5^ EUCLID ITALIAN MEETING

Giulia Rodighiero (*University of Padova)*

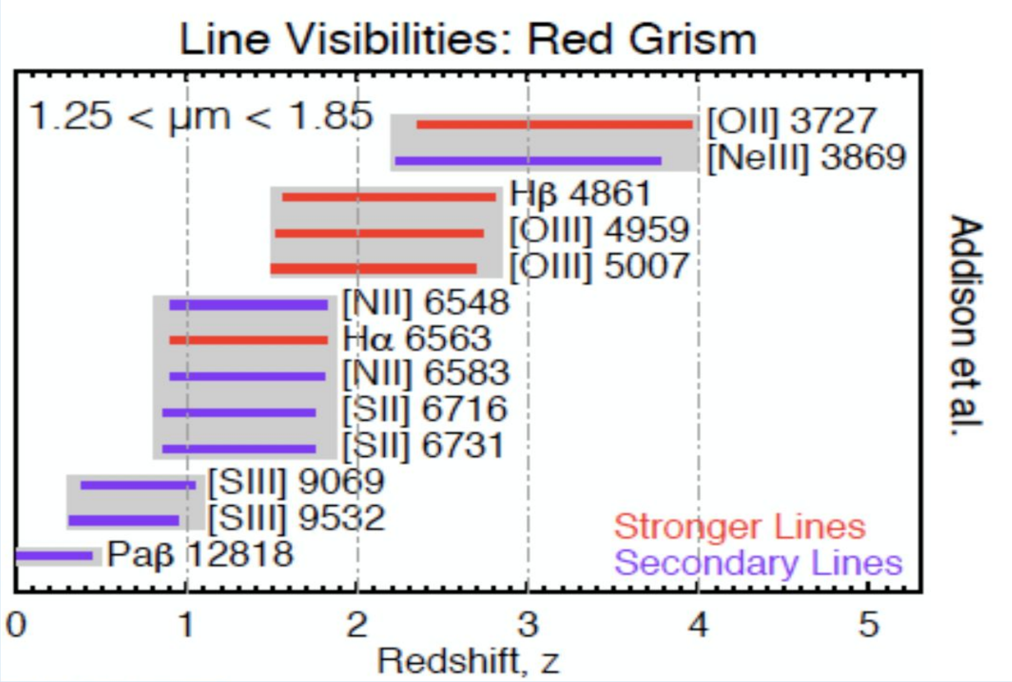
for

GAEV-SWG
WP10 – high-z galaxies

24th of February 2022

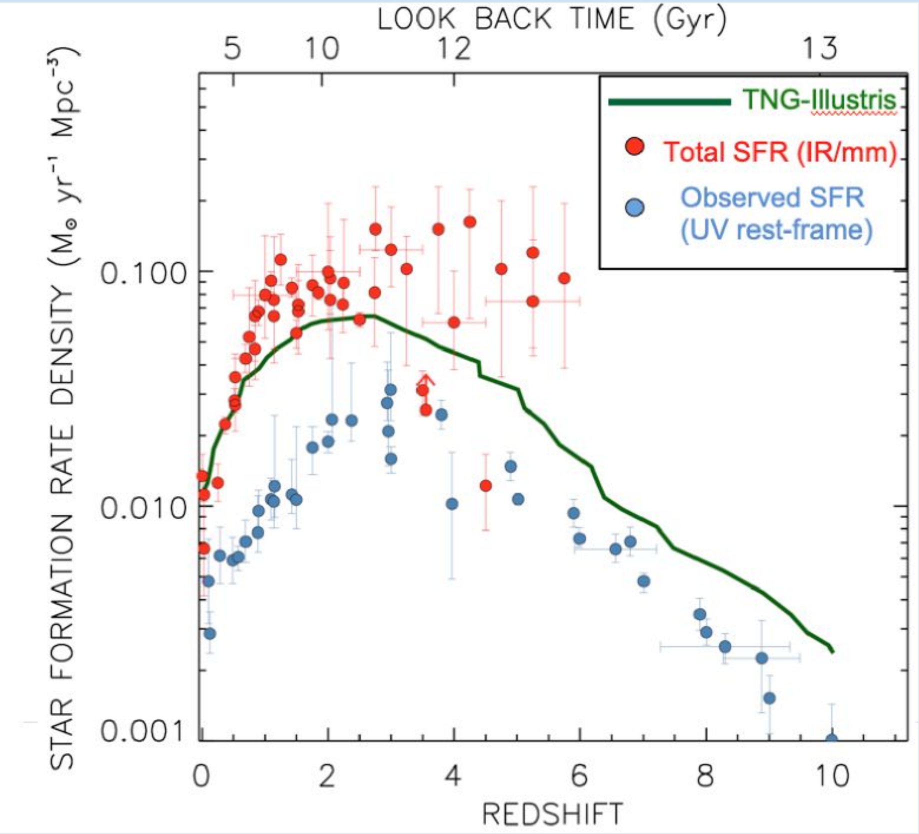
WP10 identity:

- start of activities: December 2020
- ~30 participants
- four telecons
- high-z galaxies ⇒ above z~4 and below z~7!!



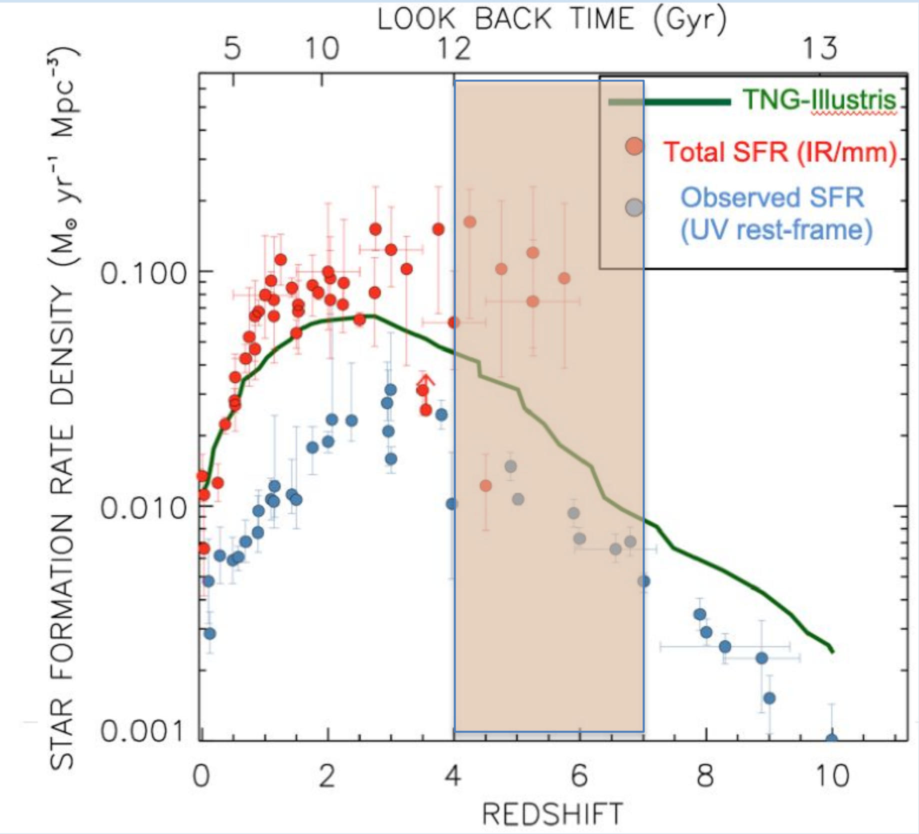
Goal of this Euclid SWG Galaxy Evolution Work Package:

To understand the potential contribution of Euclid in revealing a class of sources that are likely to represent the bulk population of massive galaxies that have been missed from previous surveys and are probably the progenitors of the largest present-day galaxies in massive groups and clusters.

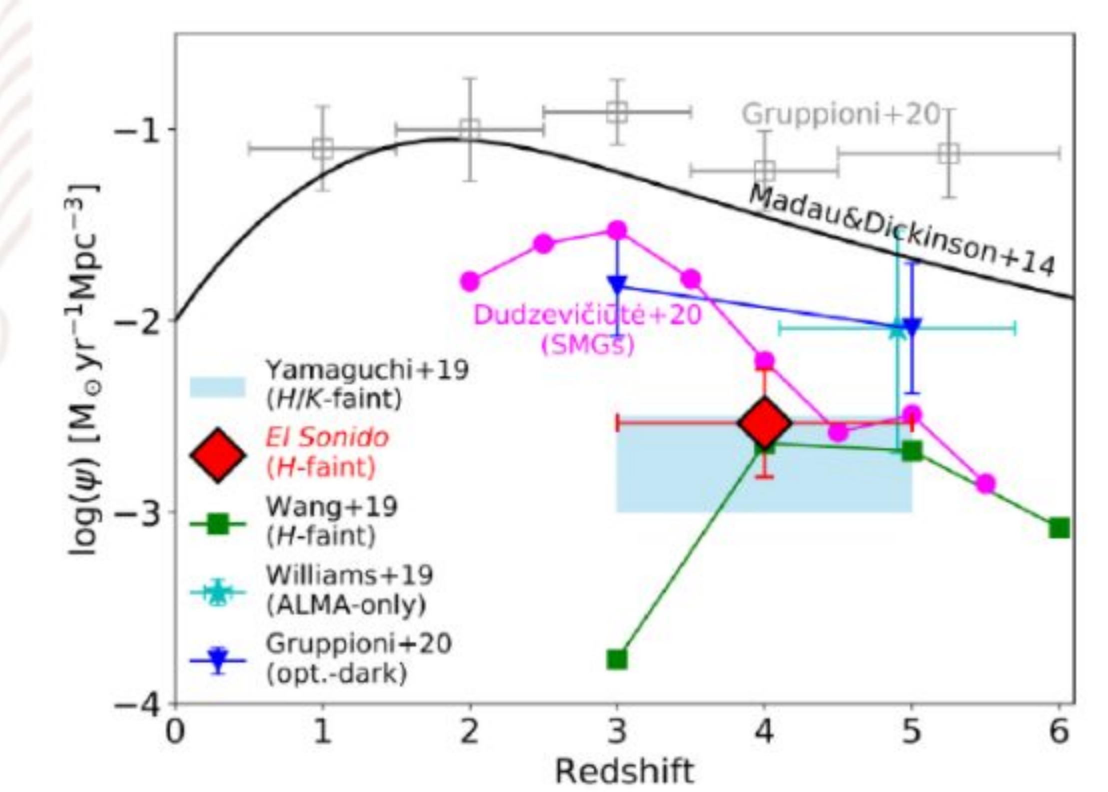


Goal of this Euclid SWG Galaxy Evolution Work Package:

To understand the potential contribution of Euclid in revealing a class of sources that are likely to represent the bulk population of massive galaxies that have been missed from previous surveys and are probably the progenitors of the largest present-day galaxies in massive groups and clusters.

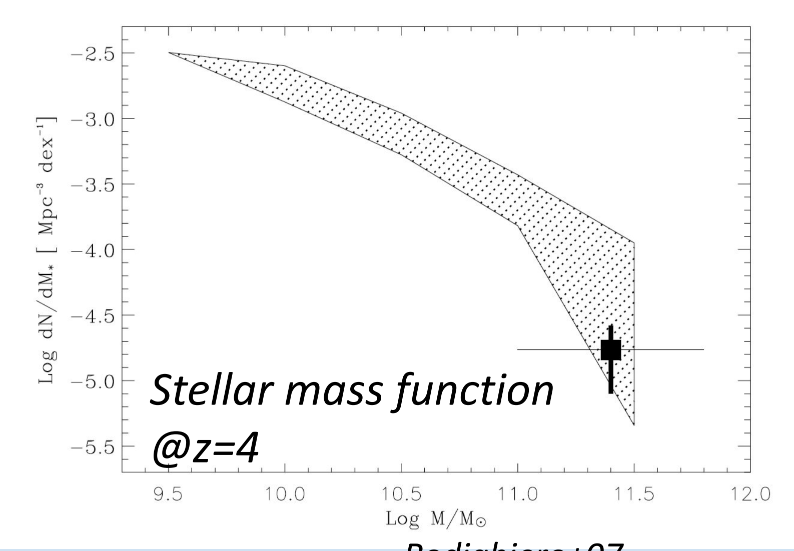


Contribution of "HST" dark sources to the stellar mass density (selection from IRAC, ALMA, radio….):



Sun+20 (but see also Talia+20, Enia+22)

Different population from LBGs!



Rodighiero+07

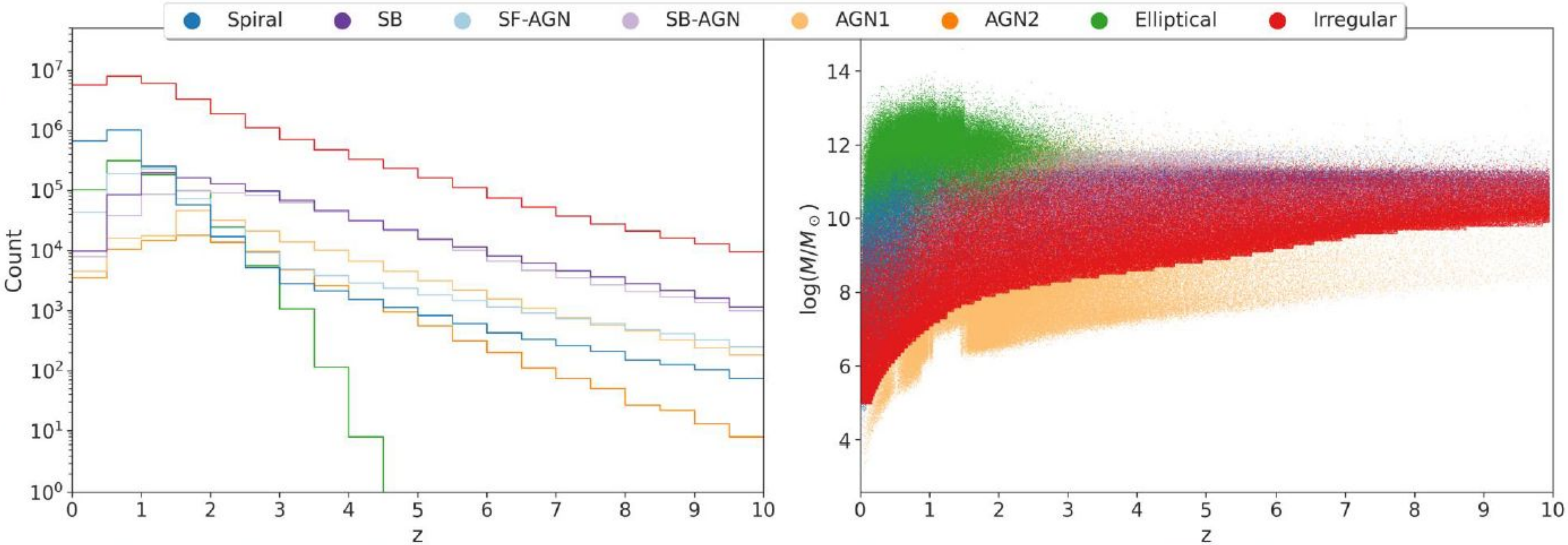
These dusty and massive galaxies show remarkable star formation activity but are **very rare and faint ⇒ Need for Deep and Wide near-IR surveys to statistically recover this population -> Euclid Deep Survey**

High-z populations in the Euclid redshift desert (always thought in combination to ancillary obs):

- Photometric redshifts (including Machine Learning)
- Colours
- Drop-outs
- Line emitters embedded in broad-band photometry

The Euclid Deep Fields Simulated Catalog

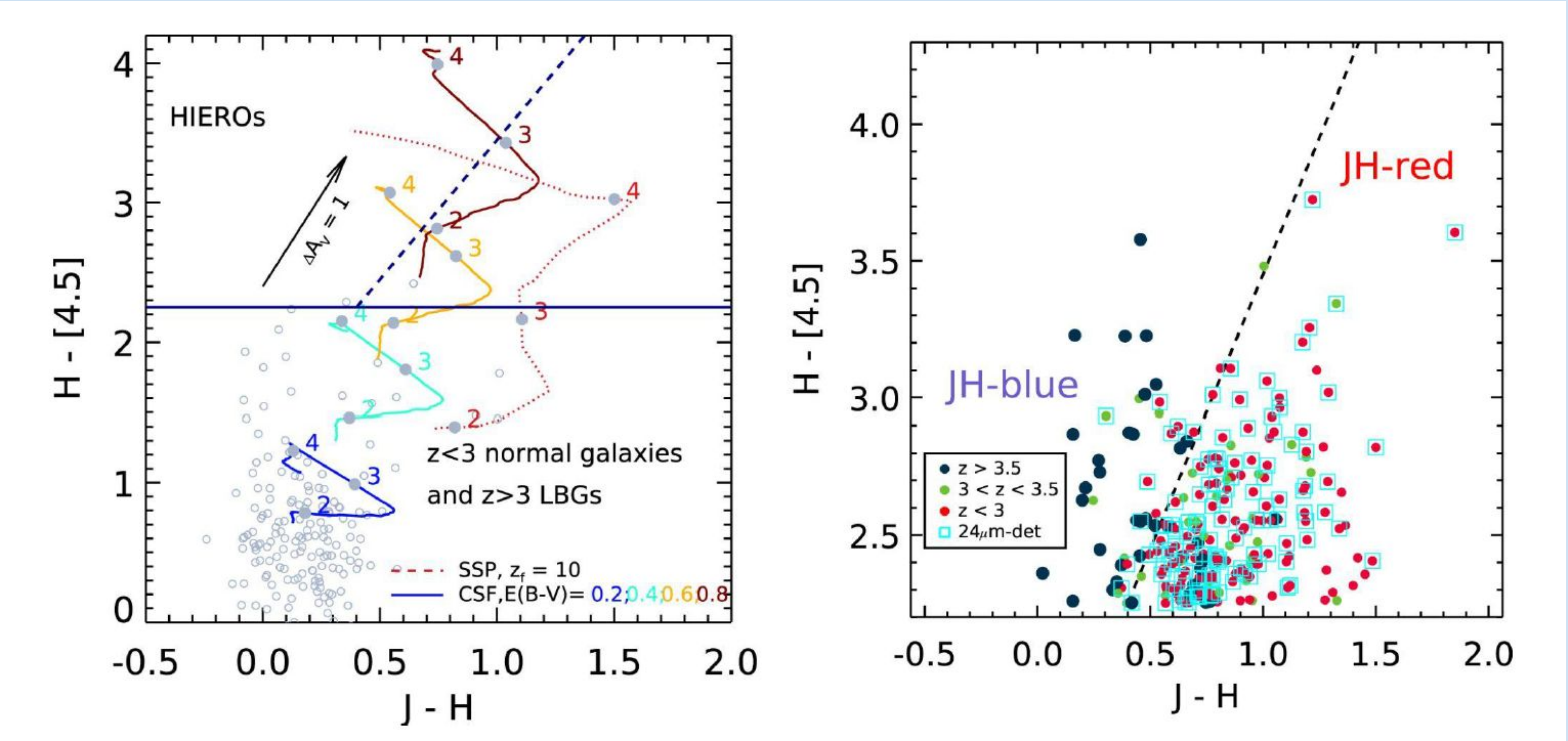
The Euclid Deep Survey combination of depth and area results in a simulated catalog with a total of more than 30 million objects with redshift from z ≈ 0 to ≈ 10



⇒ SPRITZ simulation (Bisigello et al. 2021)

•The simulation is built from the Herschel infrared luminosity functions of different galaxy populations, and is based on a wide set of empirical relations to associate a spectral energy distribution and physical properties to each simulated galaxy.

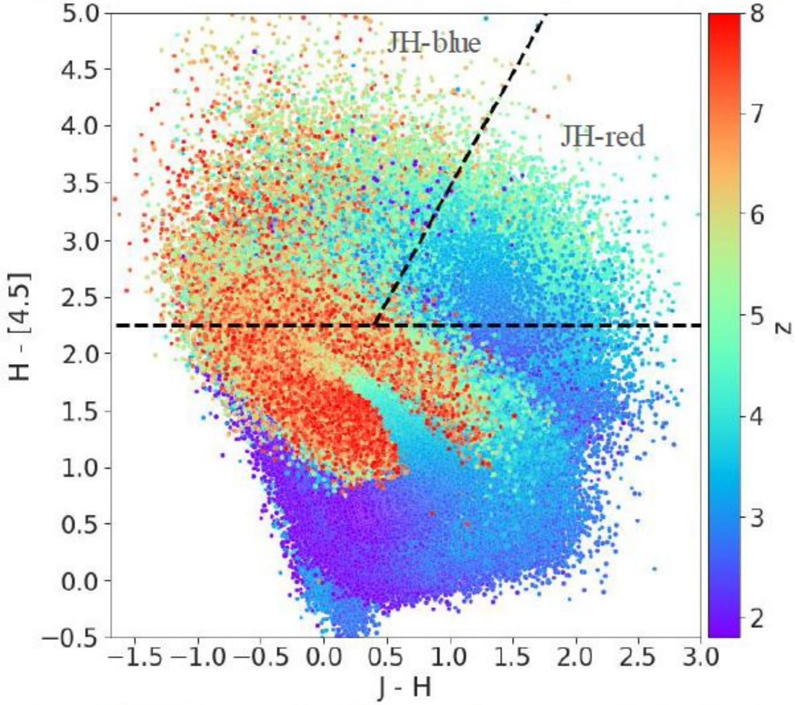
First assessment on HIEROS (Wang+16)

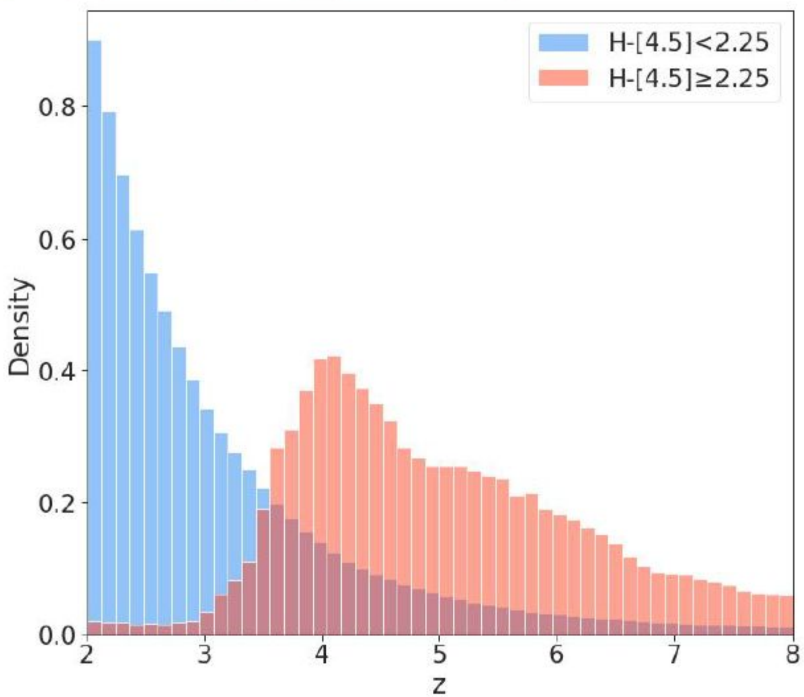


Photometric Selections

MOSTLY BASED ON THE MASTER THESIS OF *THEO SIGNOR* IN PHYSICS OF DATA (UniPD)

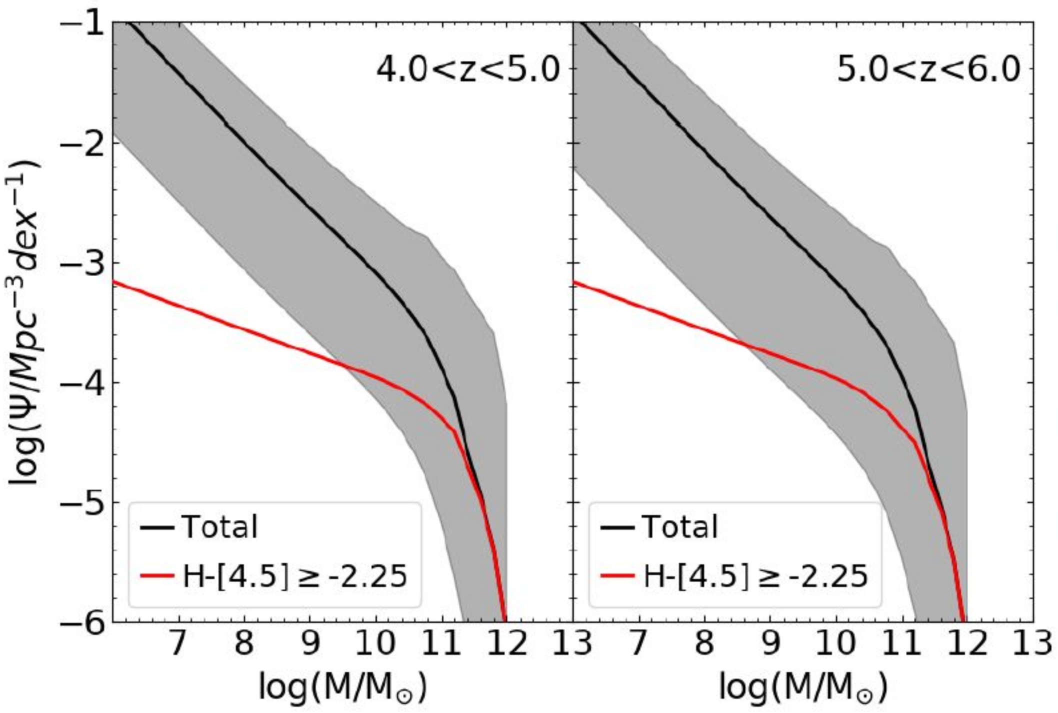
- First, we check the **simulated catalog** compatibility with a set of observed photometric diagnostics available from the literature (Laigle et al. 2016; Daddi et al. 2004; Wang et al. 2016; van Mierlo et al. 2022, in prep.)
- In particular, we check the **distributions of magnitudes**, **SED types** and **redshifts**, as a function of different color-color plots.
- HIEROs (extremely red objects; old or dusty galaxies at z>3) color selection: H-[4.5]>2.25





Photometric Selections

- First, we check the **simulated catalog** compatibility with a set of observed photometric diagnostics available from the literature (Laigle et al. 2016; Daddi et al. 2004; Wang et al. 2016; van Mierlo et al. 2022, in prep.)
- In particular, we check the **distributions of magnitudes**, **SED types** and **redshifts**, as a function of different color-color plots.
- HIEROs (extremely red objects; old or dusty galaxies at z>3) color selection: H-[4.5]>2.25



courtesy Laura Bisigello

Photo-z - Data

Gradient Boosted Trees (XGBoost) are implemented to predict the redshift of galaxies within the Euclid Deep survey simulated catalog, based on multi-band photometry.

The Dataset consists of

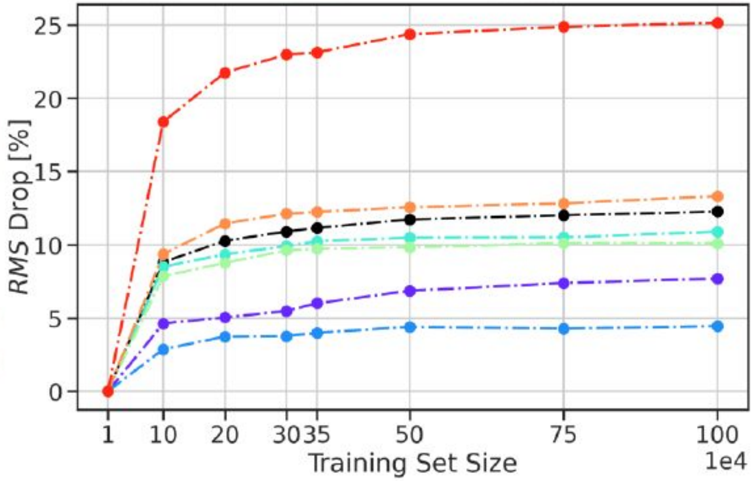
- Fluxes in 11 bands: VIS, NISP/Y, NISP/J, NISP/H, Rubin/u, Rubin/g, Rubin/r, Rubin/i, Rubin/z,IRAC/3.6µm, IRAC/4.5µm bands;
- Redshift z
- SED Type

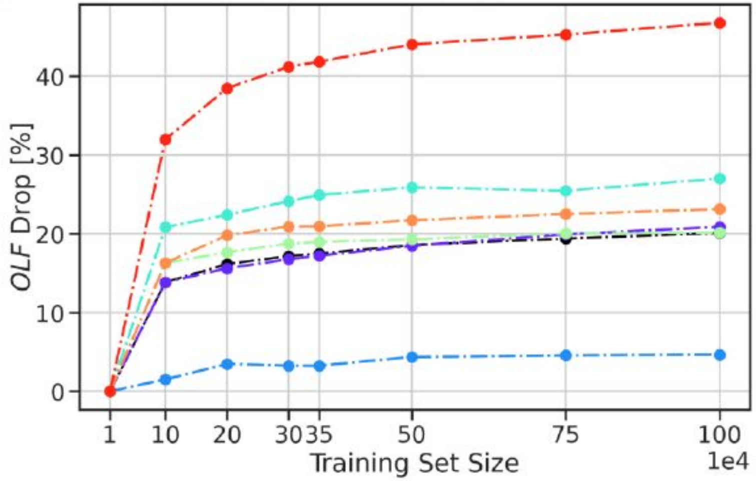
Band	5σ **Depth**	2σ **Depth**
VIS	28.2	29.2
NISP/Y	26.3	27.3
NISP/J	26.5	27.5
NISP/H	26.4	27.4
Rubin/u	25.15	26.1
Rubin/g	26.35	27.34
Rubin/r	26.45	27.4
Rubin/i	25.75	26.7
Rubin/z	25	26
IRAC/3.6μm	24.5	25.5
IRAC/4.5μm	24.5	25.5

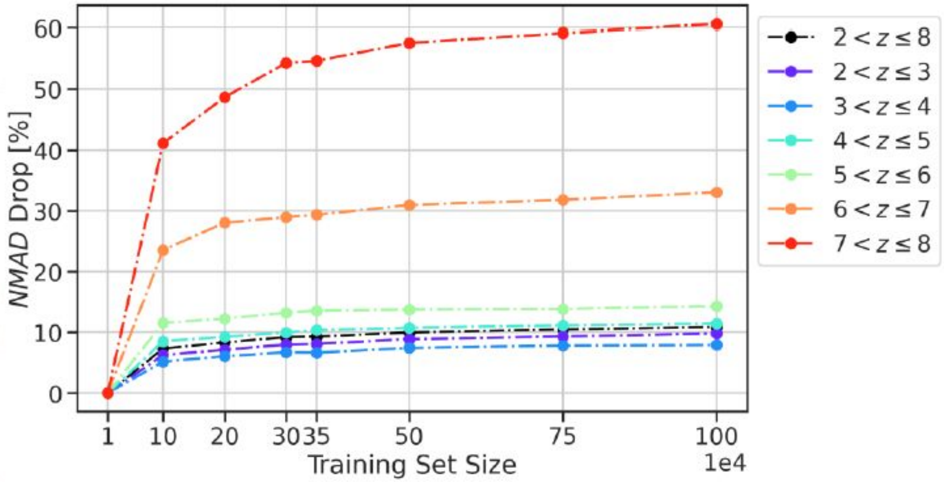
Photo-z - Training Set Size

FULL SAMPLE

In real-world observations, one will have no choice regarding the size of the training set. However, when forecasting future surveys observations, it is useful to assess what **dimension of the training set** is required to obtain a certain **redshift prediction performance**.







$$RMS = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(z_i - \tilde{z}_i)^2}$$

$$NMAD = 1.48 \times median\left(\frac{|z - \tilde{z}|}{1 + z}\right)$$

$$OLF = \frac{1}{m} card\Big(\{\tilde{z} \,:\, |z - \tilde{z}|/(1 + z) > 0.15\}\Big)$$

Photo-z - Results

Following a Bayesian optimization for the xgboost hyperparameters, the test set performances are evaluated for different groups of objects

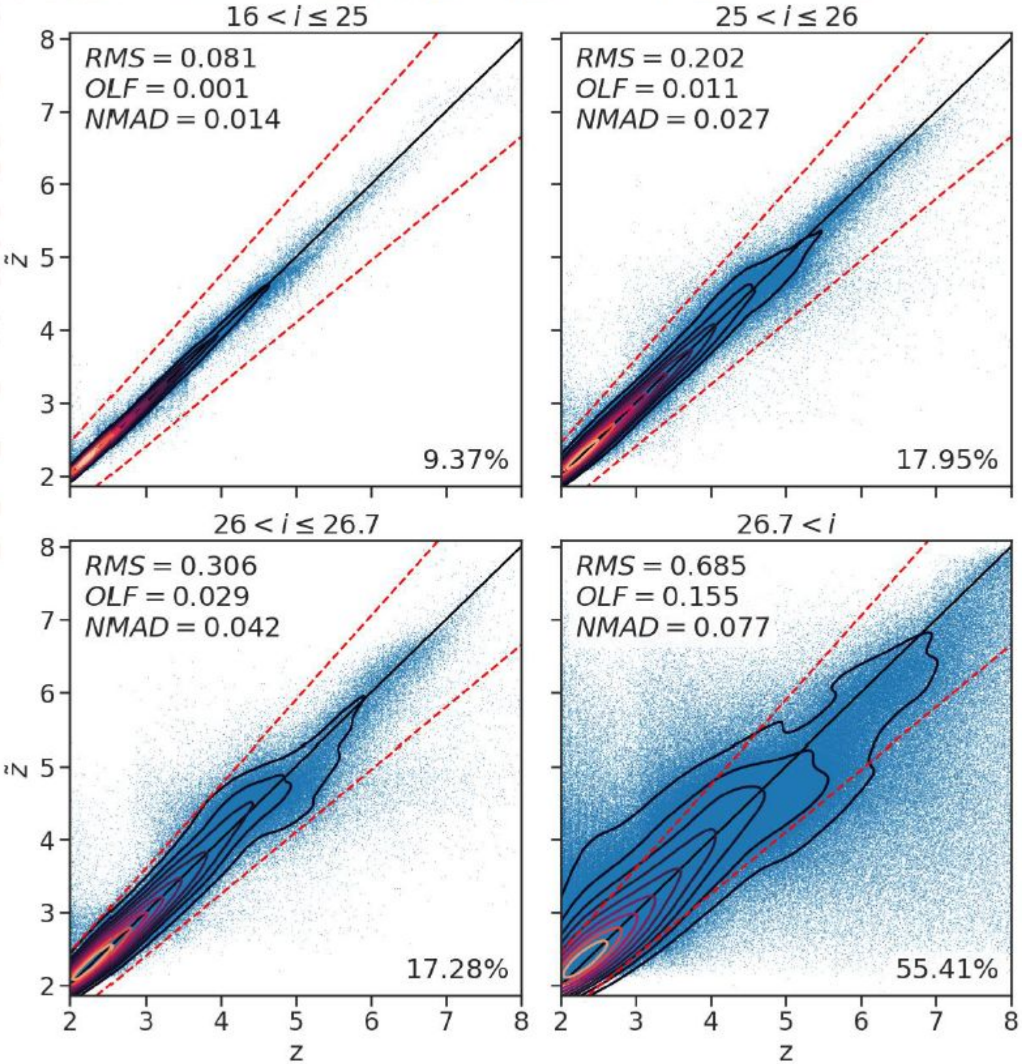


Photo-z - Results

FULL SAMPLE

Following a Bayesian optimization for the xgboost hyperparameters, the test set performances are evaluated for different groups of objects

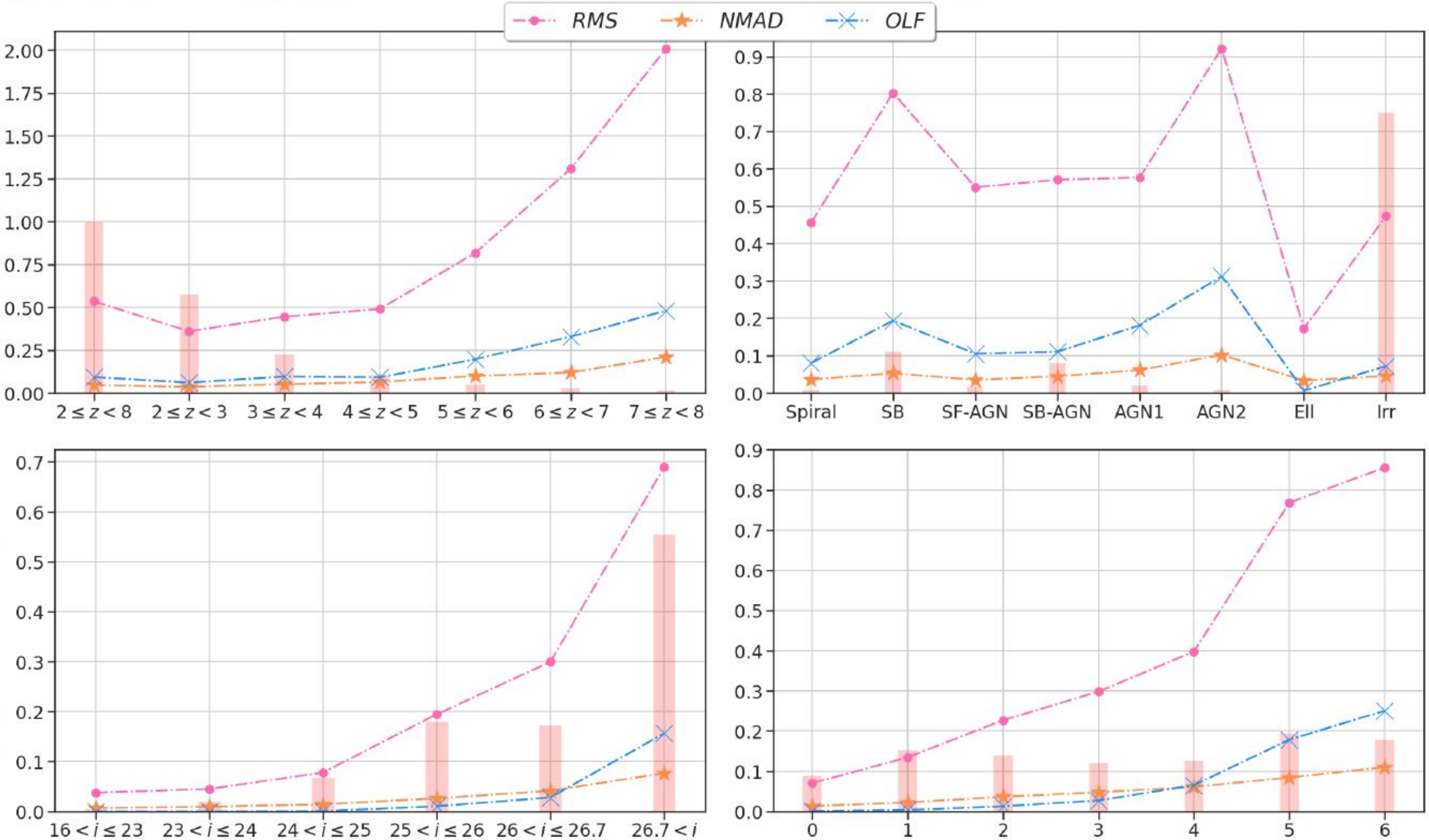


Photo-z - Results

FULL SAMPLE

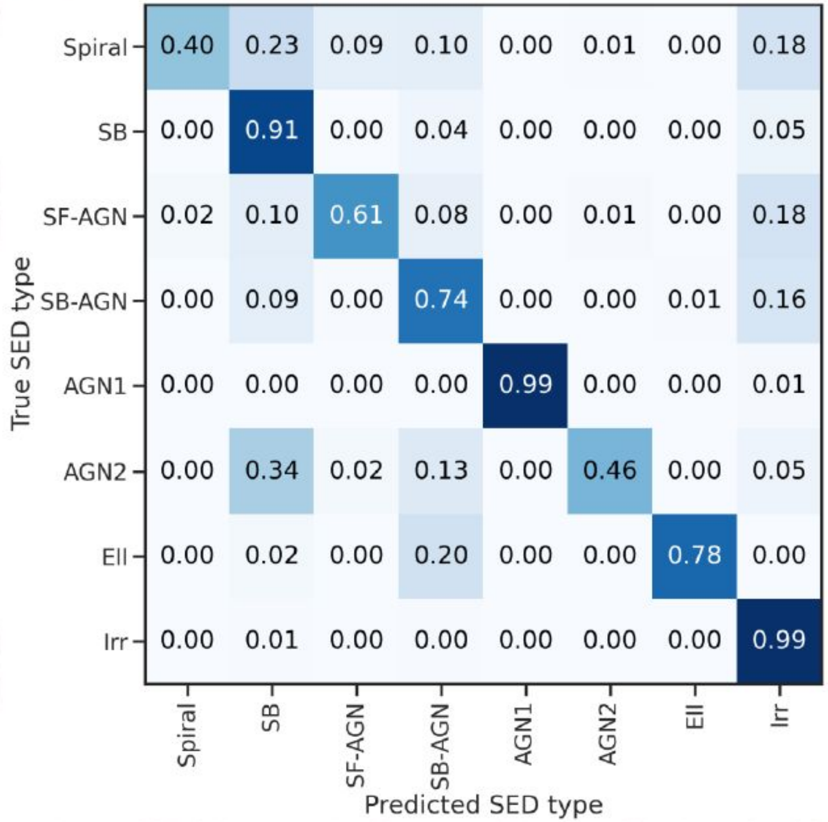
- To give a contest to the results obtained, they are **compared** to previous photometric redshift performances, in a similar z-range.
- This comparsion is made with the results reported in **Laigle 2016**, where the precision of photometric redshifts obtained via the SED fitting technique using 28 bands over the COSMOS2015 catalog is assessed againts spectroscopic ones
- Clearly the performances are very comparable

i-band Magnitude	**NMAD**			**OLF**		
	COSMOS Star-Forming	*COSMOS Quiescent*	*This Work*	*COSMOS Star-Forming*	*COSMOS Quiescent*	*This Work*
$16 < i \leq 21$	0.007	0.005	0.005	0.005	0.0	0.0
$21 < i \leq 22$	0.008	0.007	0.006	0.006	0.003	0.0
$22 < i \leq 23$	0.01	0.01	0.007	0.017	0.006	0.0
$23 < i \leq 24$	0.022	0.027	0.009	0.067	0.06	0.0
$24 < i \leq 27$	-	0.054	0.048	-	0.189	0.096

Spectral Type Classification - Results

FULL SAMPLE

- A gradient boosting approach was also taken to determine the SED type from photometry.
- **Test Set accuracy: 94.4%**
- This is clearly a simplified description, given the limited number of SED templates considered in the simulation



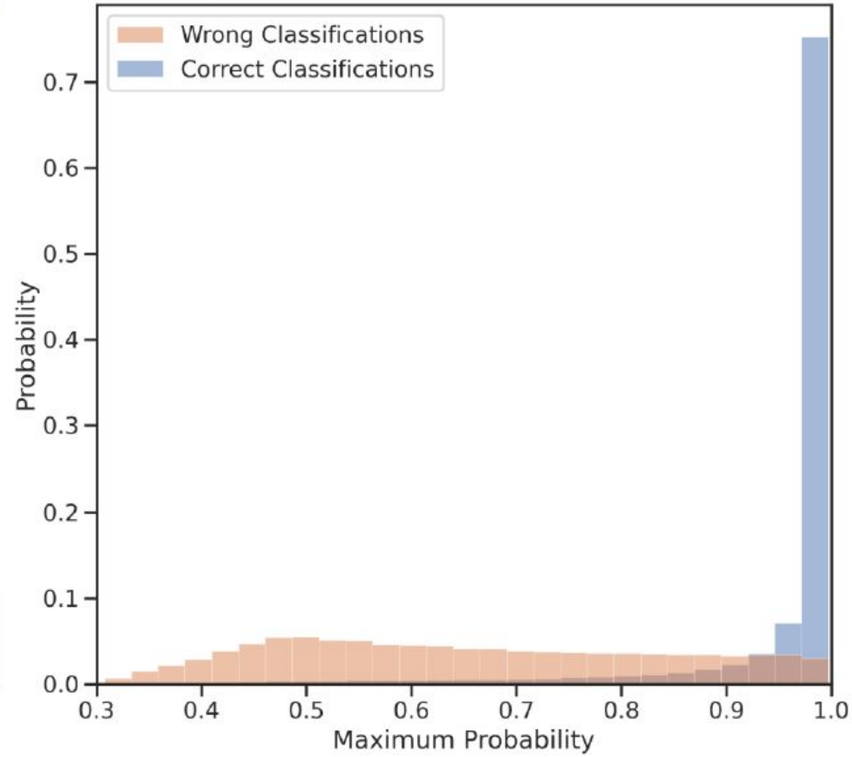
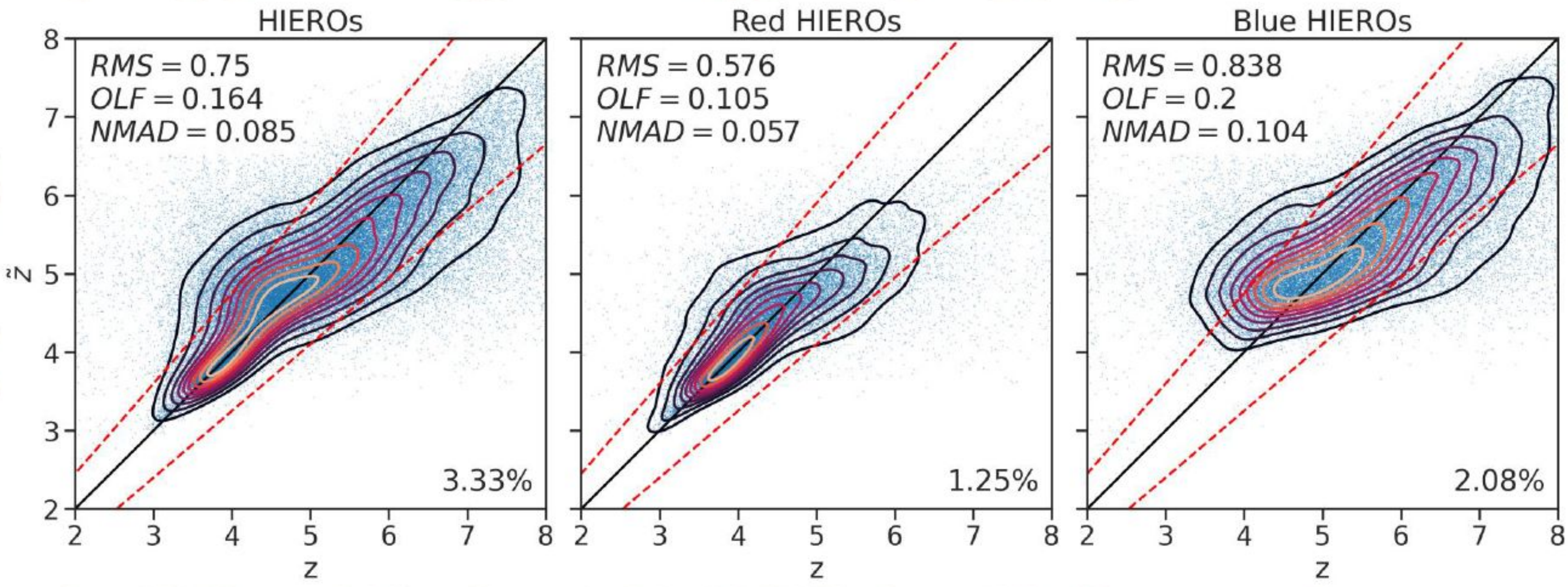


Photo-z - Results

HIEROS SAMPLE

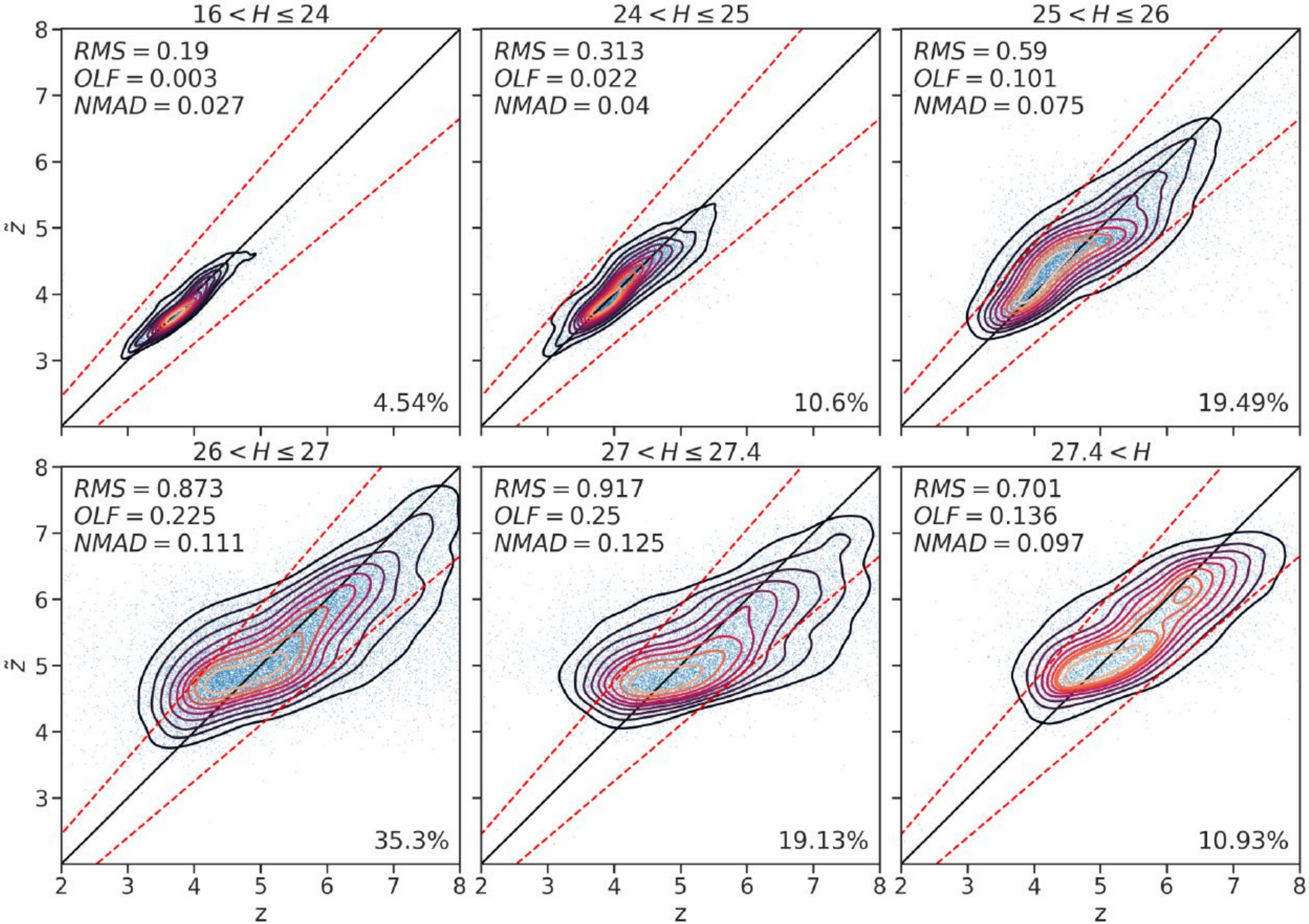
Particular focus in this work is on massive dusty galaxies, the **HIEROs** (H-[4.5]>2.25)



VS

Wang et al. 2016 HIEROs photometric redshifts

NMAD=0.11

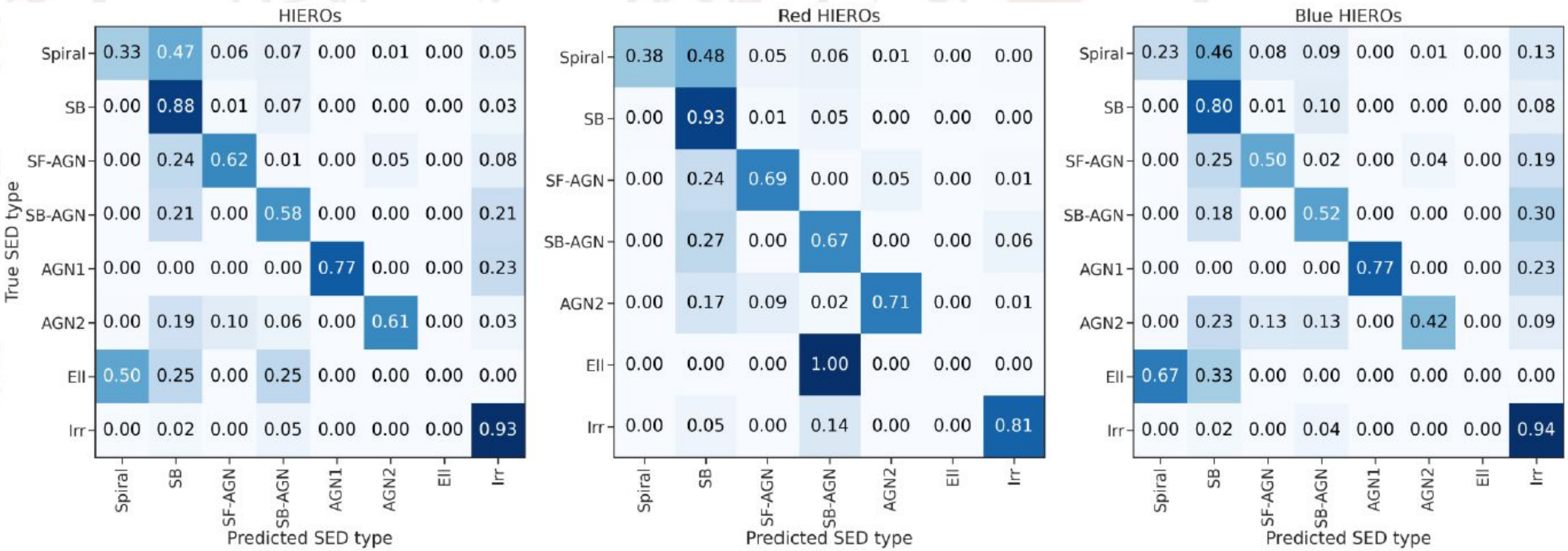


16 < H ≤ 24
RMS = 0.19
OLF = 0.003
NMAD = 0.027
4.54%
24 < H ≤ 25
RMS = 0.313
OLF = 0.022
NMAD = 0.04
10.6%
25 < H ≤ 26
RMS = 0.59
OLF = 0.101
NMAD = 0.075
19.49%
26 < H ≤ 27
RMS = 0.873
OLF = 0.225
NMAD = 0.111
35.3%
27 < H ≤ 27.4
RMS = 0.917
OLF = 0.25
NMAD = 0.125
19.13%
27.4 < H
RMS = 0.701
OLF = 0.136
NMAD = 0.097
10.93%
z̃
z

HIEROs Classification

HIEROS SAMPLE

- A gradient boosting approach was also taken to determine the SED type from photometry. In this case, the trained machine output, when a feature vector is fed to it, is a vector of probability (in this case with 8 entries, each one corresponding to a SED type). The predicted class is thus the one corresponding to the maximum value.



On going/future steps:
analysis of different simulations and photo-z from the standard Euclid pipeline (OU-photoz)

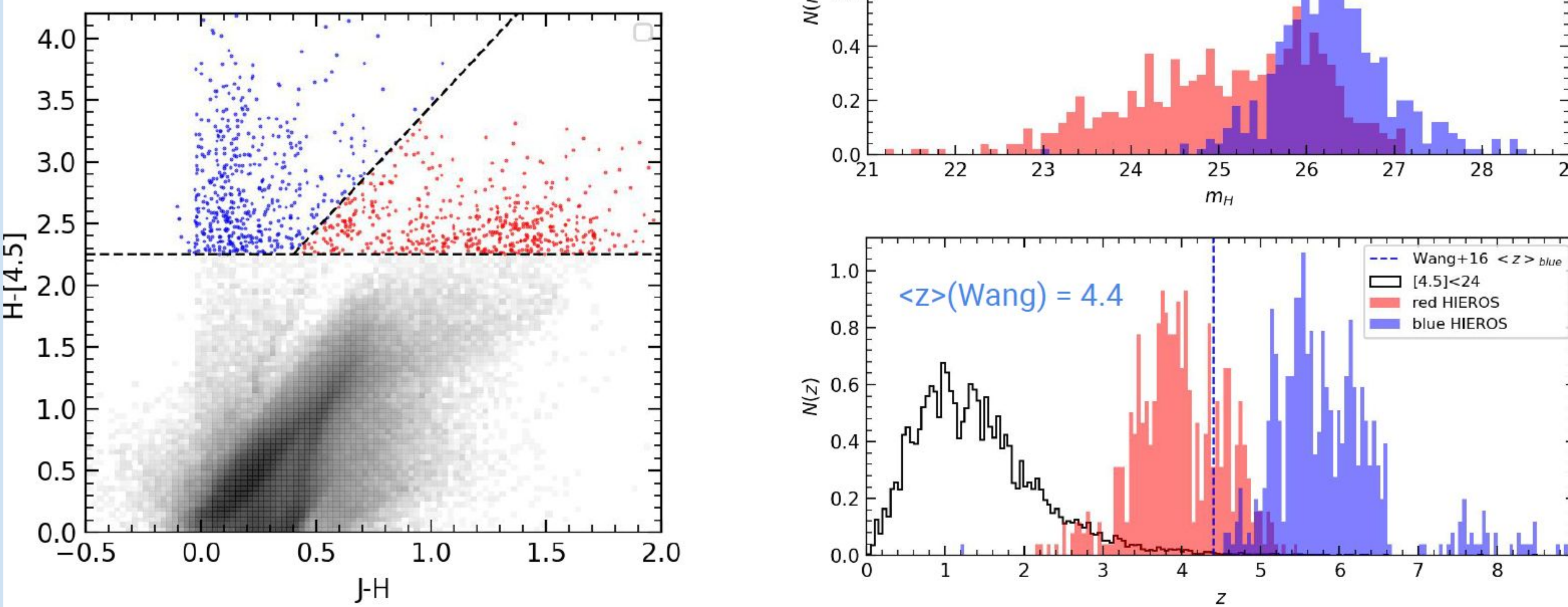
Selection in MAMBO lightcone (no photometric noise added)

[4.5] < 24.0

HIEROS: H-[4.5] > 2.25 → N = 1030 → n = 292/deg^2

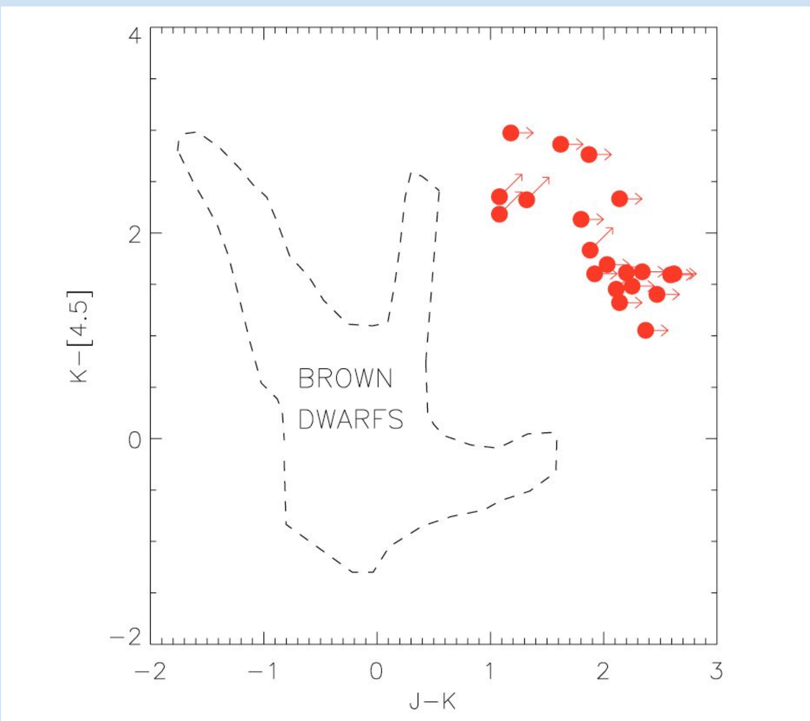
Red: H-[4.5] <= 2(J-H)+1.45 → N red HIEROS = 519

Blue: H-[4.5] > 2(J-H)+1.45 → N blue HIEROS = 511

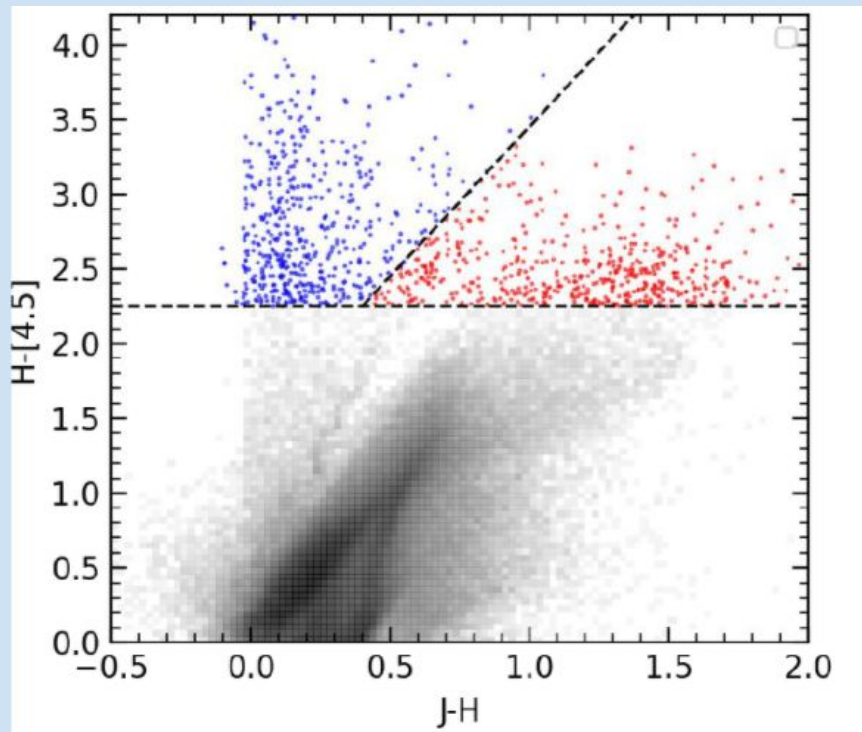


courtesy of Micol Bolzonella

On going/future steps:
Stellar contamination among the reddest drop-outs? check predictions from extensive Mock stellar catalogs (e.g. TRILEGAL, Girardi et al.)



Rodighiero et al. (2007)



Critical discussion with Primeval Universe / Cosmic Dawn team

- Taking advantage of people sitting in both Gal Evol (in particular this WP) and PU;
- Offline coordination with the WP leads in PU;
- Need to understand the depth of the multi-lambda photometric surveys available from Cosmic Dawn to address our science cases;
- Discussion about the PL-KPs from PU to see scientific and redshift overlaps with WP10 (and gal evol in general).