PDF with MLPQNA

Cavuoti Stefano,	*INAF OACN*
Brescia Massimo,	*INAF OACN*
Longo Giuseppe,	*University Federico II of Naples*
Amaro Valeria,	*University Federico II of Naples*
Vellucci Civita,	*University Federico II of Naples*

Photometric redshift PDF

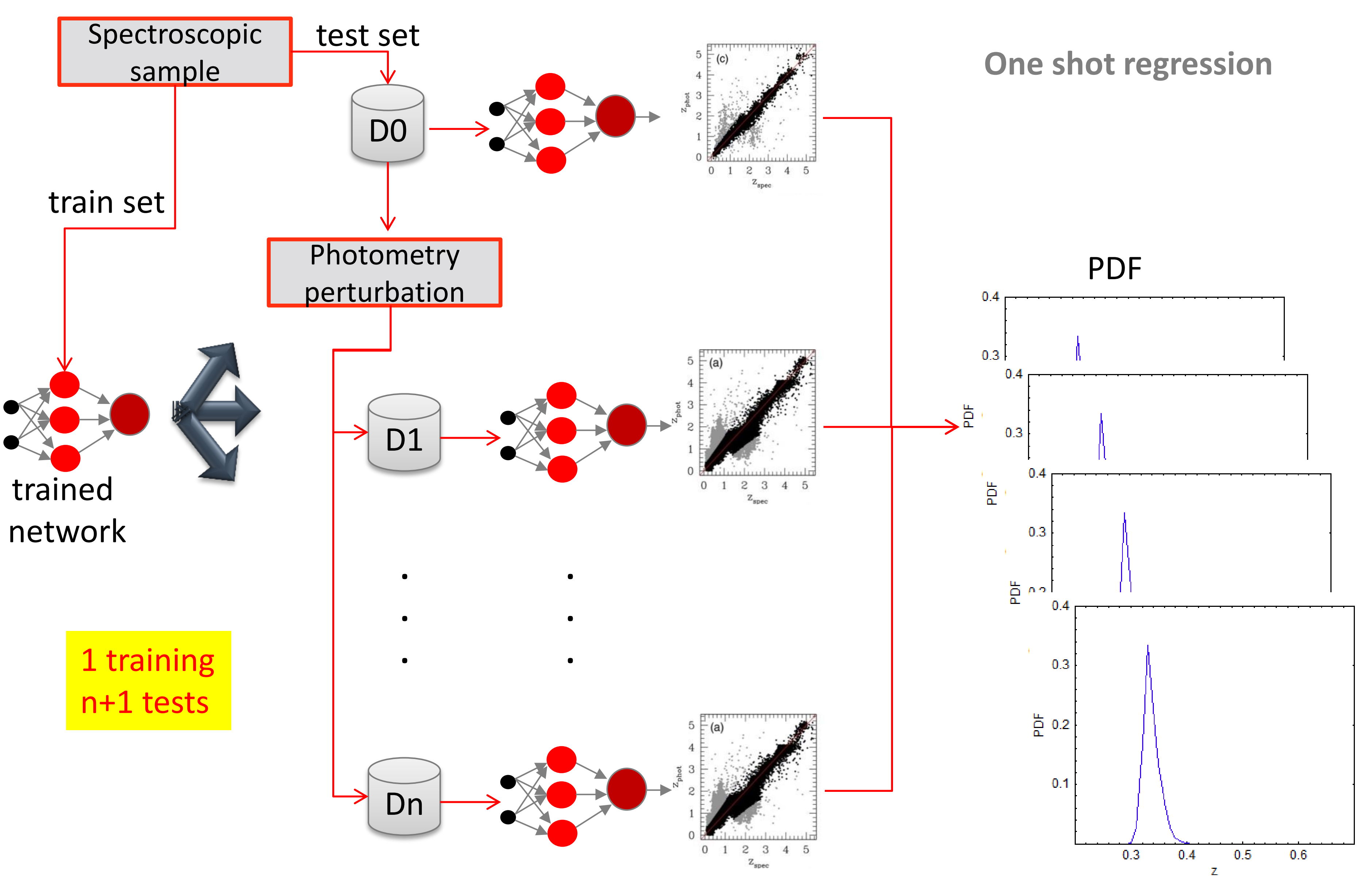
We started our R&D process by a level-0 method (called ***base algorithm***), able to provide a PDF estimation of the photo-z for each single input object of the data sample used.
Then we are still under debugging a series of more complex methods based on a post-processing of photo-z production model.

The common element of such process is the machine learning model used to derive photo-z. The model is MLPQNA (Multi Layer Perceptron trained by the Quasi Newton Algorithm), already successfully validated on several real cases.

Photo-z with MLPQNA

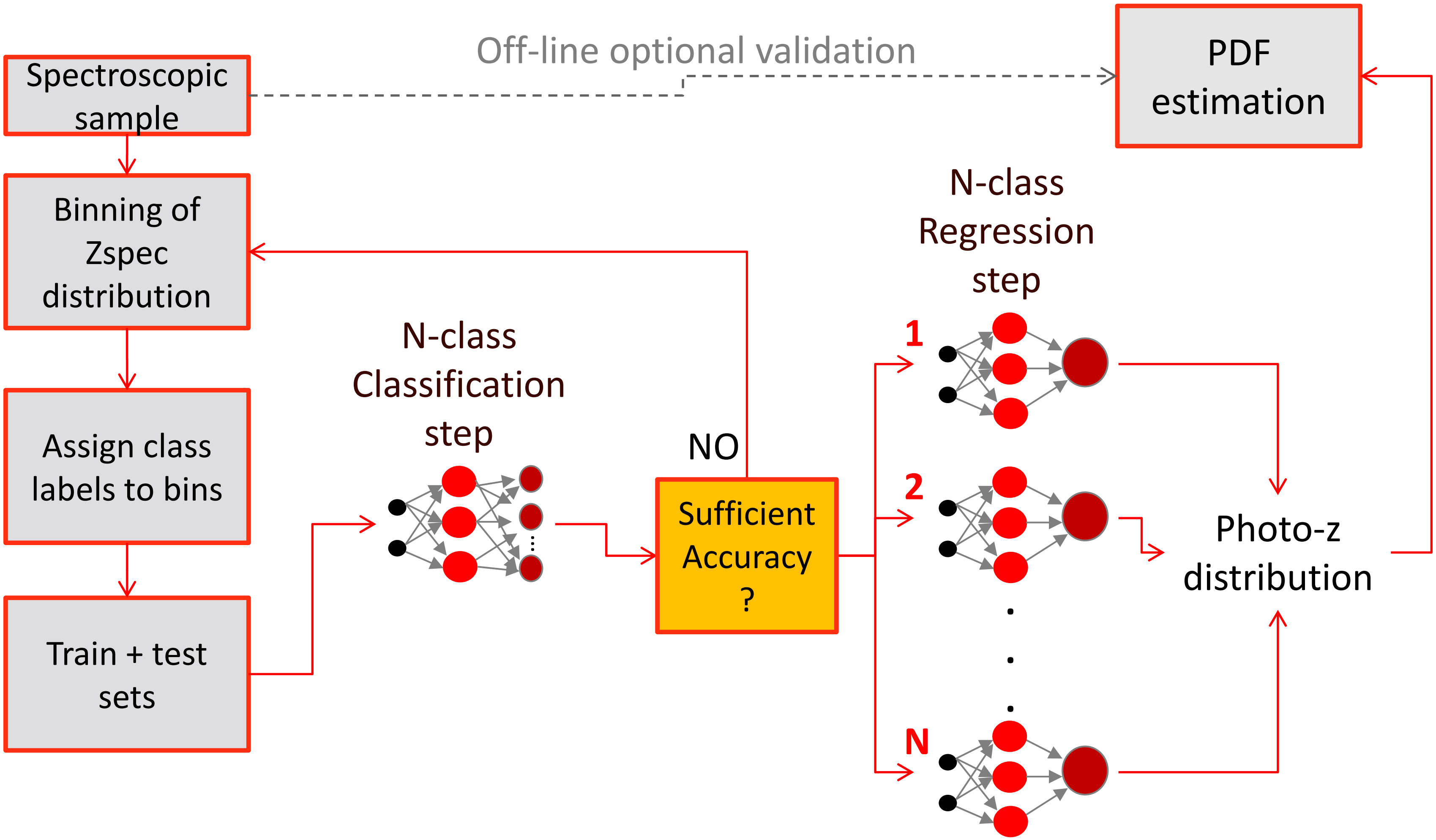
- **PHAT1 Contest** (*Cavuoti et al. 2012, A&A, 546, A13*)
- **GALEX+SDSS+UKIDSS+WISE QSOs** (*Brescia et al. 2013, ApJ, 772, 2, 140*)
- **CLASH-VLT** (*Biviano et al. 2013, A&A, 558, A1*)
- **EUCLID PHZ** (*Coupon et al. 2014, Challenge #1 internal report*)
- **SDSS DR9** (*Brescia et al. 2014, A&A, 568, A126*)
- **KiDS DR2** (*Cavuoti et al. 2015, MNRAS, accepted, in press*)
- **VST VOICE** (*Covone et al. 2015, in prep.*)
- **XMM** (*Vaccari et al. 2015, in prep.*)

PDF base algorithm processing flow

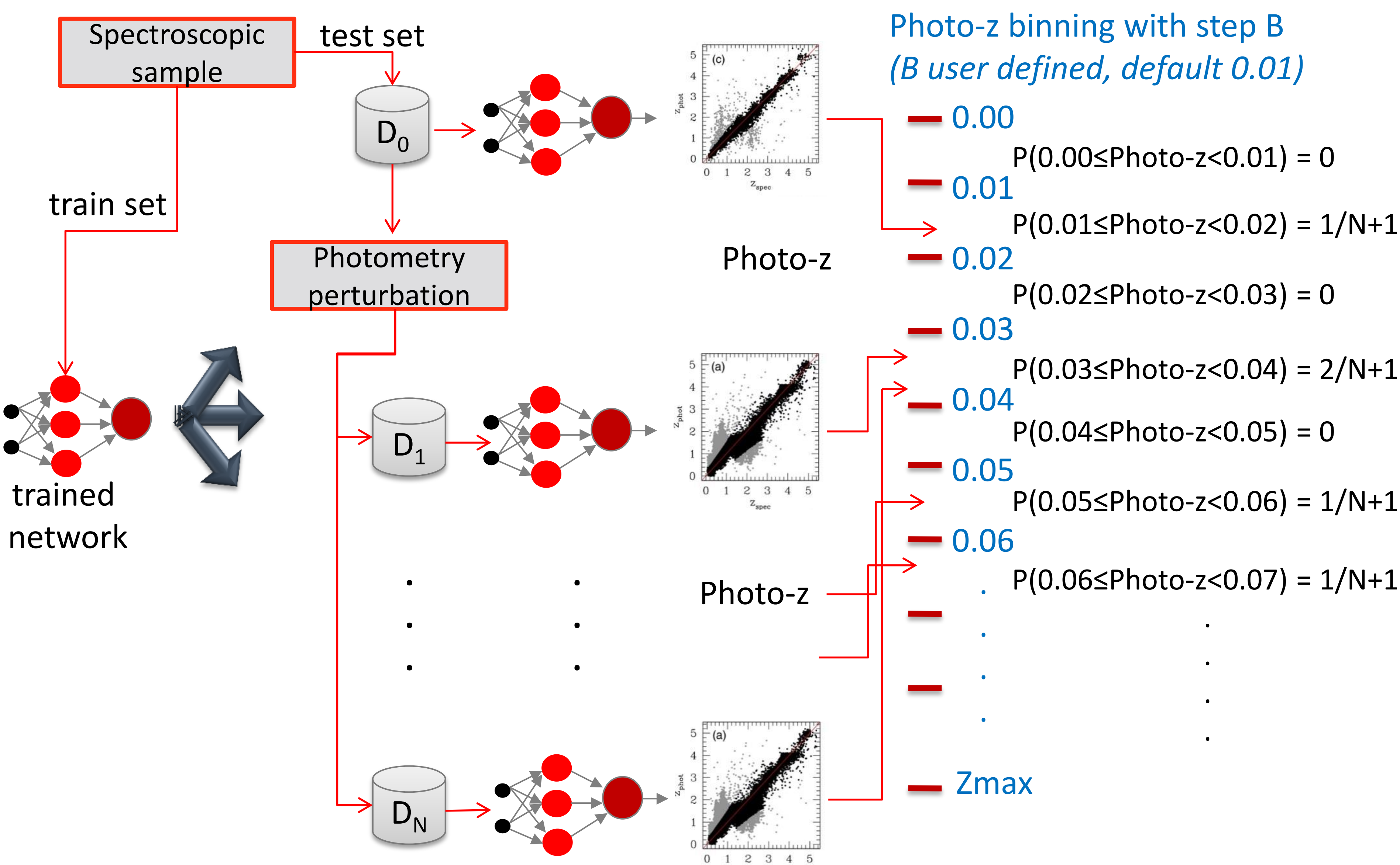


PDF base algorithm processing flow

Hierarchical approach



PDF base algorithm processing flow



$$PDF(Photo\text{-}z) = \{P(Z_i \leq Photo\text{-}z < Z_{i+B}) = C_{B,i}/N+1\}_{[Zmin,\ Zmax]}$$

Photometry perturbation

Given a dataset A, a normal distribution on A, and

$N_{samples}$ number of objects in a given dataset A

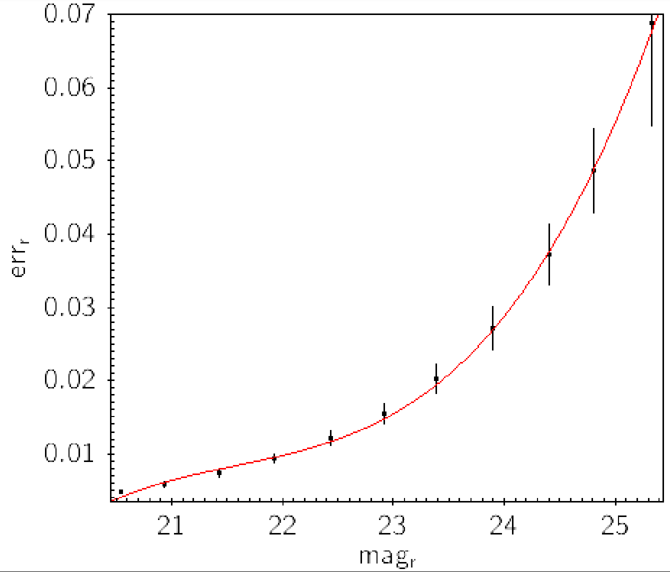
$N_{perturb}$ number of perturbations to be done

N_{mags}, number of affected magnitudes

p_b polynomial used to perturb mag of band b

$alpha_b$ perturbation constant for the band b

$mag_b(o_i)$ mag value of the band b for the object o_i

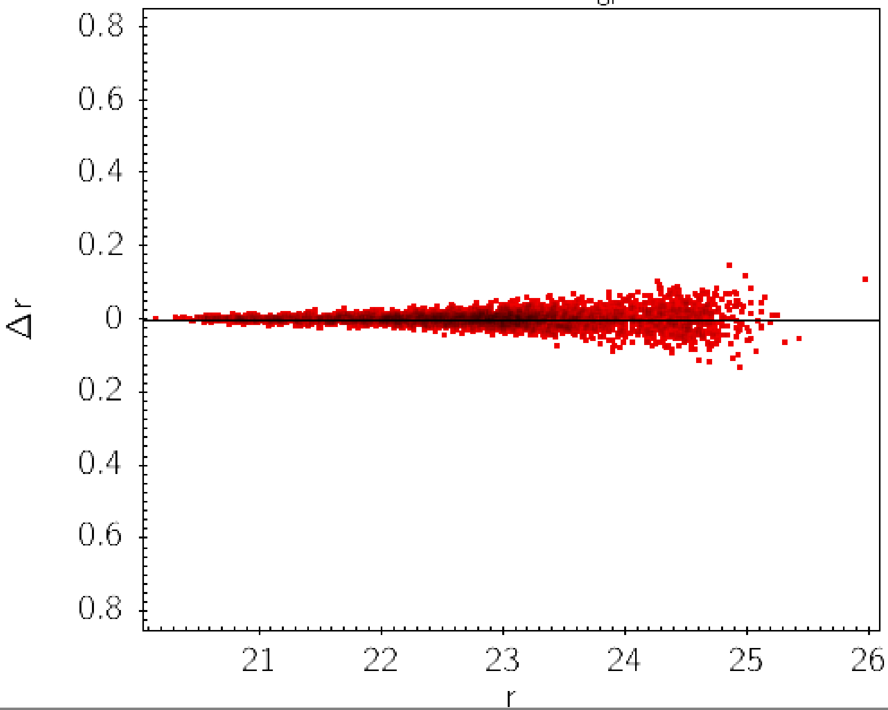


$$m_{ijperturbed}(o_i) = m_{ij} + alpha_b * p_b \circ (mag(o_i)) * N_A(0;1)$$

where the symbol "∘" stays for the scalar product,

$N_A(0;1)$ is a normal distribution with the dimension of the dataset A to be perturbed, i.e. a distribution of a number $N_{samples}$ of values in the interval (-1,1).

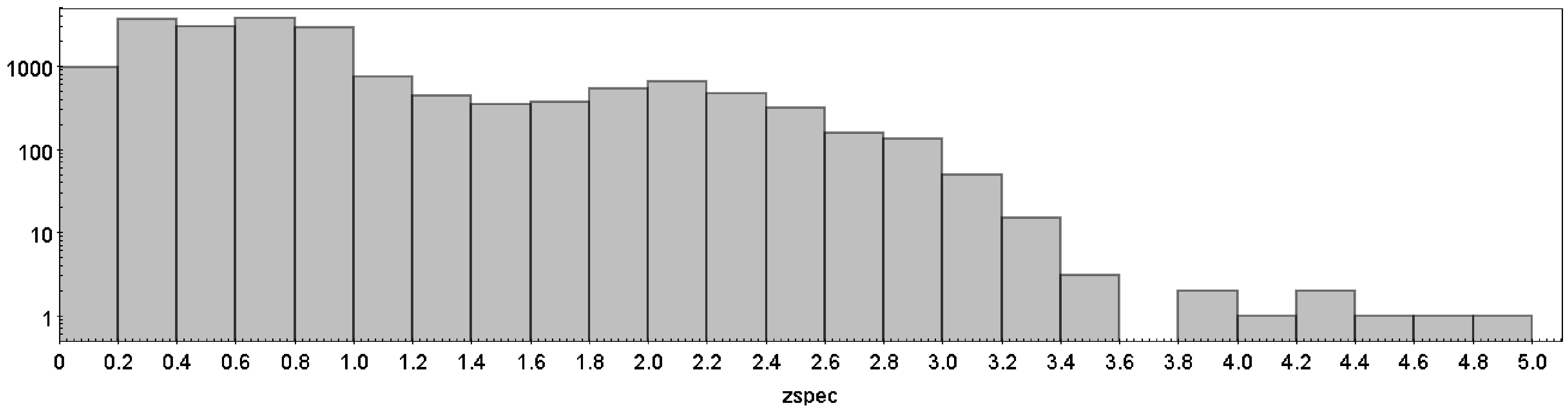
The variation of the percentage of noise is ensured by the randomly generated normal distribution at each step.



The KB

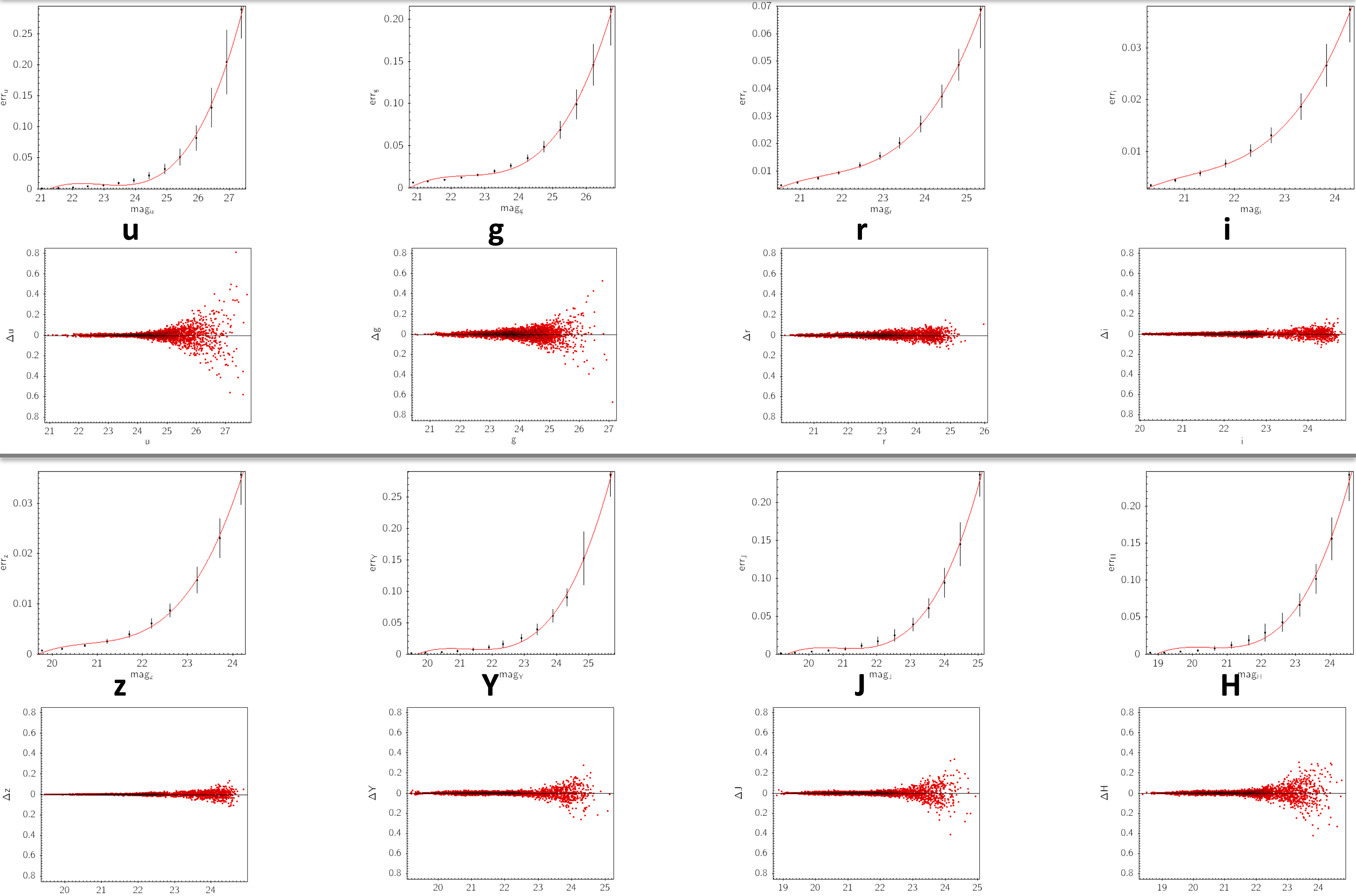
The dataset used for the current test is the same utilized by Masters et al. 2015 containing the following information, matched to the Euclid Requirements:

- u → CFHT
- griz → SUBARU
- Y,J,H → ULTRAVISTA
- zspec → Salvato 2016 (in prep)



For the following experiment we fitted the errors with a 3rd order polynomial expansion.

Photometry perturbation



Two class approach vs One Shot

First stage of Hierarchical approach – 2-class classification			
CLASSIFICATION	MAGNITUDES ONLY (8 features)	COLORS ONLY (7 features)	COLORS + MAGNITUDES (9 features)
% average efficiency	96.08	95.94	95.73
% zspec<1 purity	98.09	97.72	97.61
% zspec≥1 purity	89.06	89.58	88.99
% zspec<1 completeness	96.91	97.11	96.94
% zspec≥1 completeness	93.02	91.62	91.24
TRAIN/TEST dimensions	14,837 / 3,698		

1[st] stage: 2-class classification

2[nd] stage: multi-regression

COLORS + MAGNITUDES (9 features)			
REGRESSION	one-shot approach	2-class Hierarchical approach	
	FULL redshift range	zspec < 1	zspec >= 1
|Bias|	0.0112	0.0006	0.0089
σ	0.169	0.074	0.127
NMAD	0.036	0.020	0.082
% Outliers>0.15	8.71	3.75	18.40
TRAIN/TEST dim.	14,837 / 3,698	11,384 / 2,910	3,453 / 788

COLORS ONLY (7 features)			
REGRESSION	one-shot approach	2-class Hierarchical approach	
	FULL redshift range	zspec < 1	zspec >= 1
|Bias|	0.0198	0.0010	0.0166
σ	0.185	0.066	0.140
NMAD	0.044	0.021	0.086
% Outliers>0.15	9.44	3.44	19.03
TRAIN/TEST dim.	14,837 / 3,698	11,384 / 2,910	3,453 / 788

MAGNITUDES ONLY (8 features)			
REGRESSION	one-shot approach	2-class Hierarchical approach	
	FULL redshift range	zspec < 1	zspec >= 1
|Bias|	0.0103	0.0012	0.0178
σ	0.132	0.058	0.138
NMAD	0.037	0.014	0.076
% Outliers>0.15	8.11	3.26	16.62
TRAIN/TEST dim.	14,837 / 3,698	11,384 / 2,910	3,453 / 788

In the right tables the one-shot regression is also reported for direct comparison

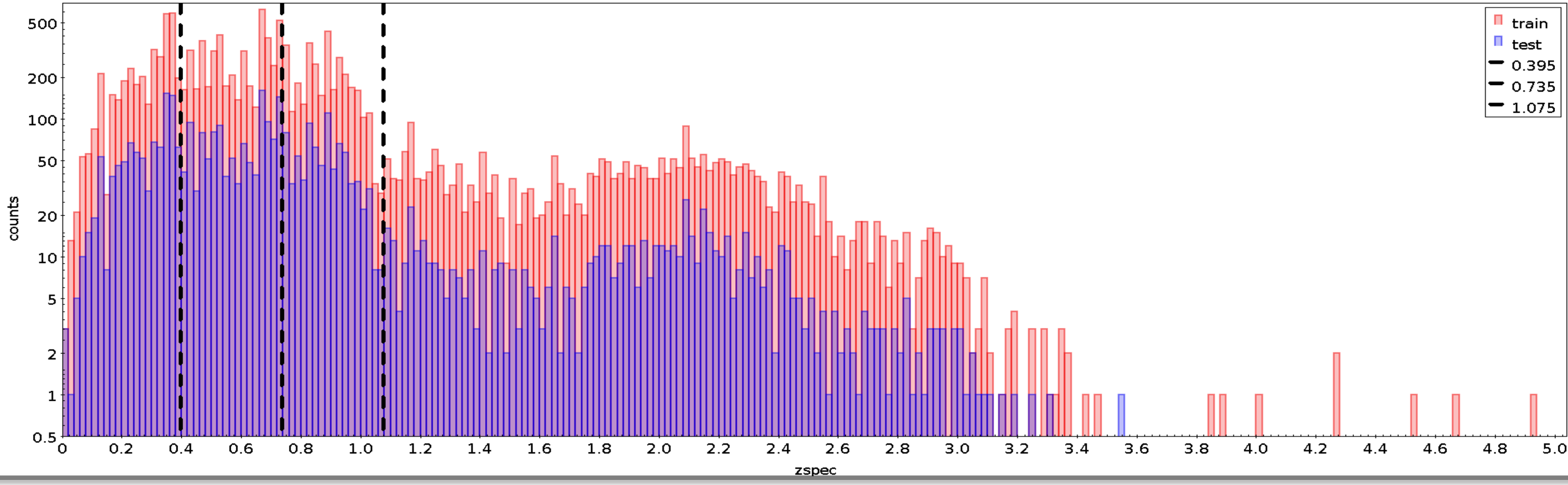
Four-class approach

In this experiment we define the classes on the base of the break at 4000 Å.
In order to properly select the redshift bins, we considered the transmission curves provided at the CALTECH web page (http://www.astro.caltech.edu/~capak/filters/index.html).
We therefore measured for each band the zspec value corresponding to the entry point of the break, resulting as follows:

Band u has the quantum efficiency peak at 4065Å;
Band g → zspec = 0.033;
Band r → zspec = 0.395;
Band i → zspec = 0.735;
Band z → zspec = 1.075;
Band Y → zspec = 1.440;
Band J → zspec = 1.915;
Band H → zspec = 2.753.

In order to maintain almost balanced the dimensions of bins and following some heuristics learned from previous experiments, we identified the following 4 classes:

Class 1: zspec < 0.395 (break of band r);
Class 2: 0.395 ≤ zspec < 0.735 (break of band i);
Class 3: 0.735 ≤ zspec < 1.075 (break of band z);
Class 4: 1.075 ≤ zspec.



Four-class approach

CONFUSION MATRIX		CLASS OUTPUT			
CLASS TARGET		1	2	3	4
	1	860	51	6	14
	2	72	1036	64	14
	3	15	63	743	41
	4	12	4	22	681

mean accuracy	mean purity	mean completeness
90%	90%	90%

We explored also a 7-class approach by simply balancing the seven bins in terms of quantity but obtaining lower results, as expected.

(no physical meaning)

MAGNITUDES ONLY (8 features)

REGRESSION	one-shot approach	4-class Hierarchical approach			
	FULL redshift range	Class 1 zspec < 0.395	Class 2 [0.395, 0.735[	Class 3 [0.735, 1.075[	Class 4 1.075 ≤ zspec
|Bias|	0.0103	5.4E-5	2.5E-5	2.9E-6	0.0172
σ	0.132	0.035	0.026	0.023	0.135
NMAD	0.037	0.017	0.017	0.015	0.075
% Outliers>0.15	8.11	1.07	0.17	0.0	15.99
TRAIN/TEST dimensions	14,837 / 3,698	3605 / 931	4700 / 1186	3347 / 862	3185 / 719

COLORS ONLY (7 features)

REGRESSION	one-shot approach	4-class Hierarchical approach			
	FULL redshift range	Class 1 zspec < 0.395	Class 2 [0.395, 0.735[	Class 3 [0.735, 1.075[	Class 4 1.075 ≤ zspec
|Bias|	0.0103	0.0009	0.0006	7.8E-5	0.0184
σ	0.132	0.035	0.027	0.024	0.144
NMAD	0.037	0.017	0.016	0.015	0.091
% Outliers>0.15	8.11	1.18	0.5	0.0	20.45
TRAIN/TEST dimensions	14,837 / 3,698	3605 / 931	4700 / 1186	3347 / 862	3185 / 719

COLORS + MAGNITUDES (9 features)

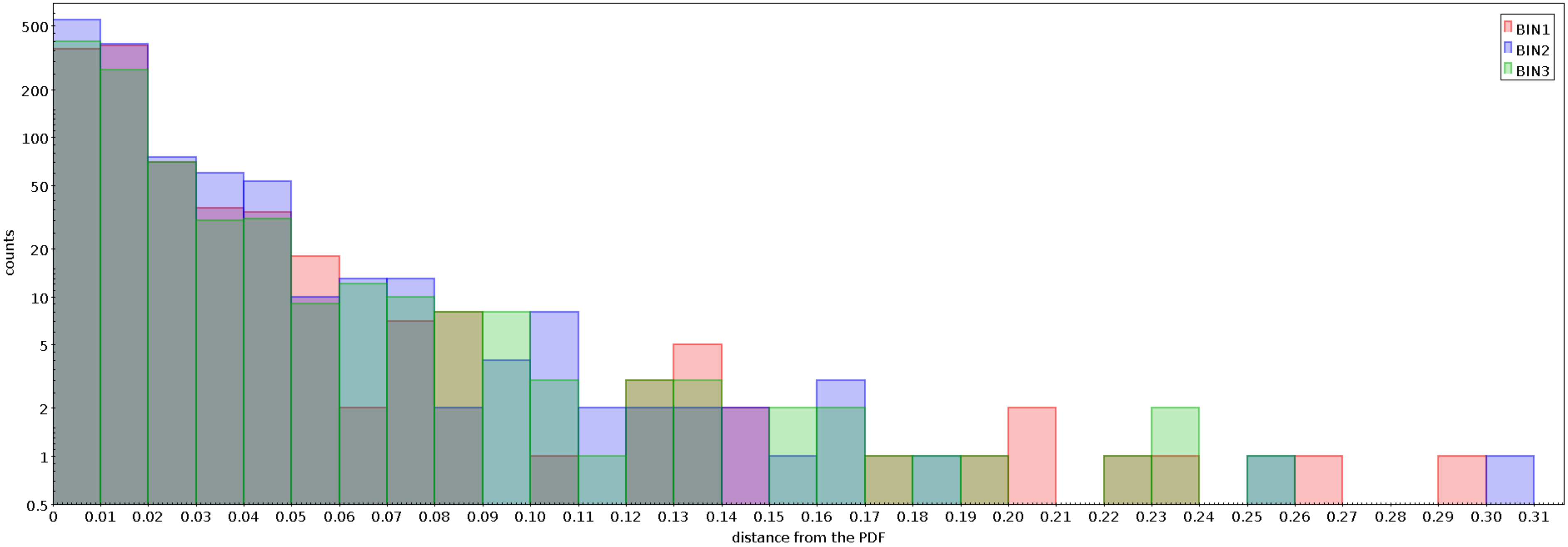
REGRESSION	one-shot approach	4-class Hierarchical approach			
	FULL redshift range	Class 1 zspec < 0.395	Class 2 [0.395, 0.735[	Class 3 [0.735, 1.075[	Class 4 1.075 ≤ zspec
|Bias|	0.0103	0.0011	0.0001	0.0011	0.0148
σ	0.132	0.039	0.029	0.026	0.158
NMAD	0.037	0.016	0.010	0.016	0.086
% Outliers>0.15	8.11	1.72	0.51	0.12	18.36
TRAIN/TEST dimensions	14,837 / 3,698	3605 / 931	4700 / 1186	3347 / 862	3185 / 719

PDFs

We derived our PDFs through ten redshift binning ranges, from 0.01 up to 0.1.
We considered the best photo-z guess the peak with the highest probability closest to the photo-z obtained without photometric perturbation.

By considering a PDF bin of 0.03:

- The 46%, 38% and 40% of objects for class 1, 2 and 3 respectively, have their zspec within the peak of the PDF;
- While the 84% 79% and 76%, have zspec falls within the PDF. By considering also the bin closest to the PDF, the percentages grow up to 87%, 85% and 85% respectively.

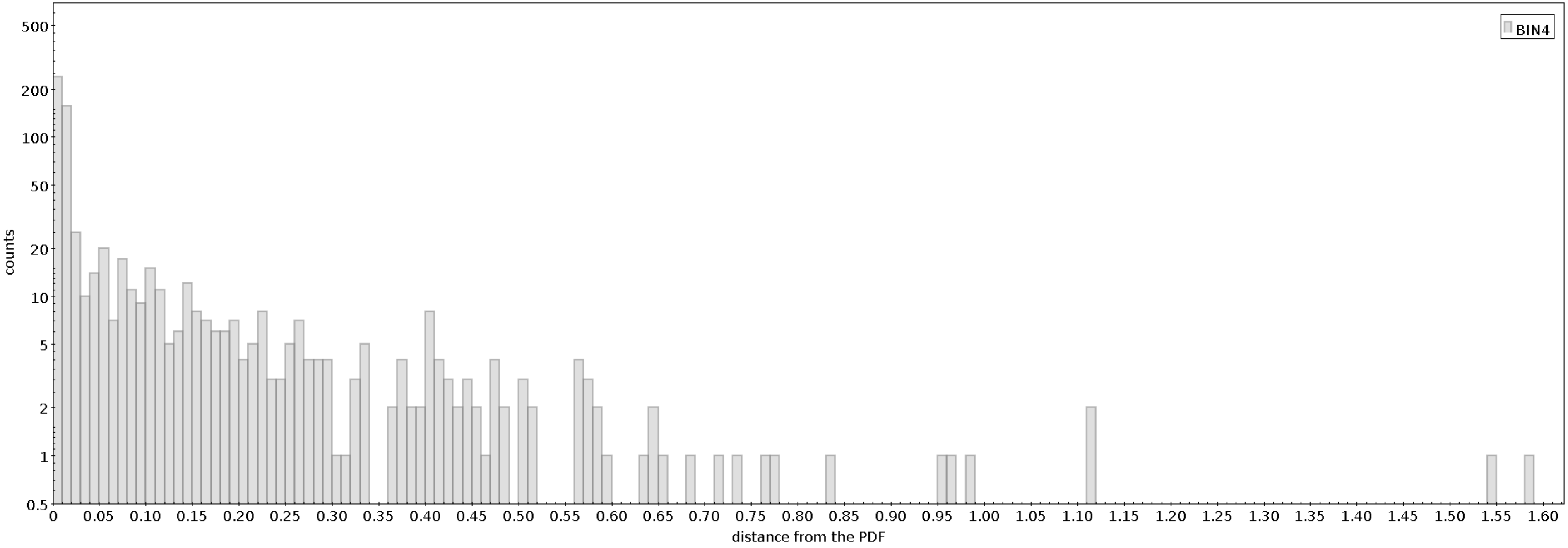


PDFs

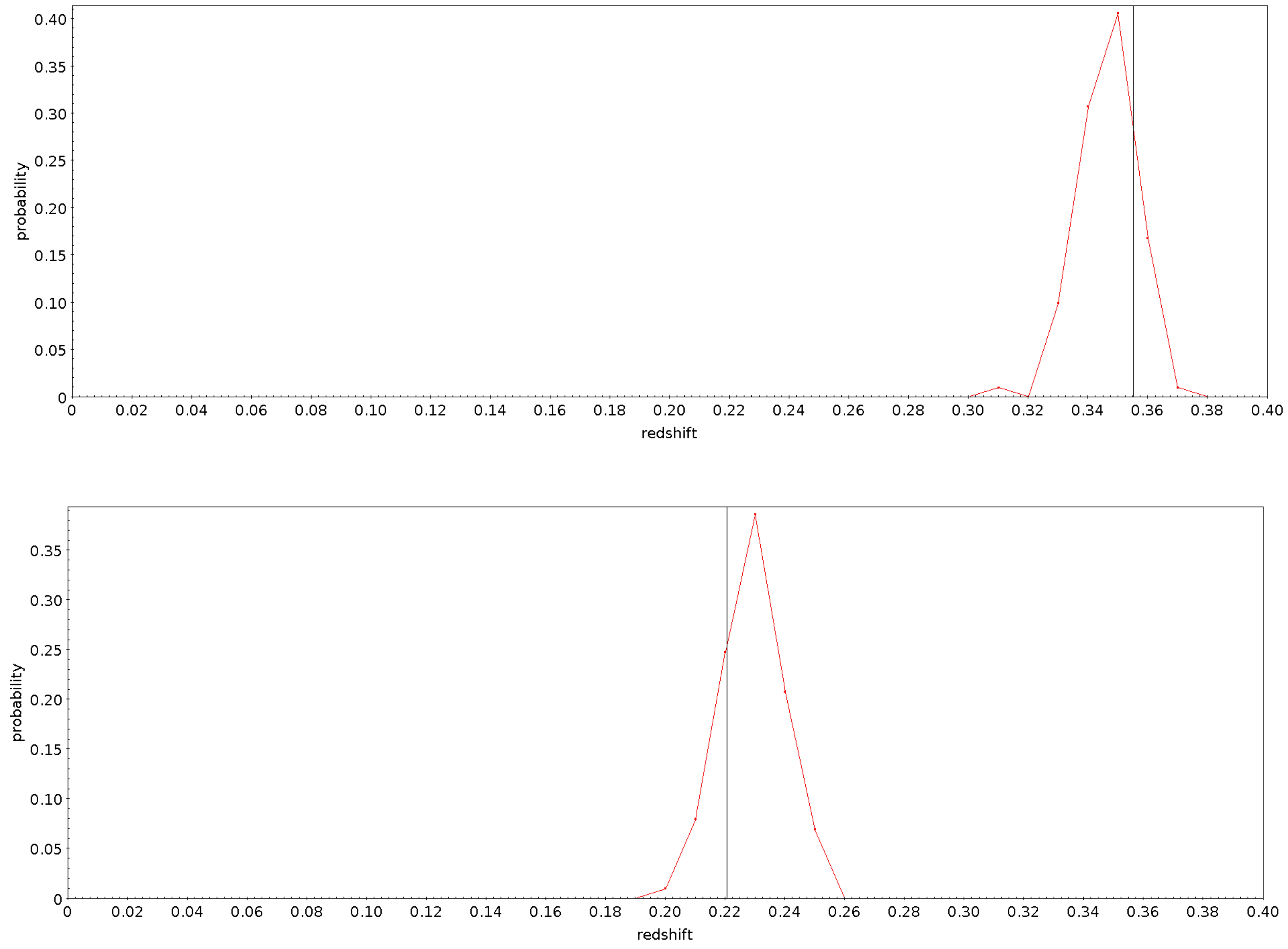
The class 4 (z>=1.075) shows a different behavior due, as expected, to the under-sampled spectroscopic KB which causes a lower quality of photo-z estimation and of the derived PDF.

By considering again a PDF bin of 0.03:

- The 6% have their zspec within the peak of the PDF;
- While the 52%, have zspec within the PDF. By considering also the bin closest to the PDF, the percentage grows up to 58%.



PDF examples



Thanks!