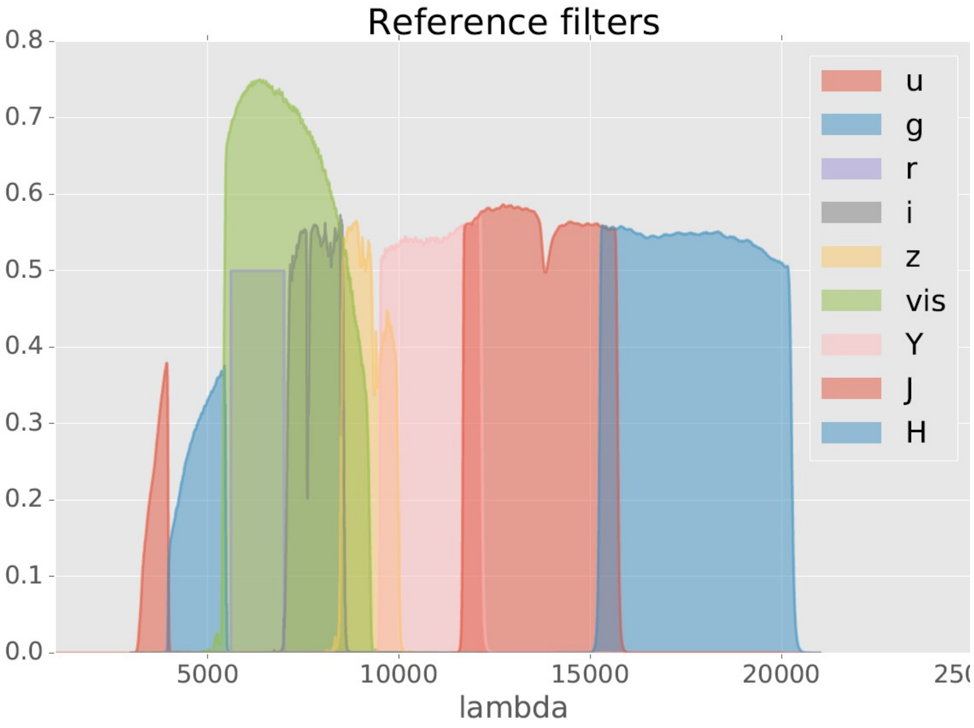


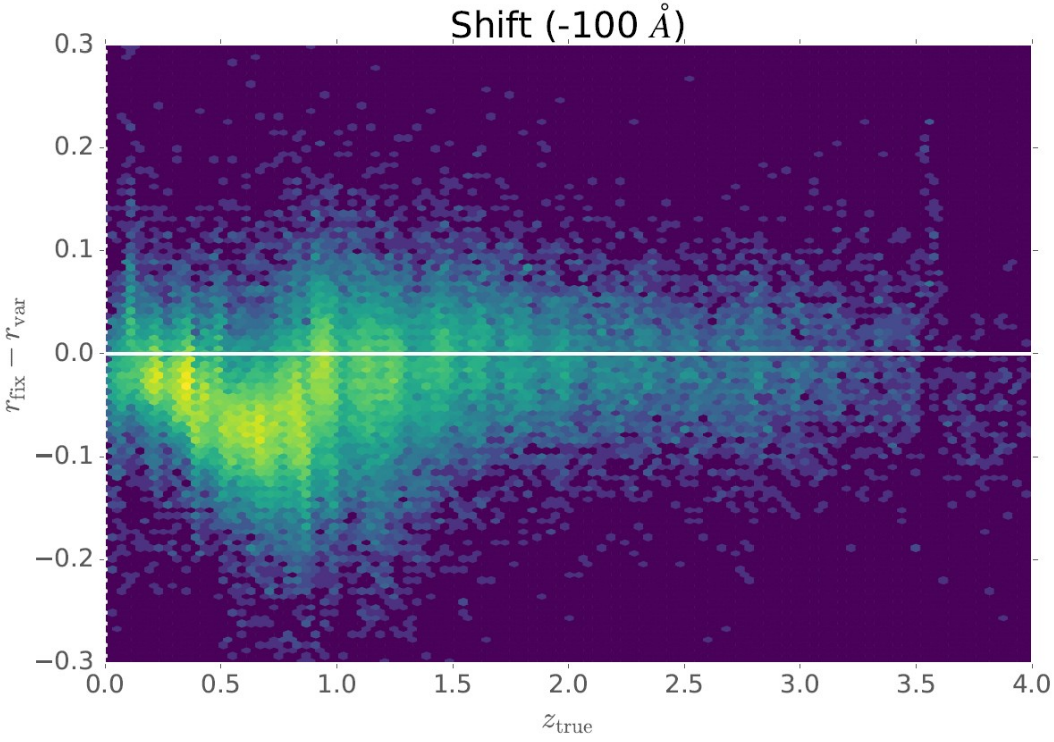


Clear effect of the off-axis angle

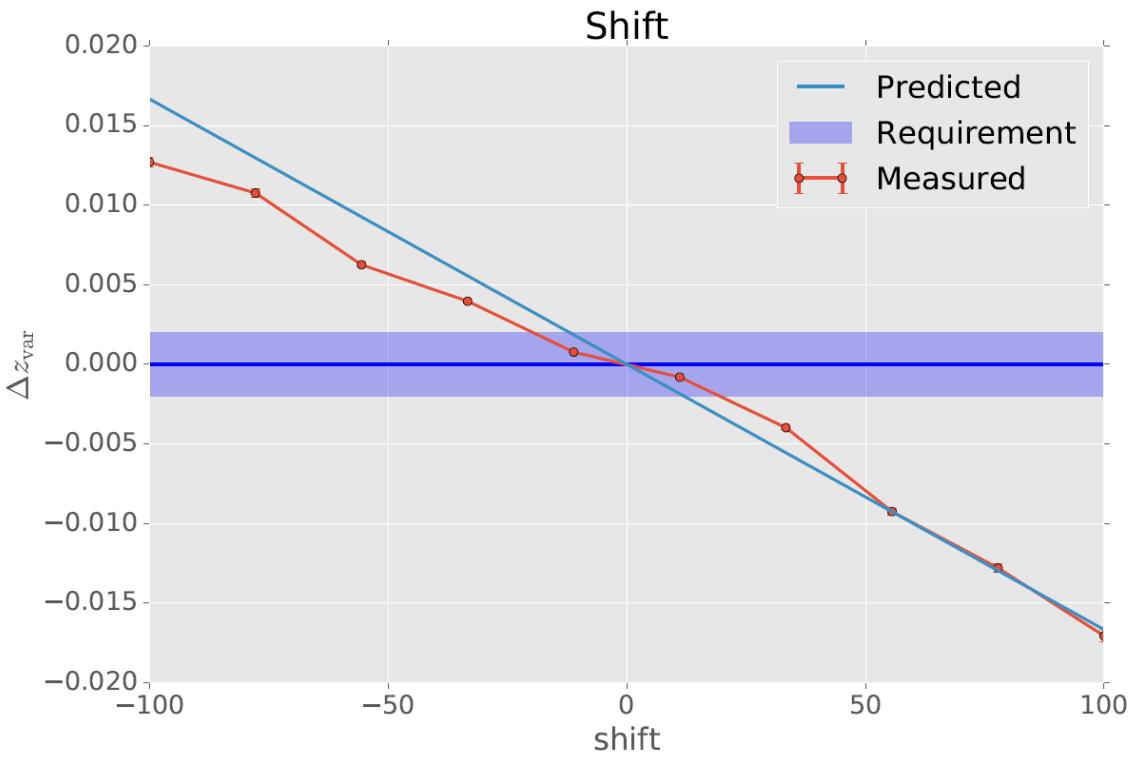
Effect of Variable Transmissions

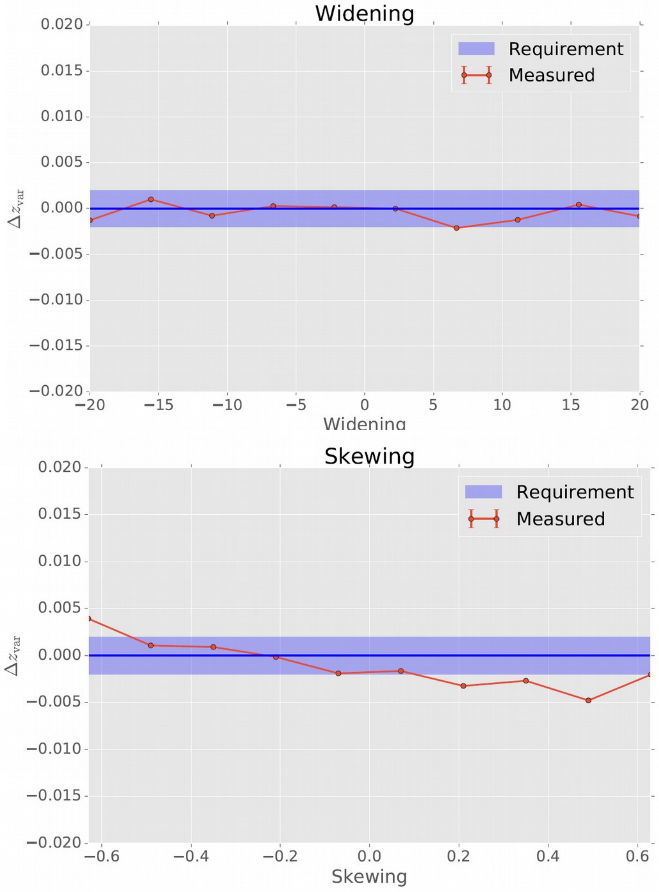
We simulate a catalogue based on COSMOS with varying *r* filter





Bias due to Variable Transmissions





Main effect is a shift of the filter central wavelength

How to deal with Variable Transmissions

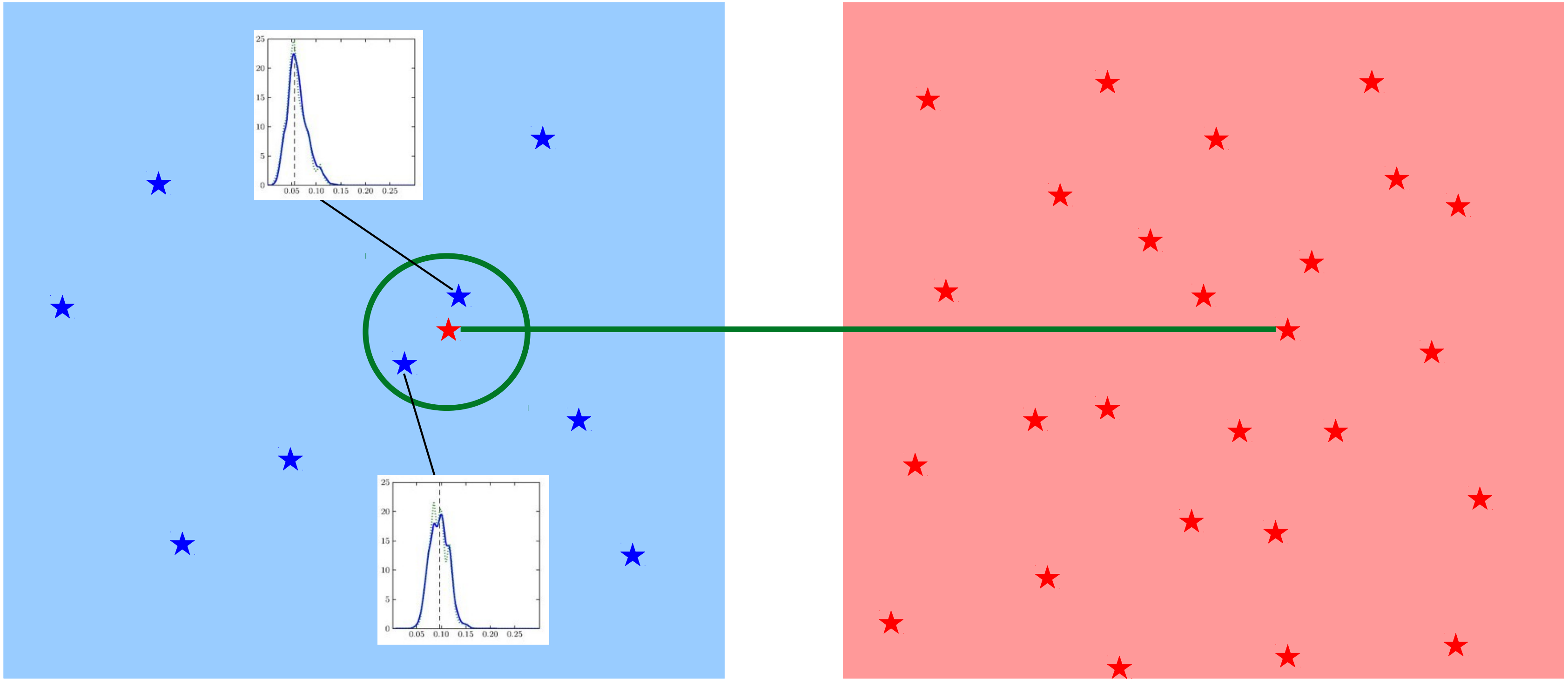
- Template fitting can deal easily with variable transmissions in principle, but prohibitive computation time, because the grid is different for each object
- ML has trouble, because colour uniformity is a pre-requisite
- But if we know the SEDs of enough reference galaxies, we can construct their colours in any colour space
  - However, we would need to train the ML algorithms for each object, making it computationally impossible
  - Unless the ML algorithm does not need training : Nearest Neighbour !!
- The reference galaxies are galaxies in well-known deep surveys, observed in many bands and a lot of redshifts: COSMOS, SXDF, VVDS-2H

Nearest Neighbour

Reference Sample
10^5 objects
Known PDZs

COLOUR SPACE

Euclid Sample
Few 10^9 objects
Unknown PDZs

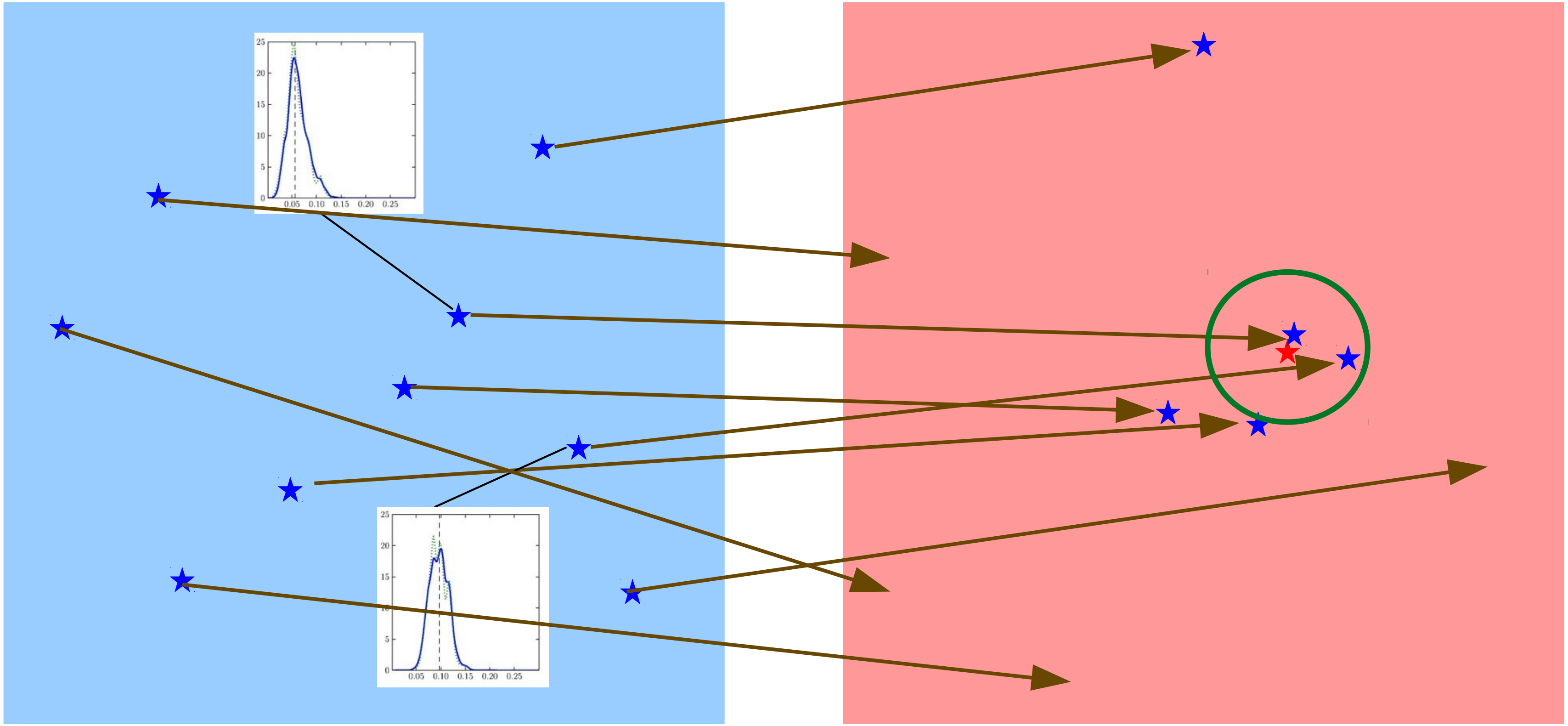


NNPZ for Euclid

Reference Sample
10^5 objects
Known PDZs
Known SEDs

COLOUR SPACE

Euclid Sample
Few 10^9 objects
Unknown PDZs



Conclusions

- Some very challenging tasks
  - Calibration
  - Filter variations
  - Classification
  - Colour reconstruction
- Good baseline, but remains close to the feasibility limit