DARK galaxies in the era of EUCLID (...and JWST)

Giulia Rodighiero
(University of Padova)

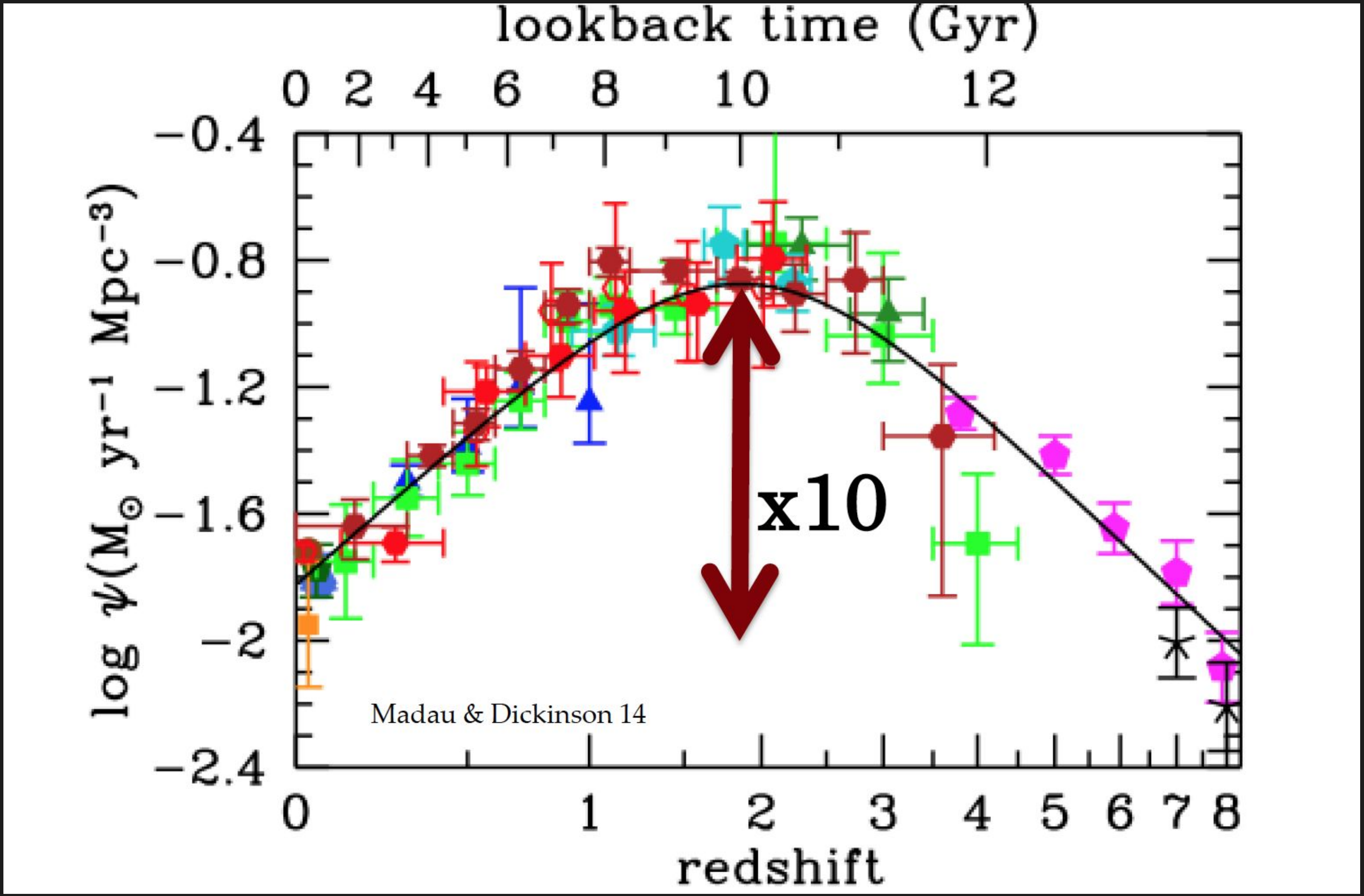
with Laura Bisigello, Theo Signor, Louis Gabarra, Chiara Mancini

DIPARTIMENTO DI FISICA E ASTRONOMIA "GALILEO GALILEI"

ESA-ESO Euclid workshop 26th of October 2022

UNIVERSITAS STUDII PADUANI · MCCXXII ·

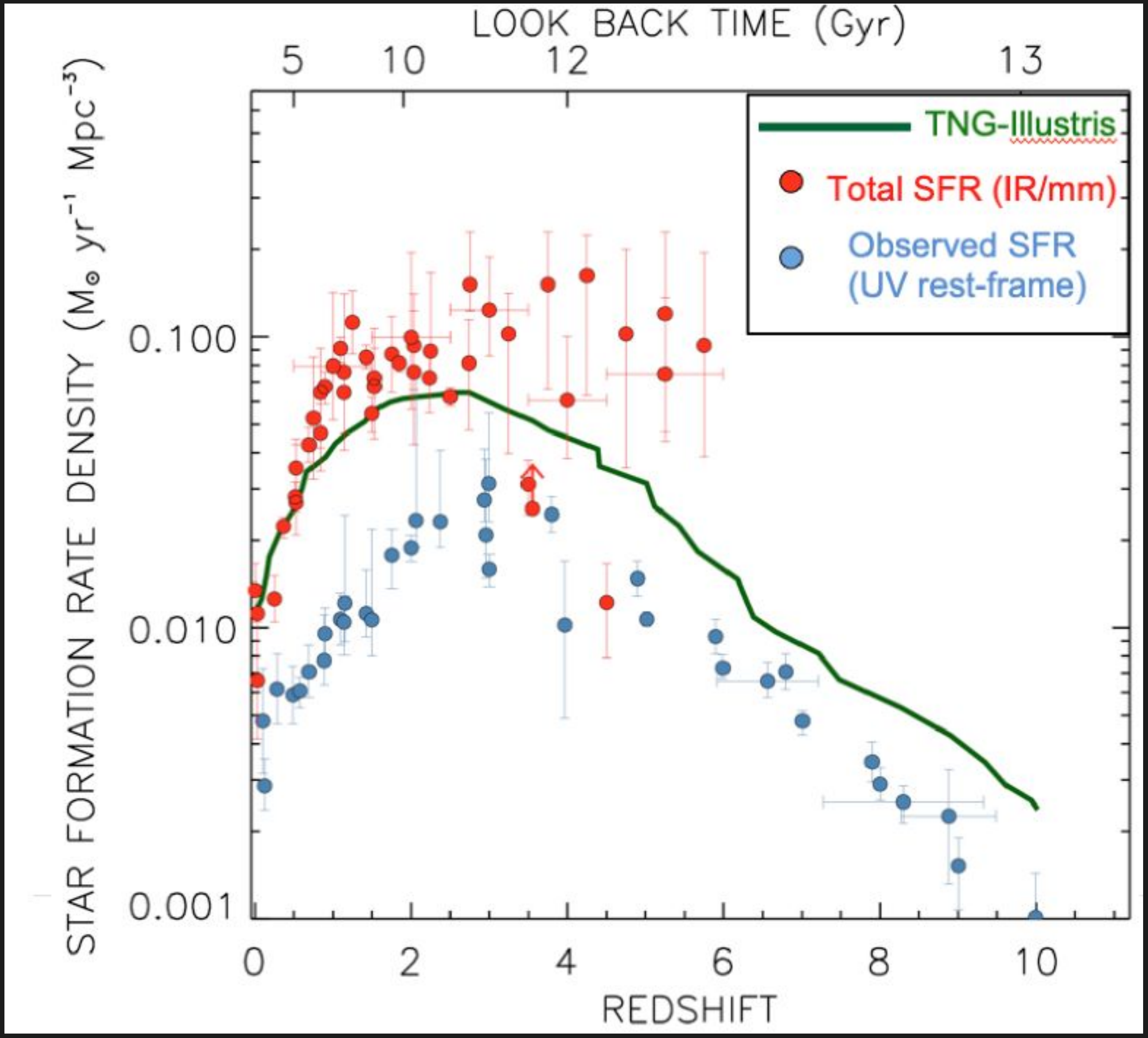
The cosmic SFRD in the *JWST* era



Mergers? Major? Minor?
(e.g. Somerville+2001, Conselice+2008)

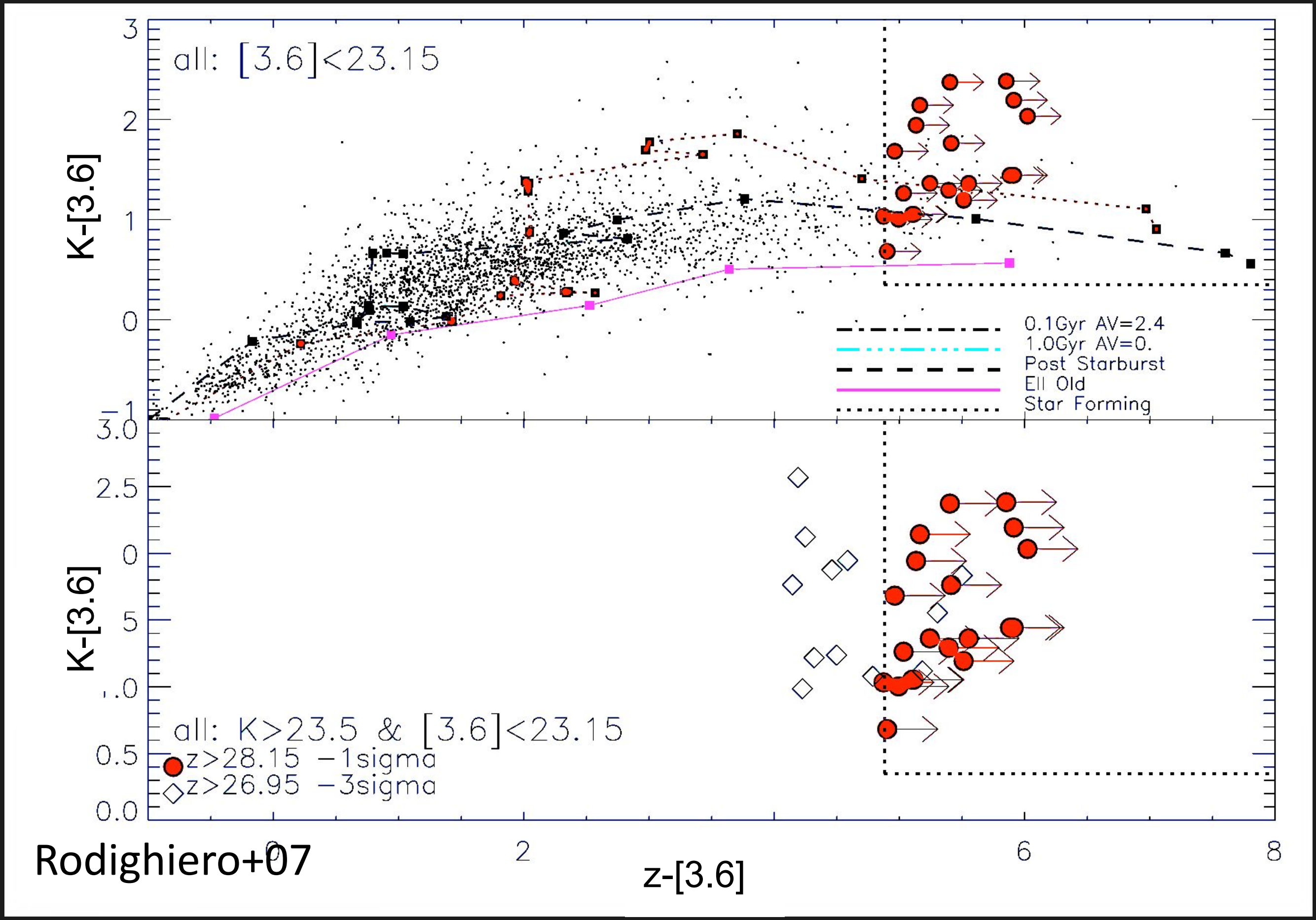
Secular processes? Gas infall rates?
(Keres+2005, Bower+2006, Dekel+2009)

The cosmic SFRD in the *JWST* era



Courtesy, Carlotta Gruppioni

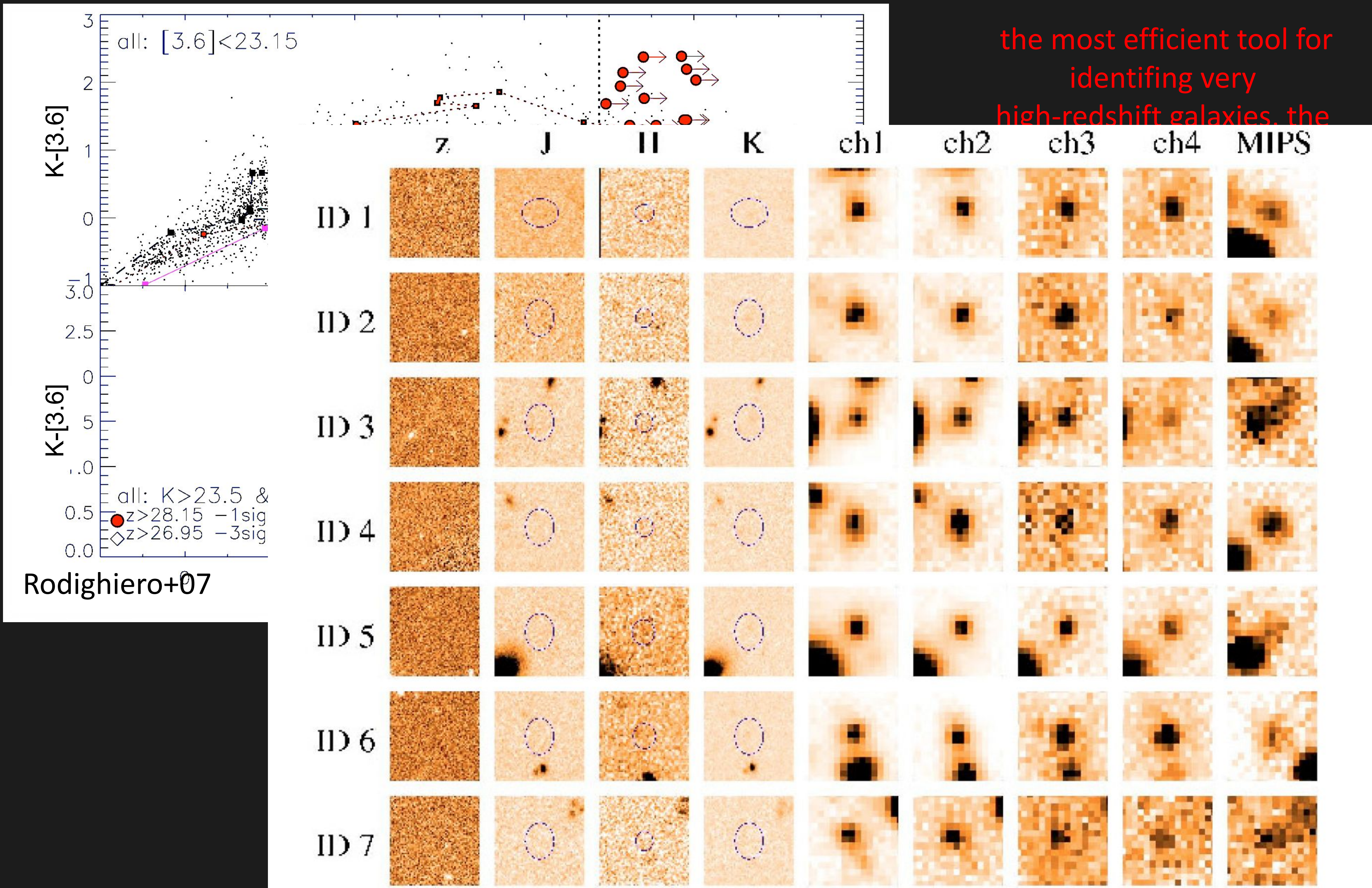
Selecting massive and dusty galaxies as missing HST objects



the most efficient tool for identifing very high-redshift galaxies, the Ly-dropout technique, is not sensitive to galaxy mass but rather to UV flux.

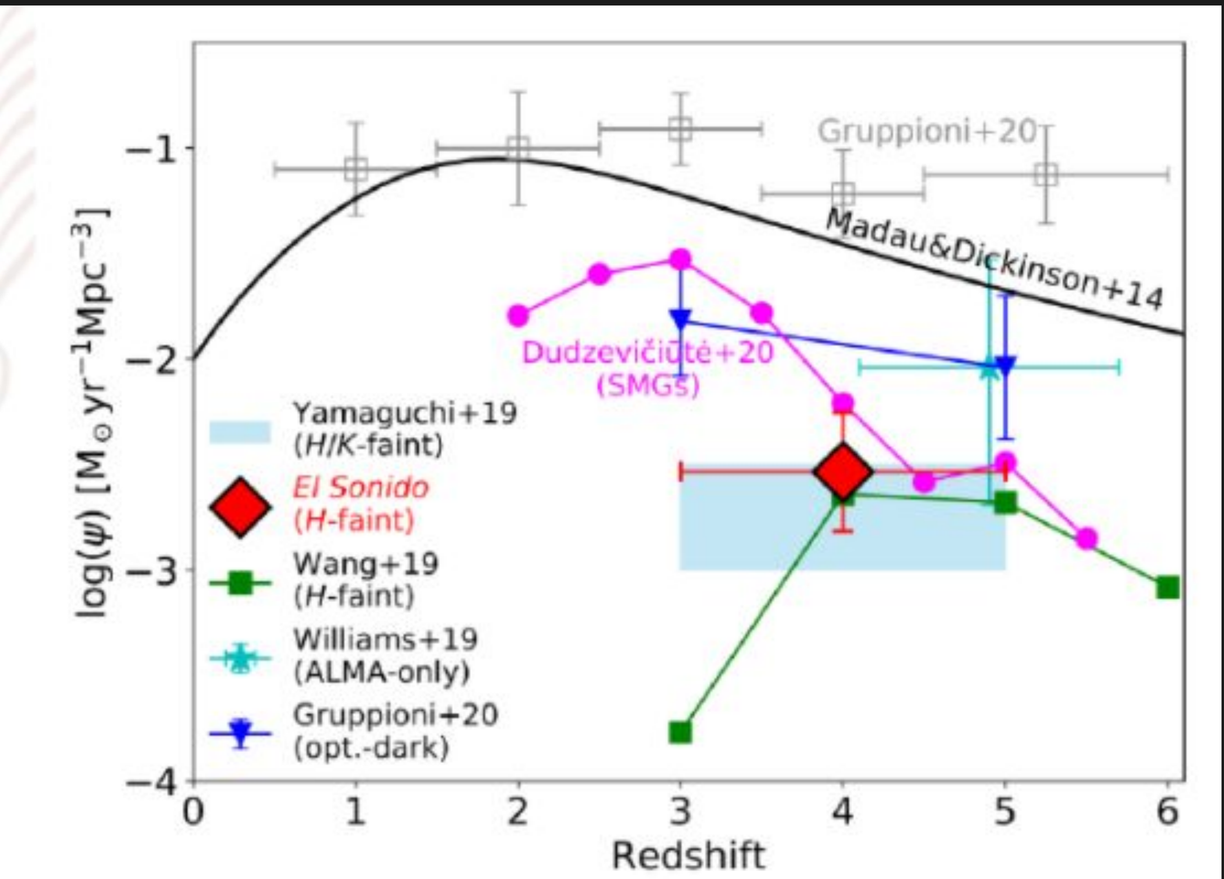
Selecting massive and dusty galaxies as missing HST objects

the most efficient tool for identifing very high-redshift galaxies, the



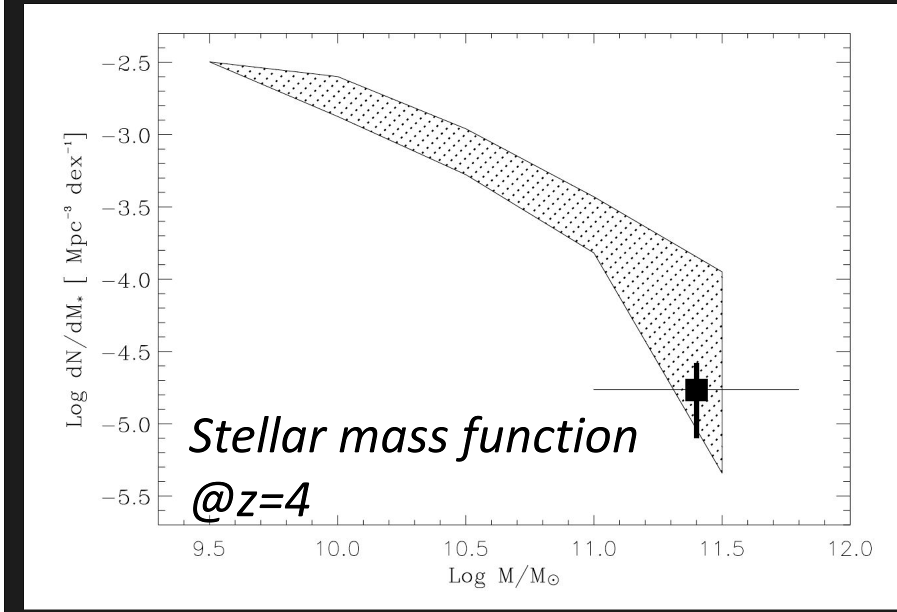
Rodighiero+07

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



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

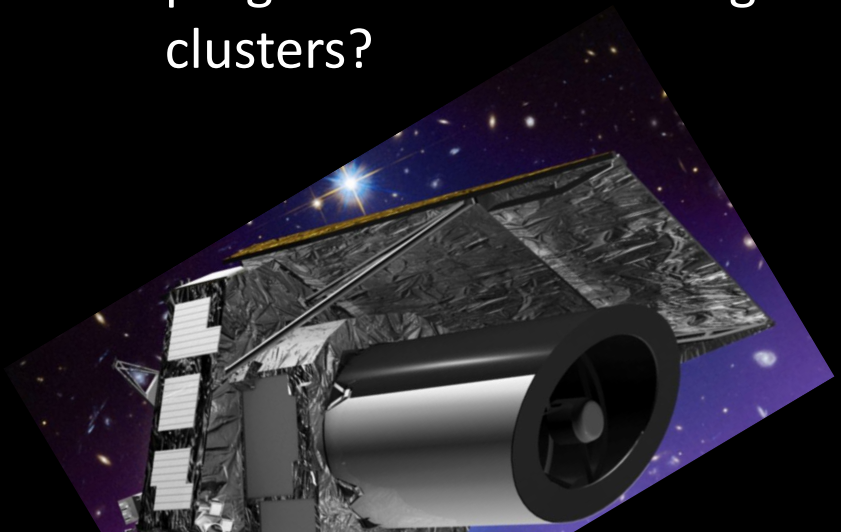
Different population from LBGs!

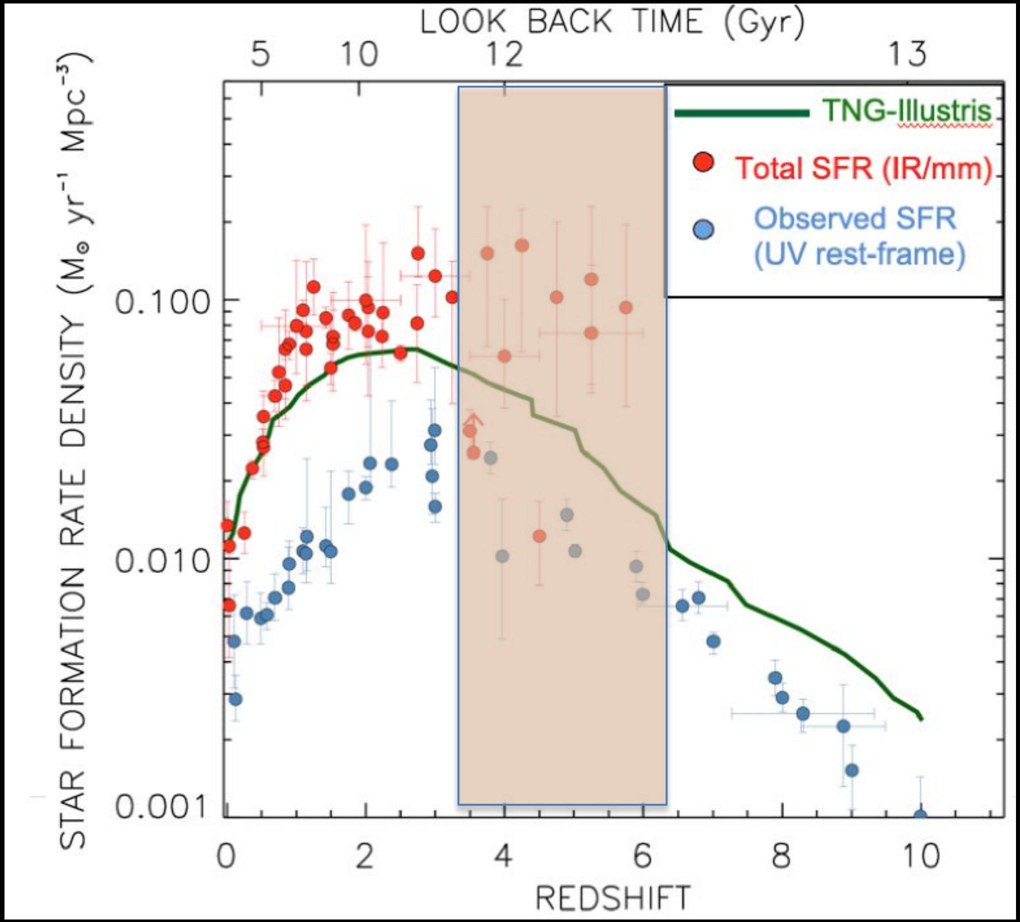


Rodighiero+07

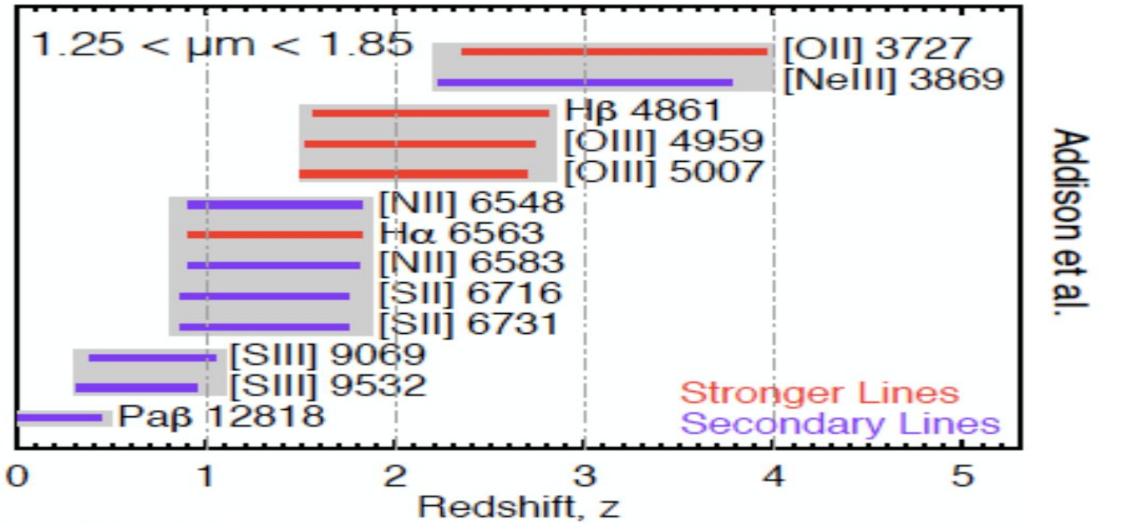
These dusty and massive galaxies show remarkable star formation activity but are **very rare and faint ⇒ Need for Deep and Wide near-IR surveys to statistically recover this population**

⇒ how 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?



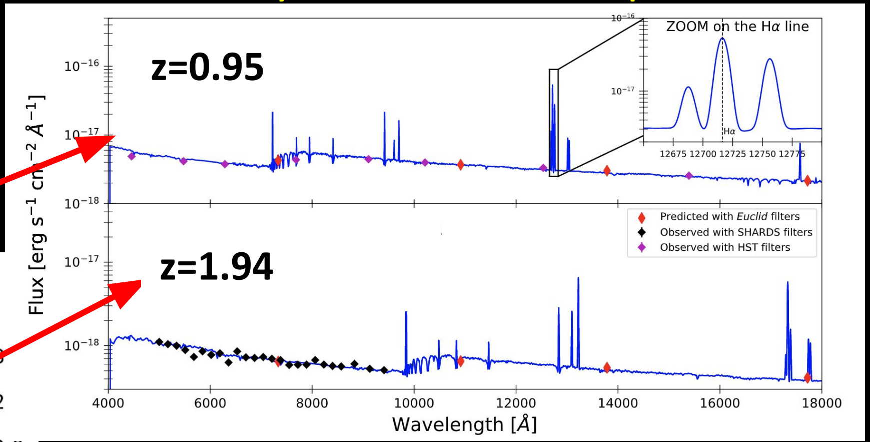


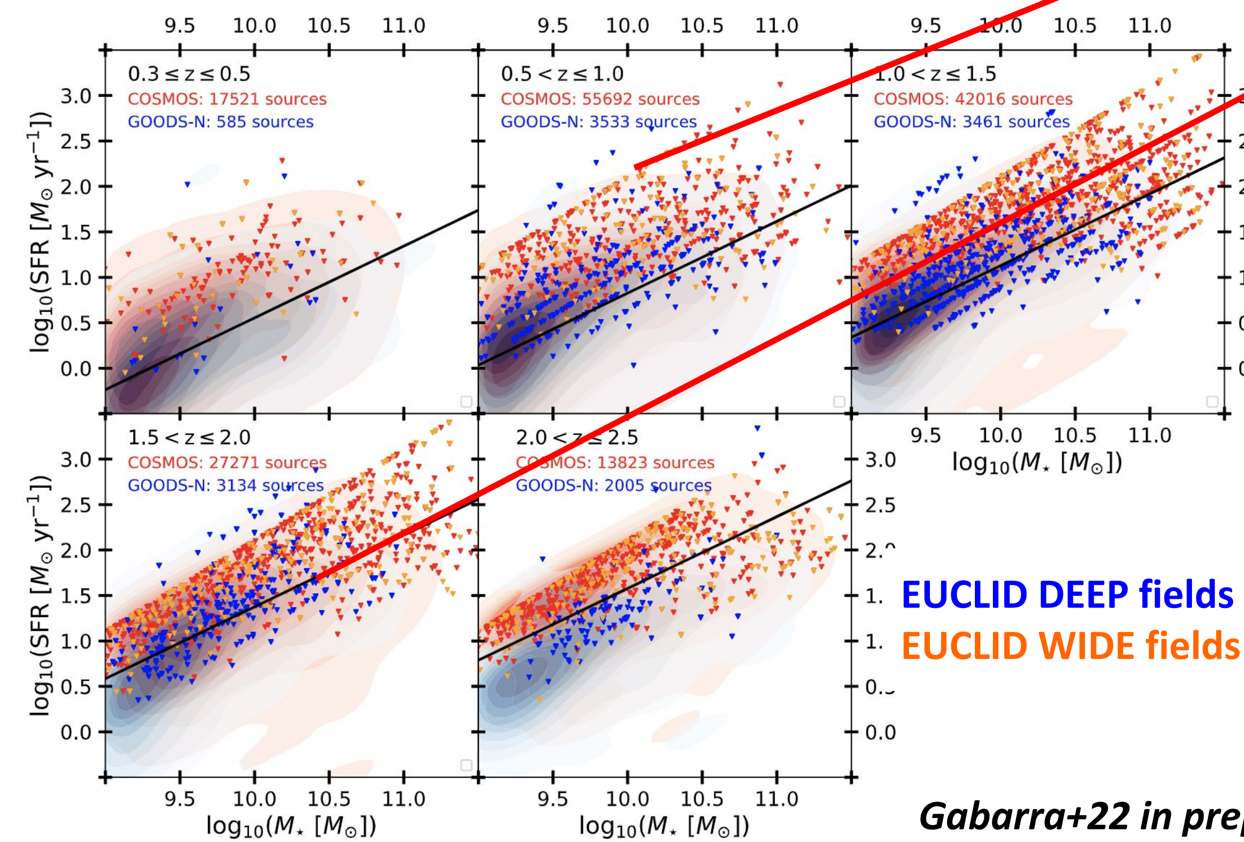
EUCLID LINE VISIBILITIES: RED GRISM



+ Lyman alpha at z>7 with Blue Grism

Purely emission lines spectra

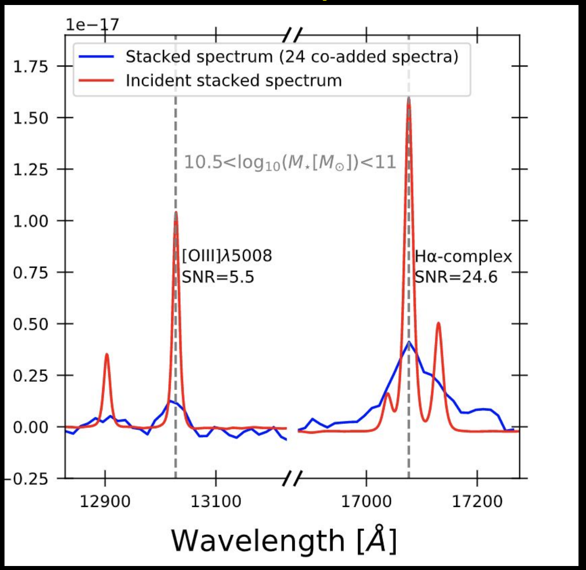




EUCLID DEEP fields

EUCLID WIDE fields

stacked spectra



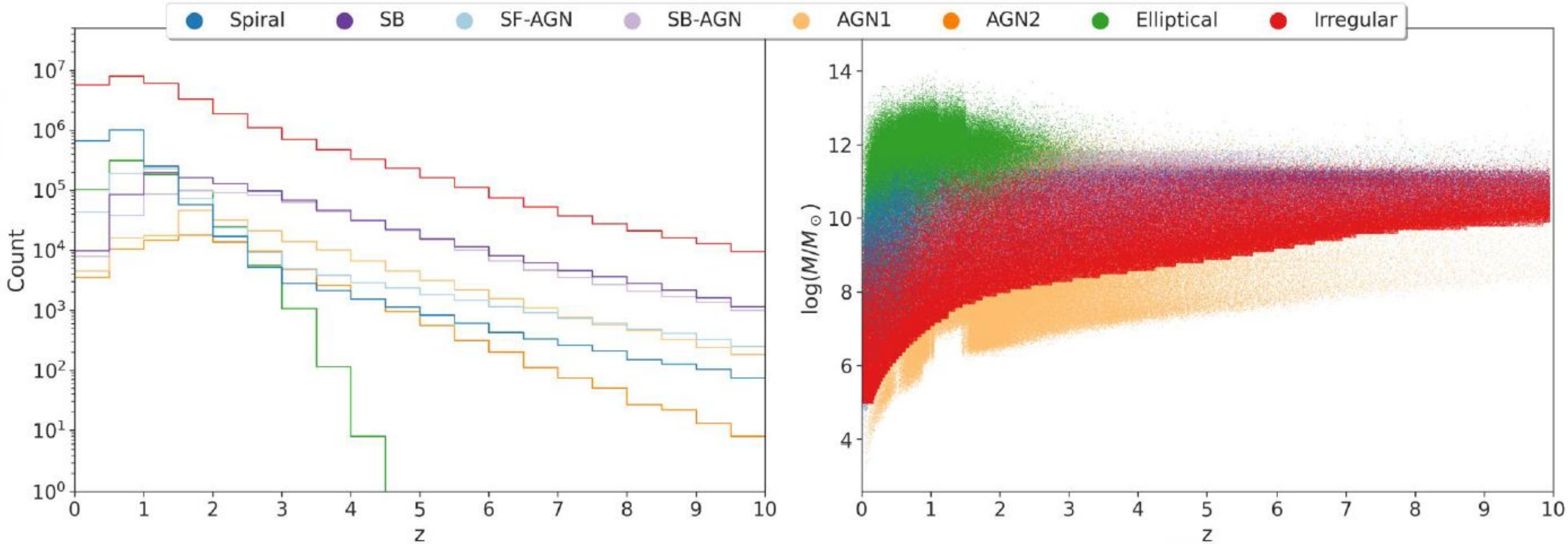
Gabarra+22 in prep

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

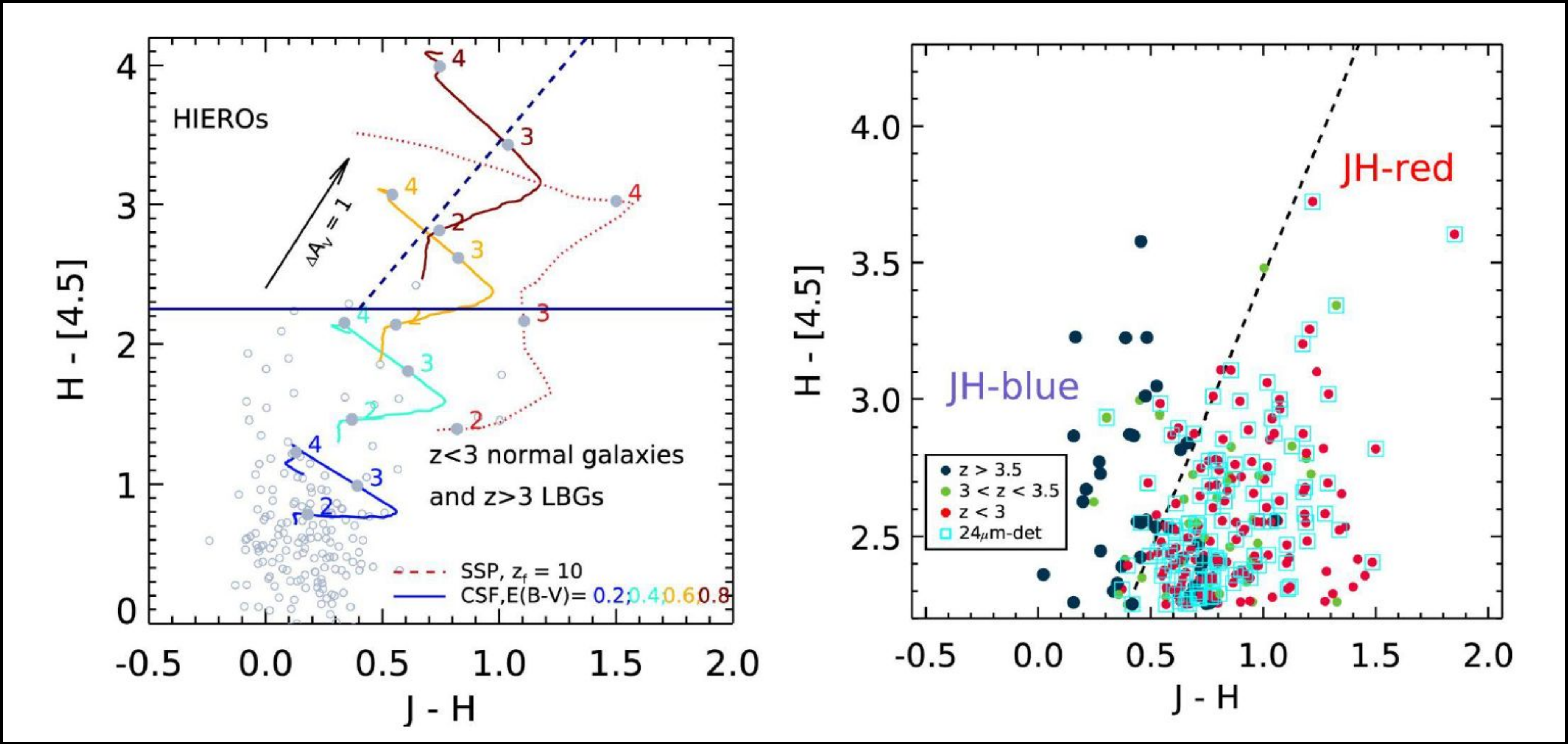
The Euclid Deep Survey combination of depth and area results in a simulated catalog with a total of more than 30 million objects with redshift from z ≈ 0 to ≈ 10



⇒ 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.**

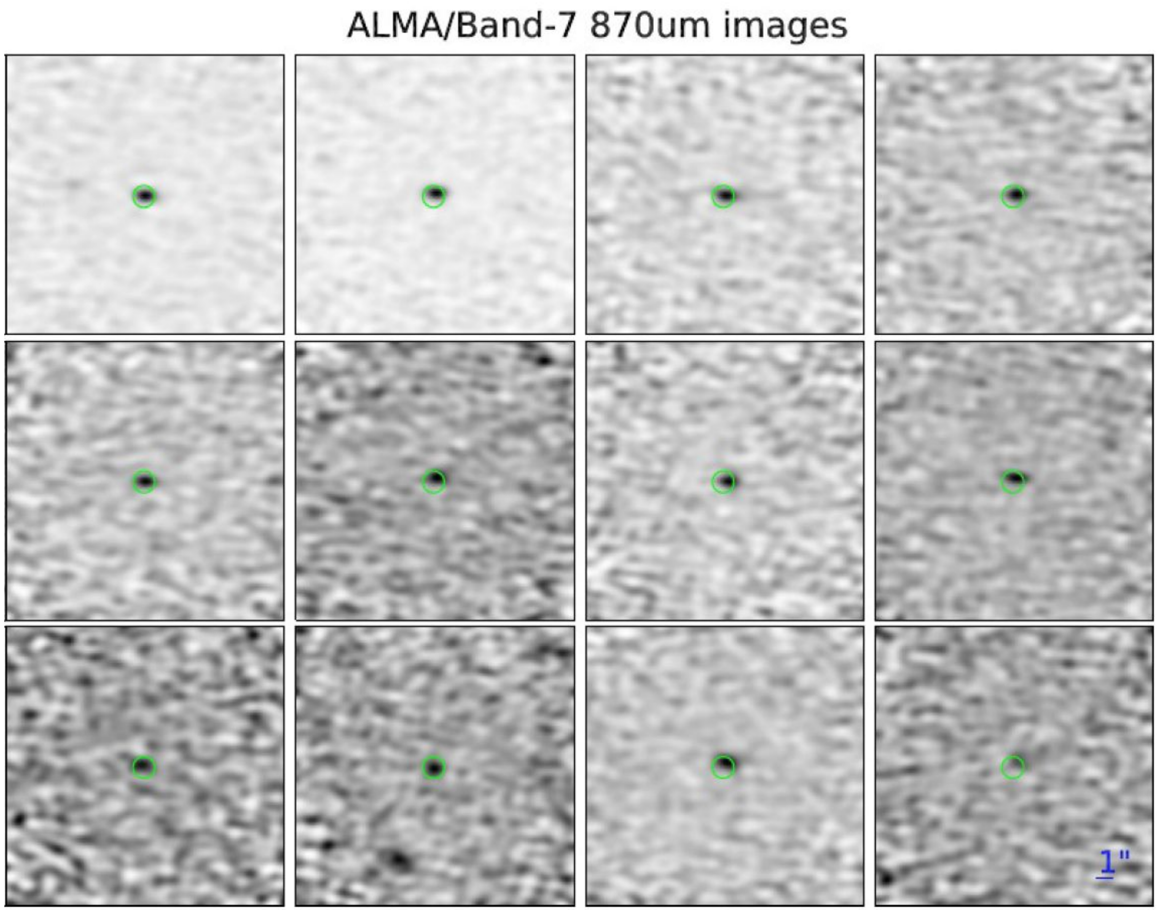
First assessment on HIEROS (Wang+16)



Most HST-dark dropout galaxies are detected in continuum by ALMA

- T. Wang: "H-dropouts" in the CANDELS catalogs
  → **62** galaxies (ch2 < 24)
- 17 of them were observed with ALMA (rest of the sample will come soon)
  → **80%** detection rate!

ALMA continuum detection favor the identification of a class of very extinguished dusty starforming sources



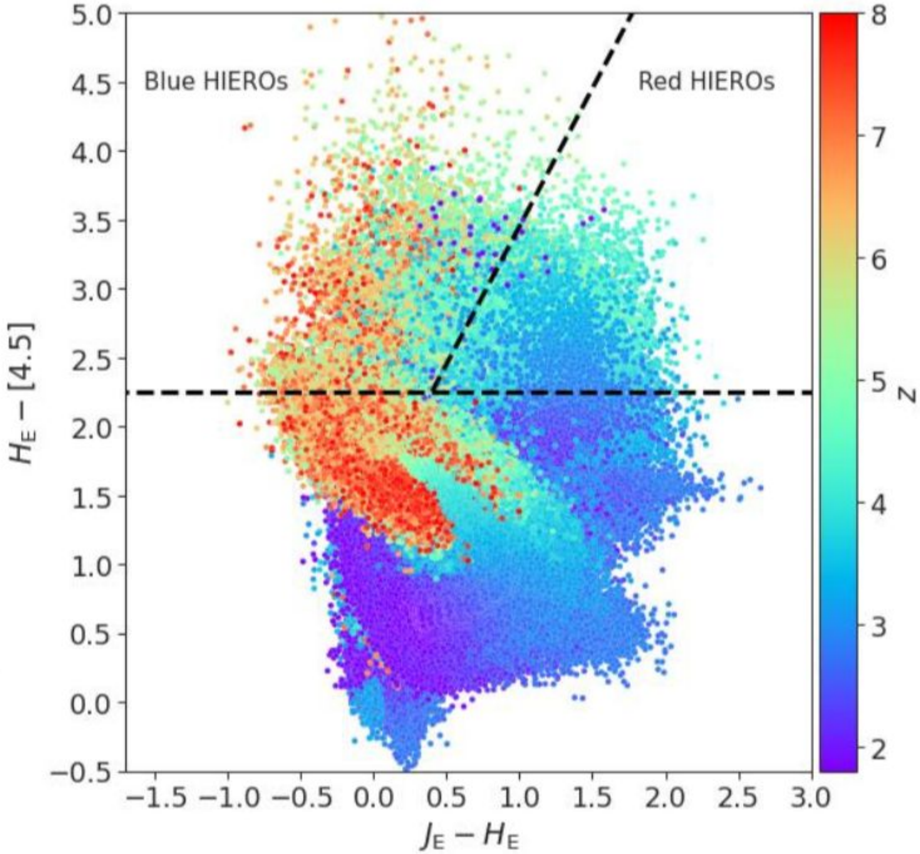
Wang+19 (Nature)

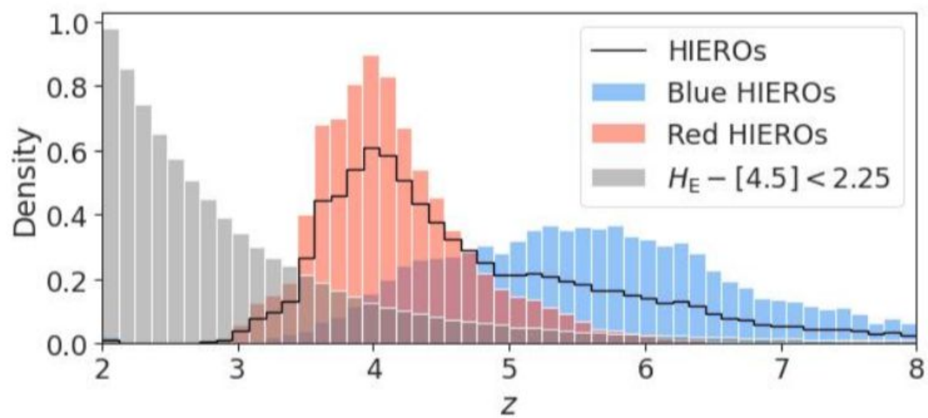
Unknown population, need to get redshifts!!!

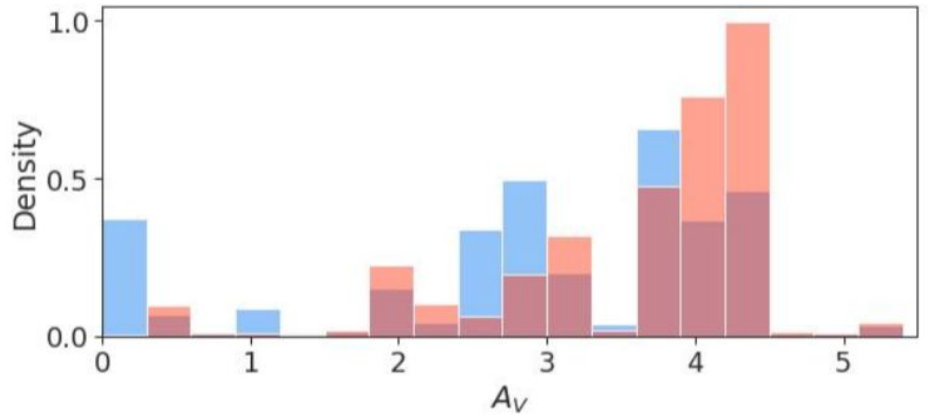
Photometric Selections

(Signor, GR, Bisigello et al. in prep.)

- 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 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

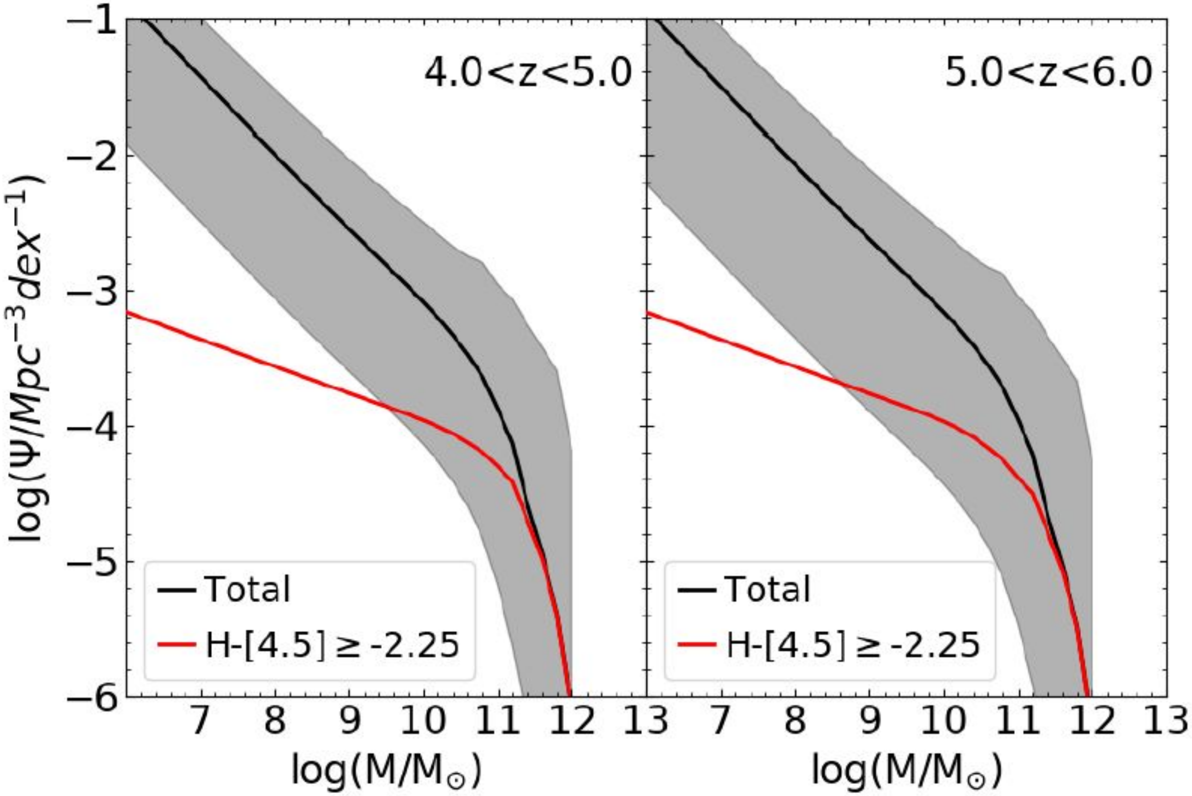






Photometric Selections

- 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: I_E, Y_E, J_E, H_E, *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
I_E	28.2	29.2
Y_E	26.3	27.3
J_E	26.5	27.5
H_E	26.4	27.4
Rubin/u	26.8	27.8
Rubin/g	28.4	29.4
Rubin/r	28.5	29.5
Rubin/i	28.3	29.3
Rubin/z	28.0	29.0
IRAC/3.6 µm	24.8	25.8
IRAC/4.5 µm	24.7	25.7

Photo-z: **Data Preprocessing**

The **input features** have been **preprocessed** before starting the subsequent analysis

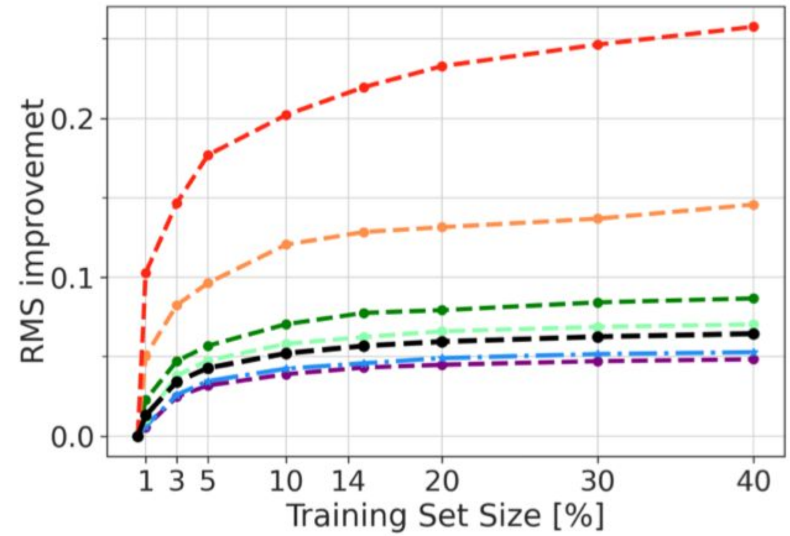
- First, the magnitude depth is set to 2σ.
- Only objects detected in at least 4 bands and with redshift from 2 to 8 are considered
- Some derived features are also included:

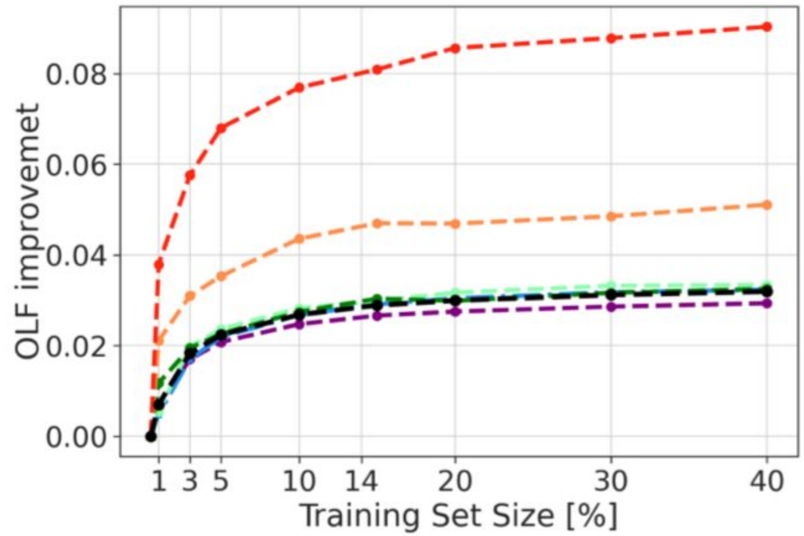
  1. **colors**: pairwise differences of the magnitudes;
  2. **ratios**: pairwise ratios between magnitudes;
  3. **errors**: parametric photometric errors associated to each band.

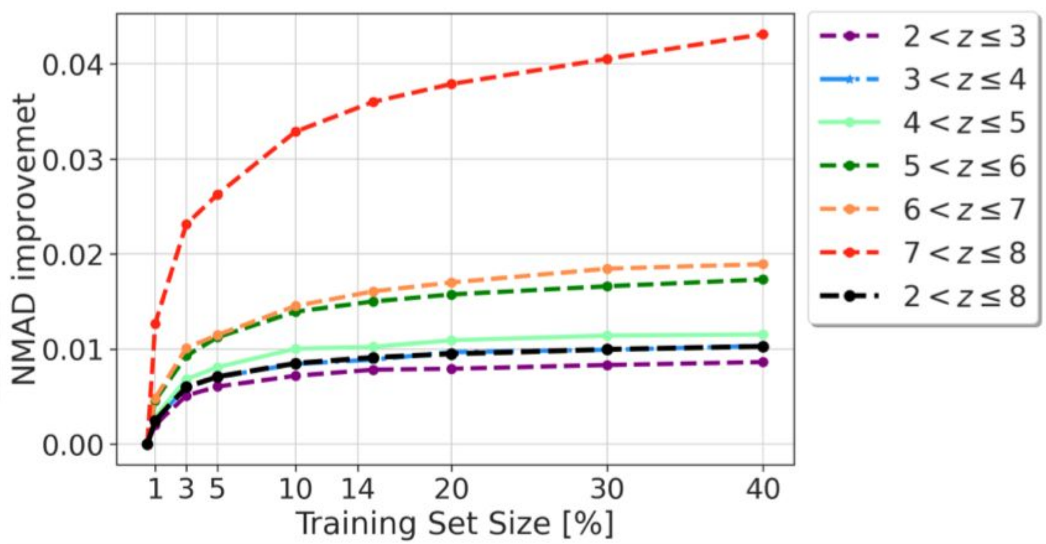
We are now ready to perform out analyses using XGBoost

Photo-z: **Training Set Size**

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**.







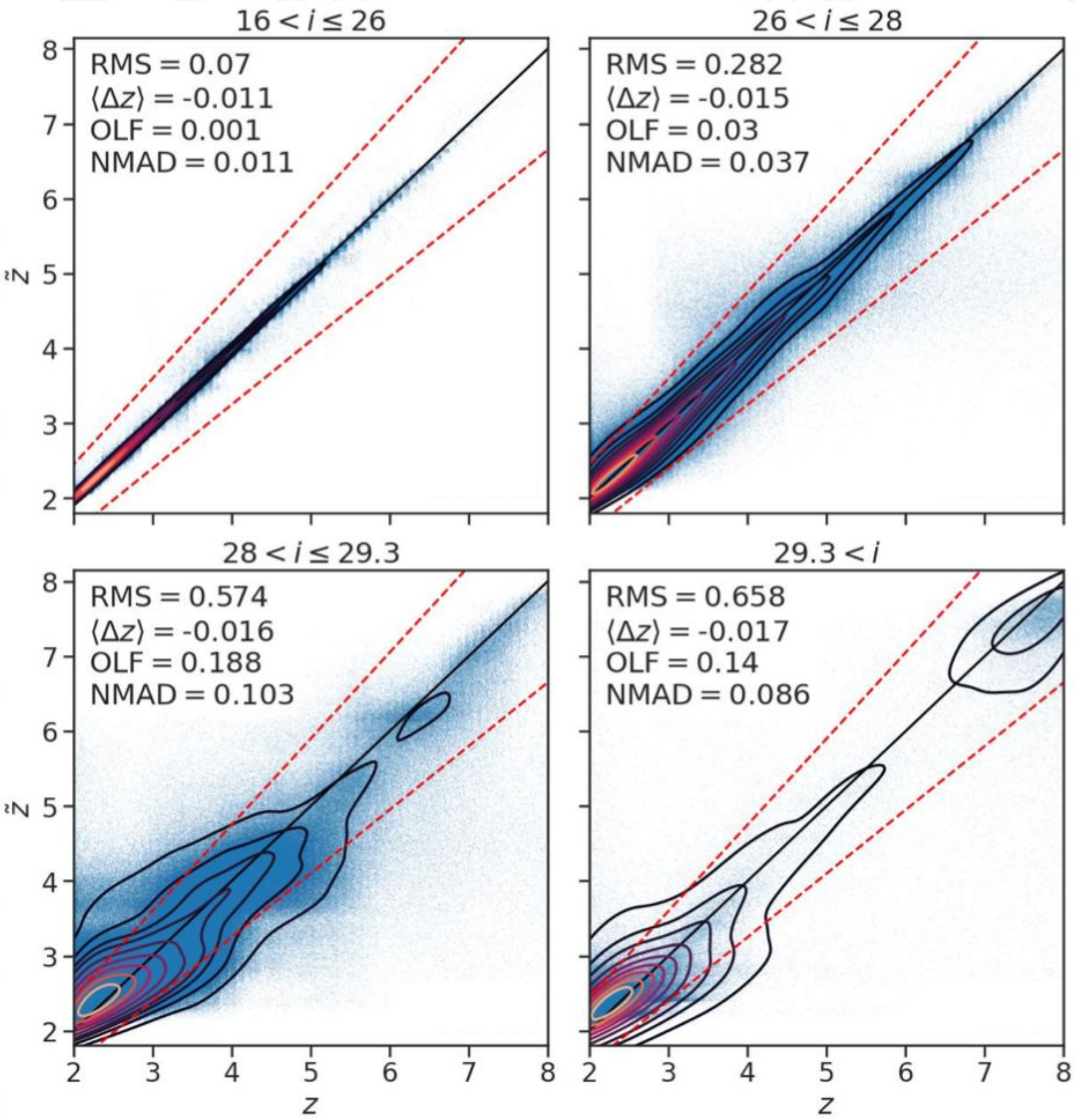
$$RMS = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(z_i - \tilde{z}_i)^2}$$

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

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

Photo-z: **Results**

Following a Bayesian optimization for the XGBoost hyperparameters, we evaluate its performance



- XGBoost takes ≈ 70 seconds to train on a data set with 517 498 samples and 132 features, and ≈ 40 seconds to predict the redshift for the 4+ million galaxies in the test set*
- Doing the same with LePHARE would take approximately 12 days.

*Timed on a workstation with a 2.20 GHz Intel Xeon CPU and a 16 GB Tesla P100-PCIE GPU

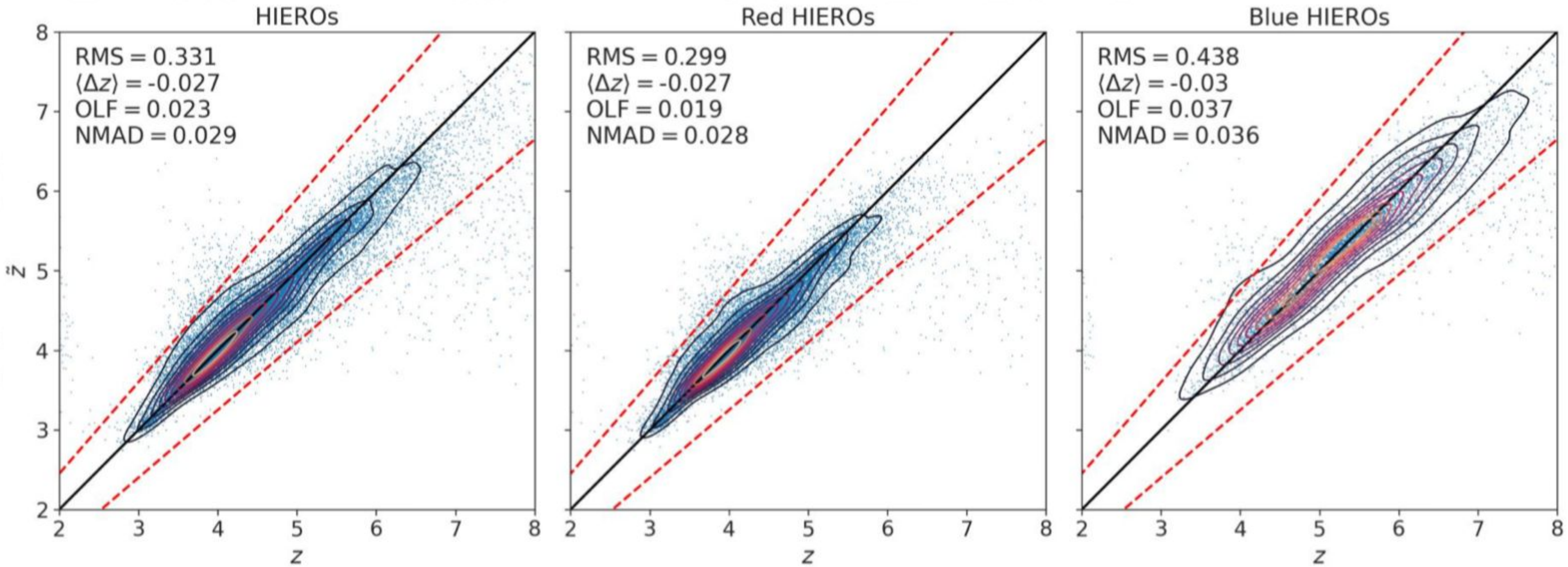
Photo-*z*: **Results**

- To give a contest to the results obtained, they are **compared** to previous photometric redshift performance, in a similar *z*-range.
- This comparsion is made with the results reported in **Weaver et al. 2022**, where the precision of photometric redshifts obtained via the SED fitting technique using 39 bands over the COSMOS2020 catalog is assessed againts spectroscopic ones
- Clearly the performance is very comparable

i-band magnitude	**NMAD**		**OLF**	
	COSMOS	*This work*	*COSMOS*	*This work*
$17 < i \leq 22.5$	0.008	0.011	0.006	0
$22.5 < i \leq 24$	0.015	0.007	0.032	0
$24 < i \leq 25$	0.024	0.008	0.063	0
$25 < i < 27$	0.044	0.019	0.204	0.005

Photo-z: Results

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

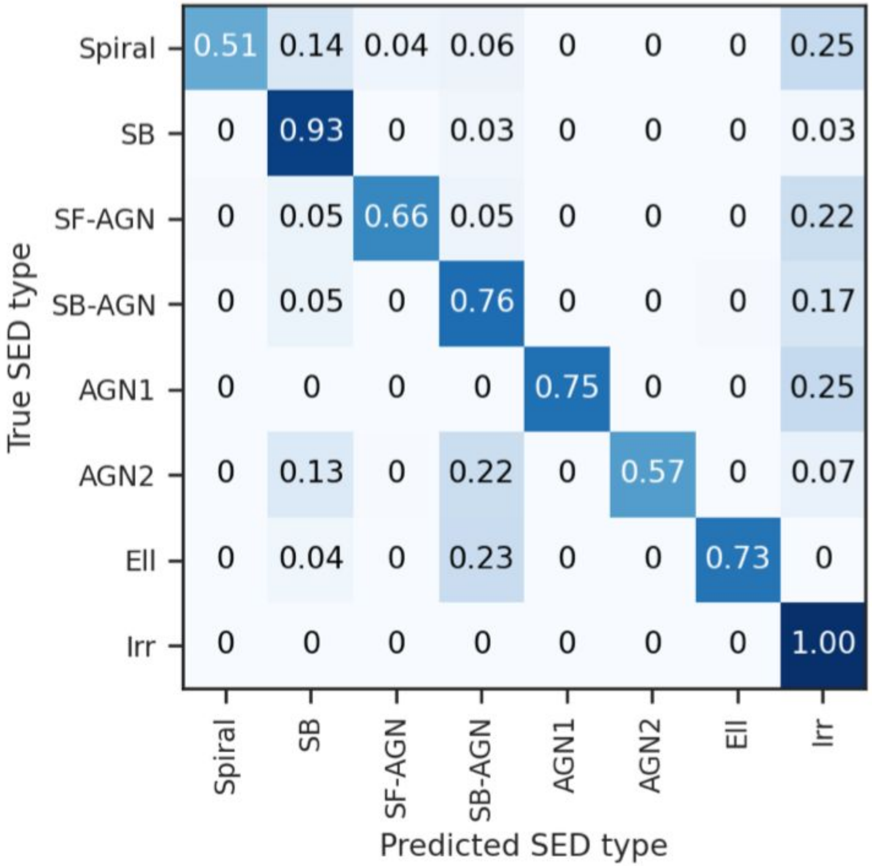


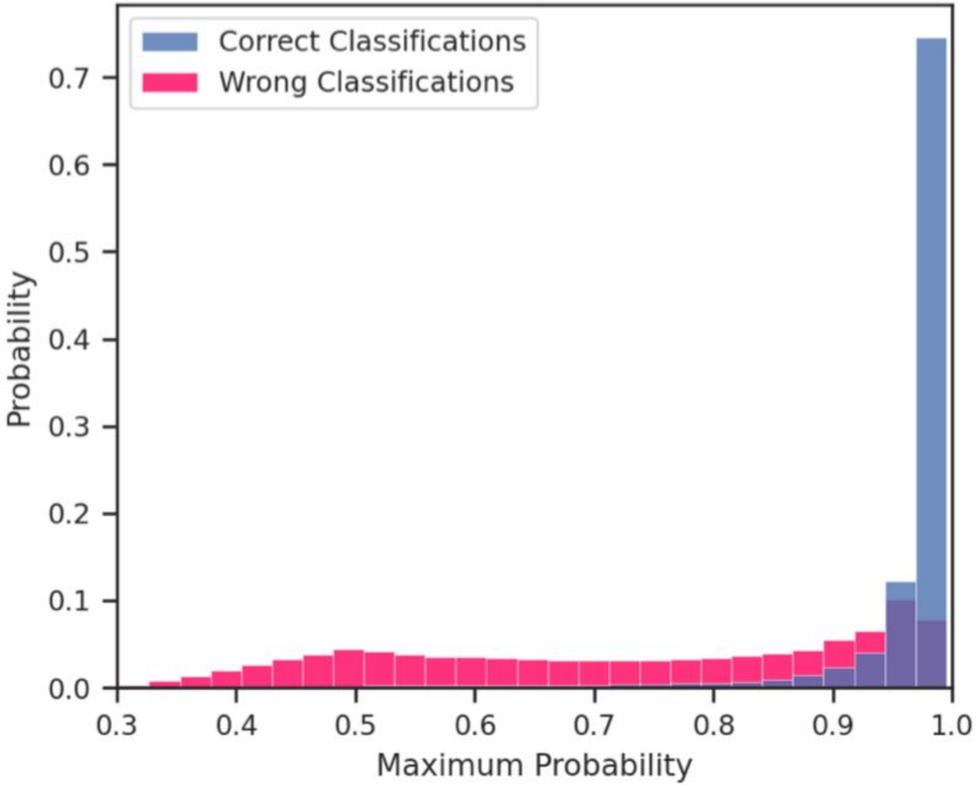
vs

Wang et al. 2016 HIEROs photometric redshifts
NMAD=0.11

Spectral Type Classification

- A gradient boosting approach was also taken to determine the SED type from photometry.
- Test Set accuracy: 96.3% - HIEROs accuracy: 86.45%
- This is clearly a simplified description, given the limited number of SED templates considered in the simulation





Conclusions & Future Developments

Summarizing:

- The study of distant obscured galaxies is fundamental to understand the build up of large scale structures, but, given their faintness and rarity, wide and deep surveys are required.
- We have provided a set of tests (colors diagnostics and photometric redshifts determination), based on simulations, suggesting that Euclid will allow to do the job, as we simultaneously identify and classify sources at $z>2$.
- In particular, the dusty population at $3 \leq z \leq 6$, which misses spectral features required to obtain a reliable spectroscopic redshift, is satisfyingly identified (for $H_E<26$).

Future Developments:

- Although XGBoost trained with additional features and cleaner data turned out to be already very powerful, some additional strategies can be taken, for example, the gradient boosting algorithm is extendable in order to provide confidence intervals for redshifts.
- It is also possible to perform a regression on the Stellar Mass
- Then, further work would include comparing the results obtained in this work with the official Euclid pipelines and on different simulated data-sets.

JWST unveils heavily obscured (active and passive) sources up to $z \sim 13$

Giulia Rodighiero[1,2]★, Laura Bisigello[1,2], Edoardo Iani[3], Antonino Marasco[2], Andrea Grazian[2], Francesco Sinigaglia[1,2], Paolo Cassata[1,2] and Carlotta Gruppioni[4]

[1]*Dipartimento di Fisica e Astronomia, Università di Padova, Vicolo dell'Osservatorio, 3, I-35122, Padova, Italy*
[2]*INAF–Osservatorio Astronomico di Padova, Vicolo dell'Osservatorio 5, I-35122, Padova, Italy*
[3]*Kapteyn Astronomical Institute, University of Groningen, P.O. Box 800, 9700AV Groningen, The Netherlands*
[4]*INAF–Osservatorio di Astrofisica e Scienza dello Spazio di Bologna, Via Gobetti 93/3 - 40129 Bologna - Italy*

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT

A wealth of extragalactic populations completely missed at UV-optical wavelengths has been identified in the last decade, combining the deepest *HST* and *Spitzer* observations. These dark sources are thought to be dusty and star-forming systems at $3 < z < 5$, and major contributors to the stellar mass build up. In this Letter we report an investigation of the deep *JWST* survey in the SMACS0723 cluster, analysing NIRCam and MIRI images. We search for sources in the F444W band that are undetected in the F200W catalogues. We characterise the properties of these sources via detailed SED modelling, accounting for a wide set of parameters and star formation histories, after a careful determination of their photometry. Among a robust sample of 20 candidates, we identify a mixed population of very red sources. We highlight the identification of evolved systems, with stellar masses $M_* \sim 10^{9-11} M_\odot$ at $8 < z < 13$ characterized by unexpectedly important dust content at those epochs (A_V up to $\sim$ 5.8mag), challenging current model predictions. We further identify an extremely red source (F200W-F440W$\sim$7mag) that can be reproduced only by the spectrum of a passive, quenched galaxy of $M_* \sim 10^{11.56} M_\odot$ at $z \sim 5$, filled of dust ($A_V \sim$ 5mag).

METHODOLOGY

Field: SMACS0723 NIRCAM + MIRI ⇒ magnification complicates but helps!

Selection: F444W detection / F200W non-detection in blind SExtractor matched catalogs (any a priori color cut)

Ad hoc photometry: refined photometry accounting for local background and contamination around each source ⇒ Marasco+22
Some no-detections might become very faint detections!

SED fitting: BAGPIPES (Carnall+18) with parametric SF histories (delayed declining + rising), wide range of parameter space: Av → 6mag

Position in the M*-SFR plane with redshift

How does this compare to recent JWST high-z candidates searches?

⇒ LBG selection! privileges UV blue and bright spectral types

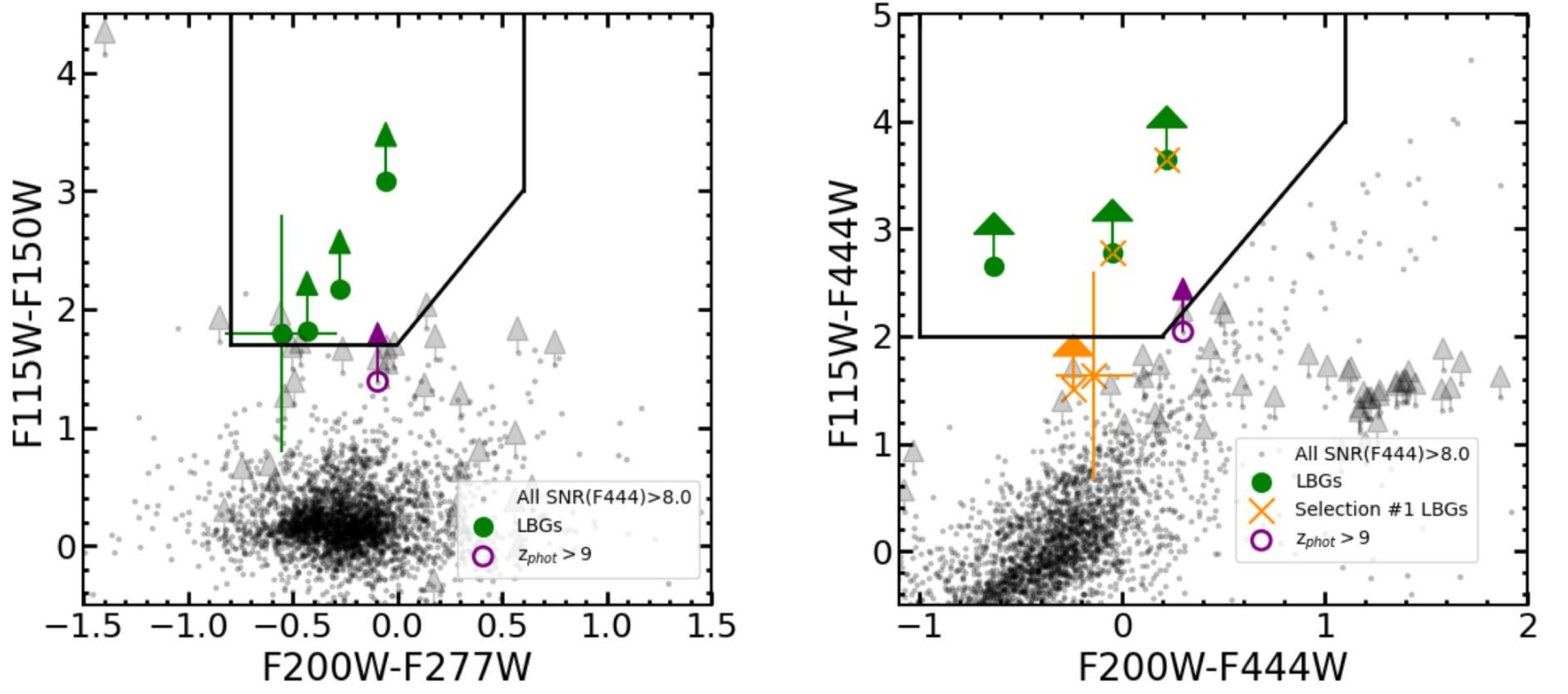
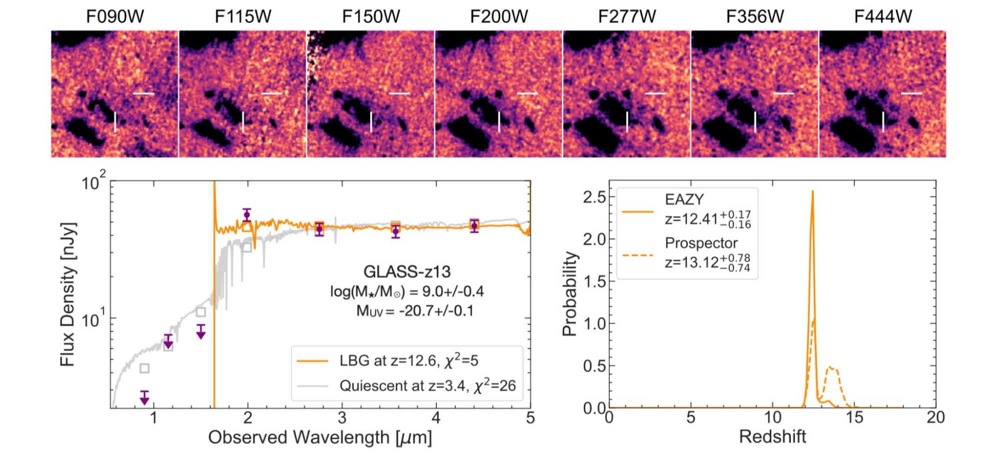


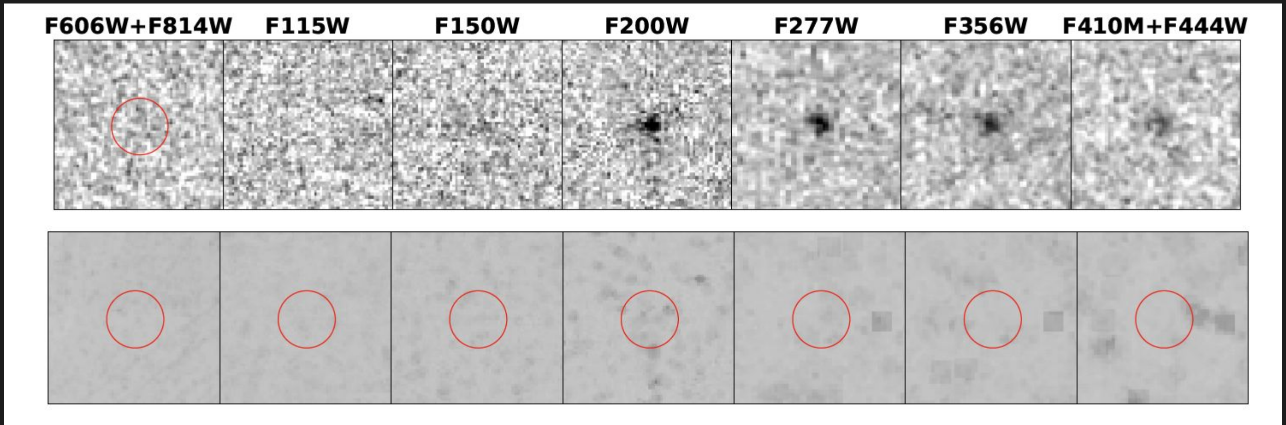
Figure 2. Observed color selection diagrams for LBGs at z~9-11 (left, Selection #1) and z~9-15 (right, Selection #2) in GLASS-JWST. Grey points show all objects with SNR(F444W)>8 in the GLASS catalog. Green circles indicate the color-selected candidates. The additional candidate selected on the basis of photometric redshift is shown as a purple empty circle. The z~9-11 LBGs from the Selection #1 diagram are shown as dark orange crosses in the Selection #2 one. Upper limits are indicated by arrows. All error-bars and upper limits are at 1σ.

Castellano+22

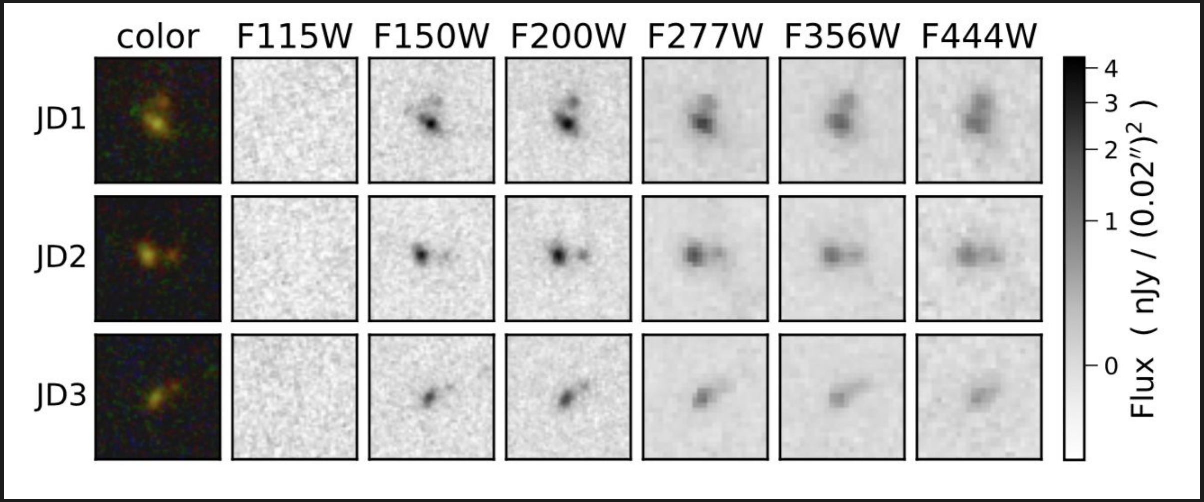
The famous ones



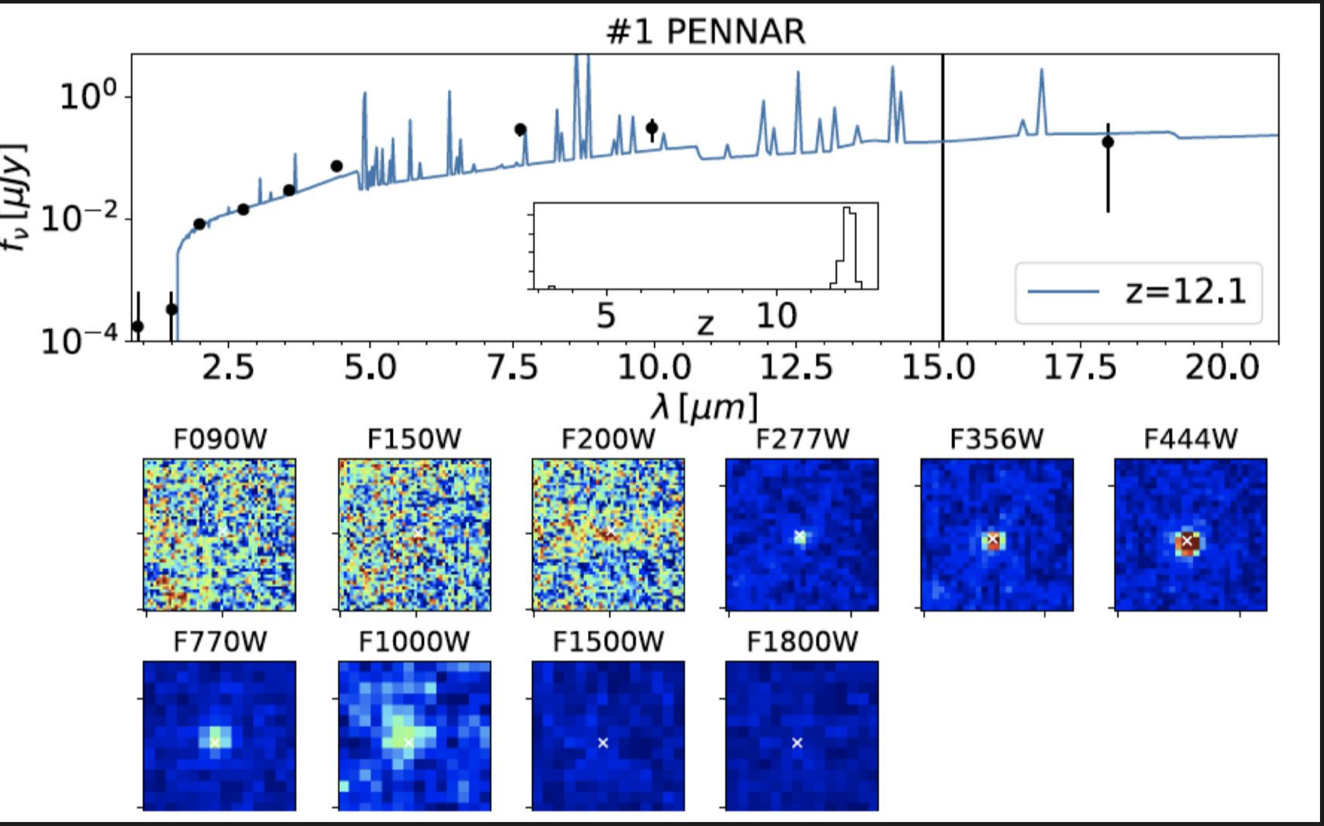
GHz11 @z~11
GHz13 @z~13
Naidu+22



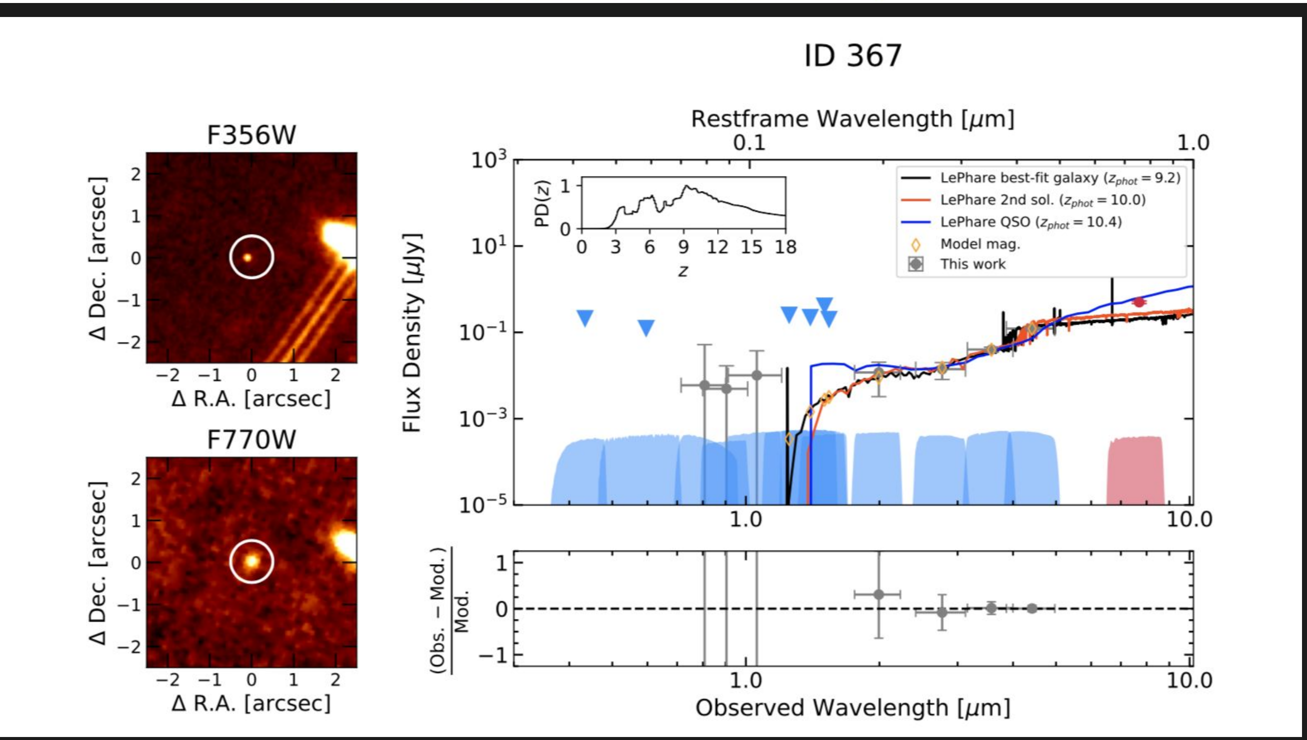
Maisie @z~12
Finkelstein+22



Triply lensed merger @z~11
Hsiao+22 (today on arXiv)



PENNAR at z~12: the highest redshift dusty galaxy?



PENNAR as in Iani+22: ID 367

the highest redshift candidate in the MIRI 7.7um SMACS0723 catalog

The hidden side of cosmic star formation at $z > 3$

Bridging optically-dark and Lyman break galaxies with GOODS-ALMA

M.-Y. Xiao[1,2,3★], D. Elbaz[2], C. Gómez-Guijarro[2], L. Leroy[2], L.-J. Bing[4], E. Daddi[2], B. Magnelli[2], M. Franco[5], L. Zhou[1,3], M. Dickinson[6], T. Wang[1,3], W. Rujopakarn[7,8,9], G. E. Magdis[10,11,12], E. Treister[13], H. Inami[14], R. Demarco[15], M. T. Sargent[16,17], X. Shu[18], J. S. Kartaltepe[19], D. M. Alexander[20], M. Béthermin[4], F. Bournaud[2], R. Chary[21], L. Ciesla[4], H. C. Ferguson[22], S. L. Finkelstein[5], M. Giavalisco[23], Q.-S. Gu[1,3], D. Iono[24,25], S. Juneau[6], G. Lagache[4], R. Leiton[15], H. Messias[26,27], K. Motohara[28], J. Mullaney[29], N. Nagar[15,30], M. Pannella[31,32], C. Papovich[33,34], A. Pope[26], C. Schreiber[35], and J. Silverman[9]

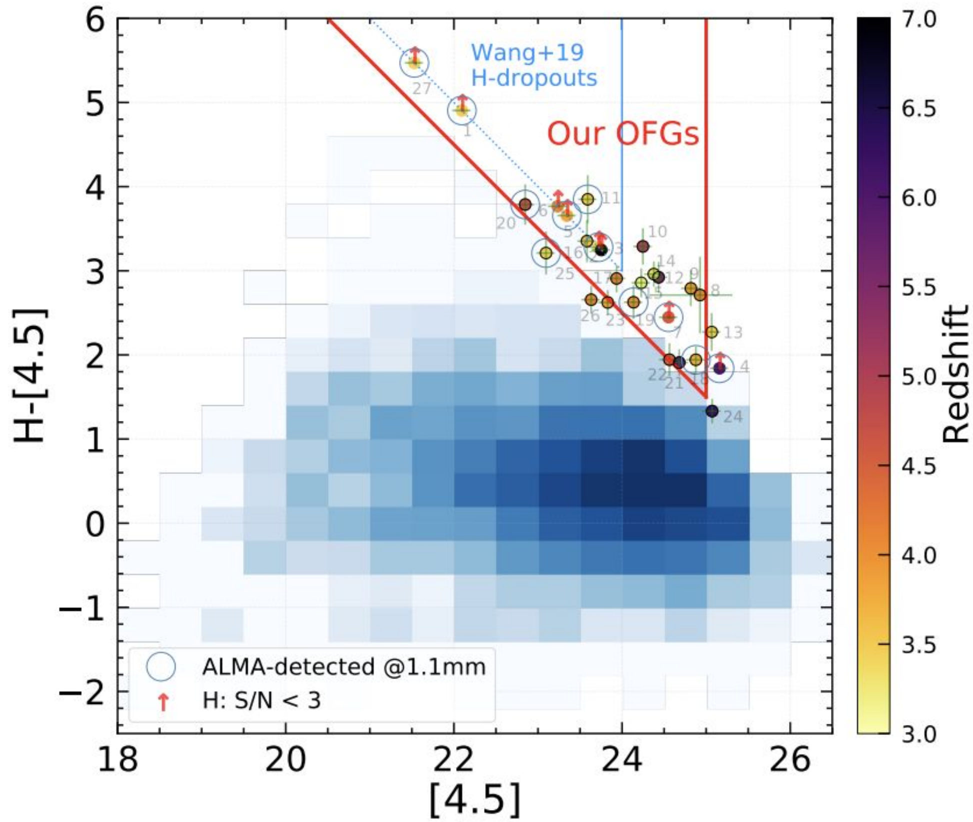


Fig. 3. Color-magnitude diagram color-coded by photometric redshift. Our criteria (in red) for selecting OFGs are: $H > 26.5$ mag & $[4.5] < 25$ mag. We note that we include sources outside the wedge whose 1σ photometric uncertainties overlap the wedge, so we have points outside the red triangle. The arrows show our H-dropouts (S/N < 3), with the typical depth of H = 27 mag (5σ) in the shallowest region of the HLF survey as their lower limits. The blue shaded area describes the distribution of all the IRAC detected sources (see Appendix A for more details) in the GOODS-ALMA field. The blue and red triangular regions are the same as in Fig. 2.

HST-dark Xiao+22:
[H]>26.5 mag &
[4.5] < 25 mag

Stacked SED

visible at short wavelengths

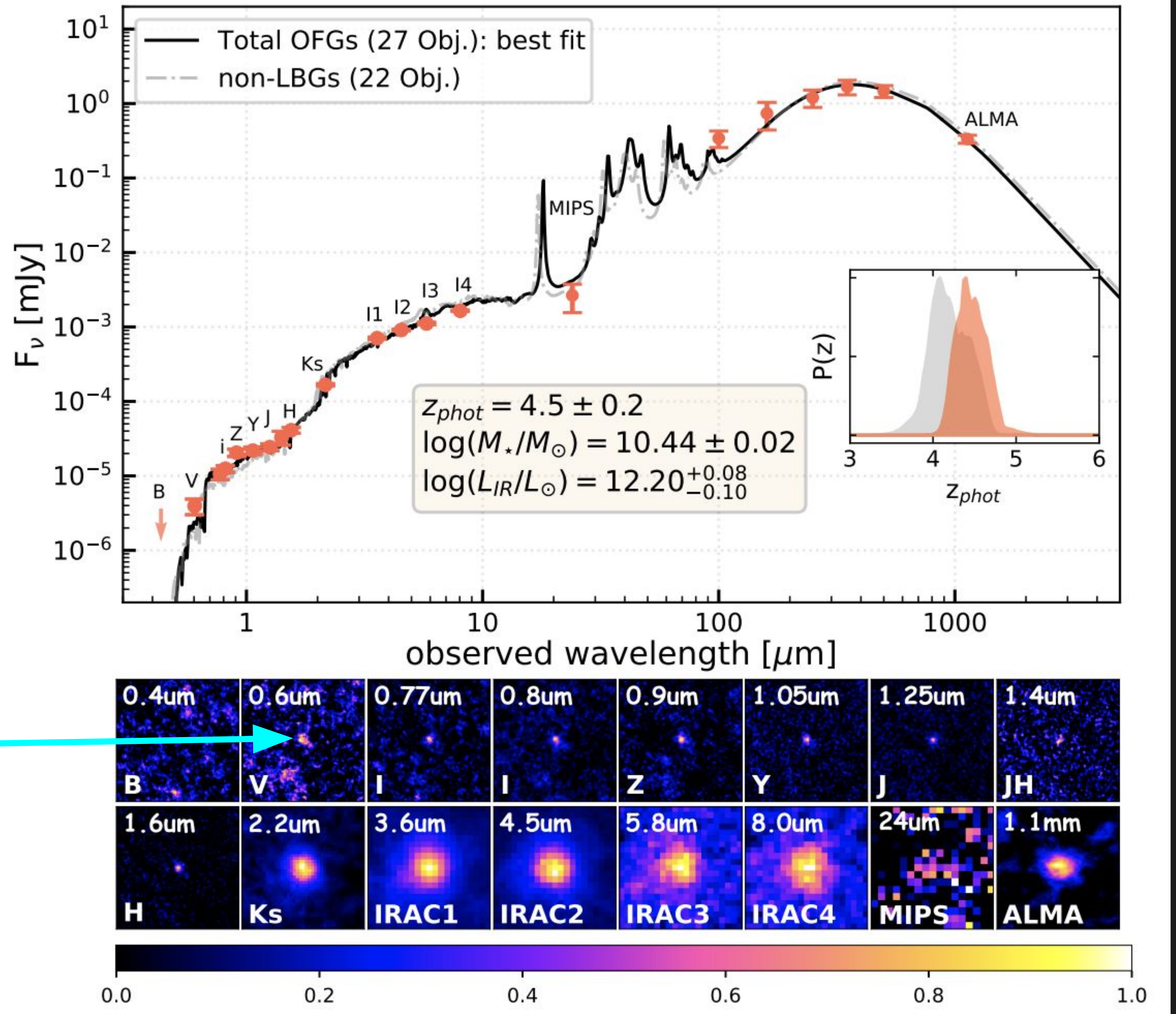
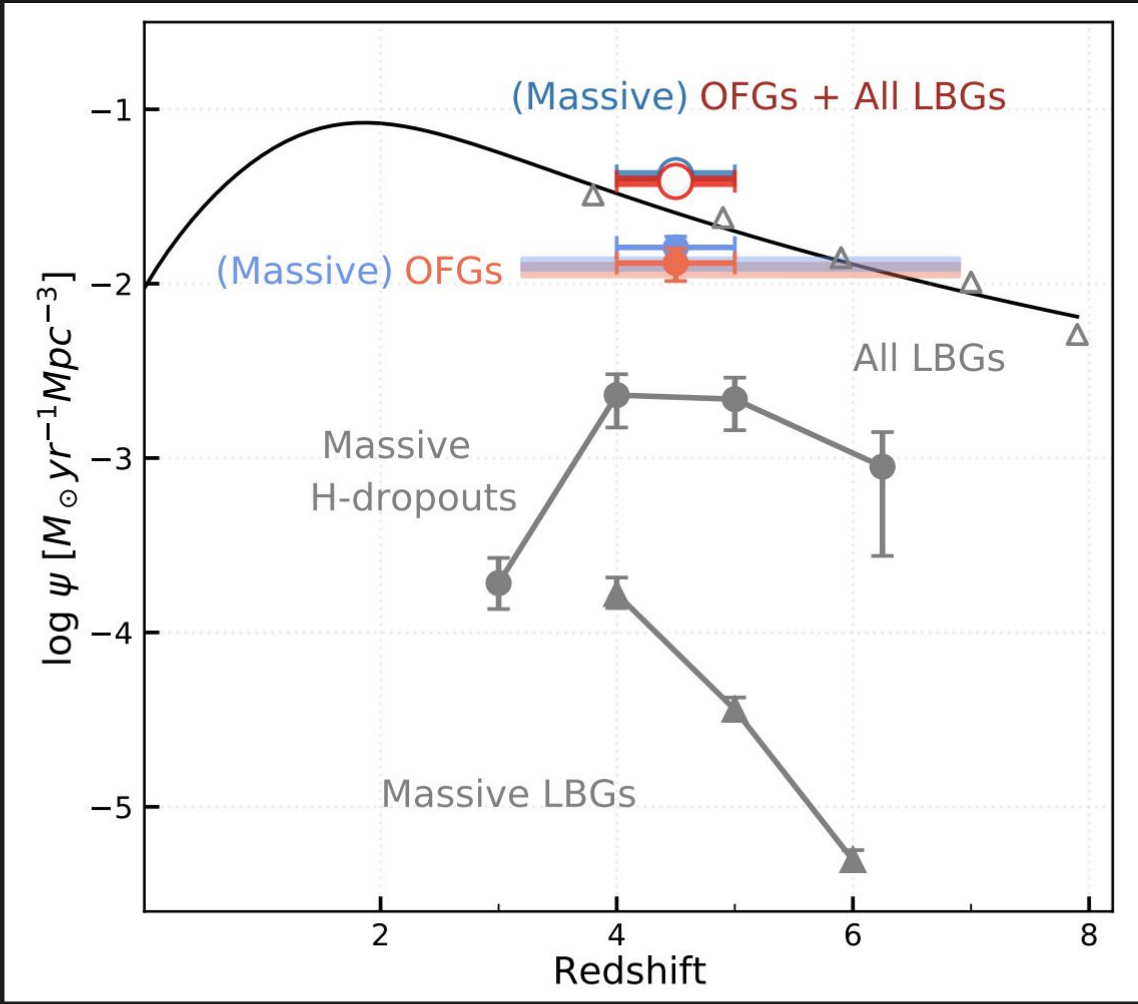


Fig. 6. Median stacked SED and images of the total sample of 27 OFGs in this work. *Top*: best-fit SED of the total sample (black line). The measured fluxes (red points) are derived from the stacked images. Error bars (1σ) and upper limits (3σ) are obtained from the Monte Carlo simulation (except *Herschel*) and bootstrap approach (*Herschel*; see §5.2). We also show the best-fit SED for 22 non-LBGs (grey line). These 22 non-LBGs will be used to calculate the cosmic SFRD. The inset shows the likelihood distributions of the photometric redshift of our samples (total sample in red, 22 non-LBGs in grey), based on the UV to MIR SED fitting from EAzY, which is normalized to the peak value. The redshift obtained from the maximized likelihood is $z \sim 4.5$ for the total 27 OFGs and $z \sim 4.2$ for the 22 non-LBGs. *Bottom*: stacked images of the total sample with peak fluxes normalized. Each panel is $6'' \times 6''$ except for the MIPS 24 μm, which is $24'' \times 24''$.

SFR density contribution

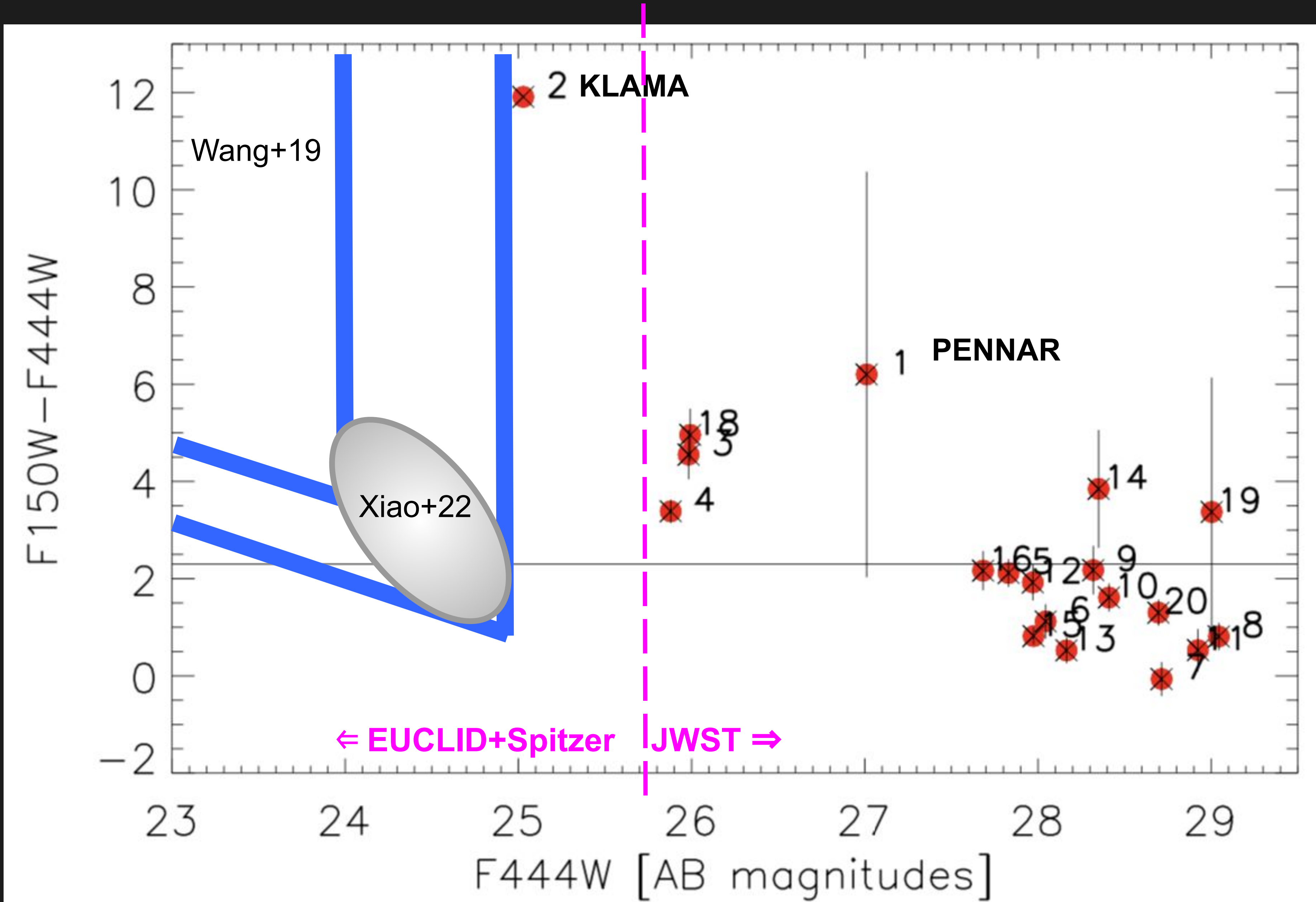


Xiao+22

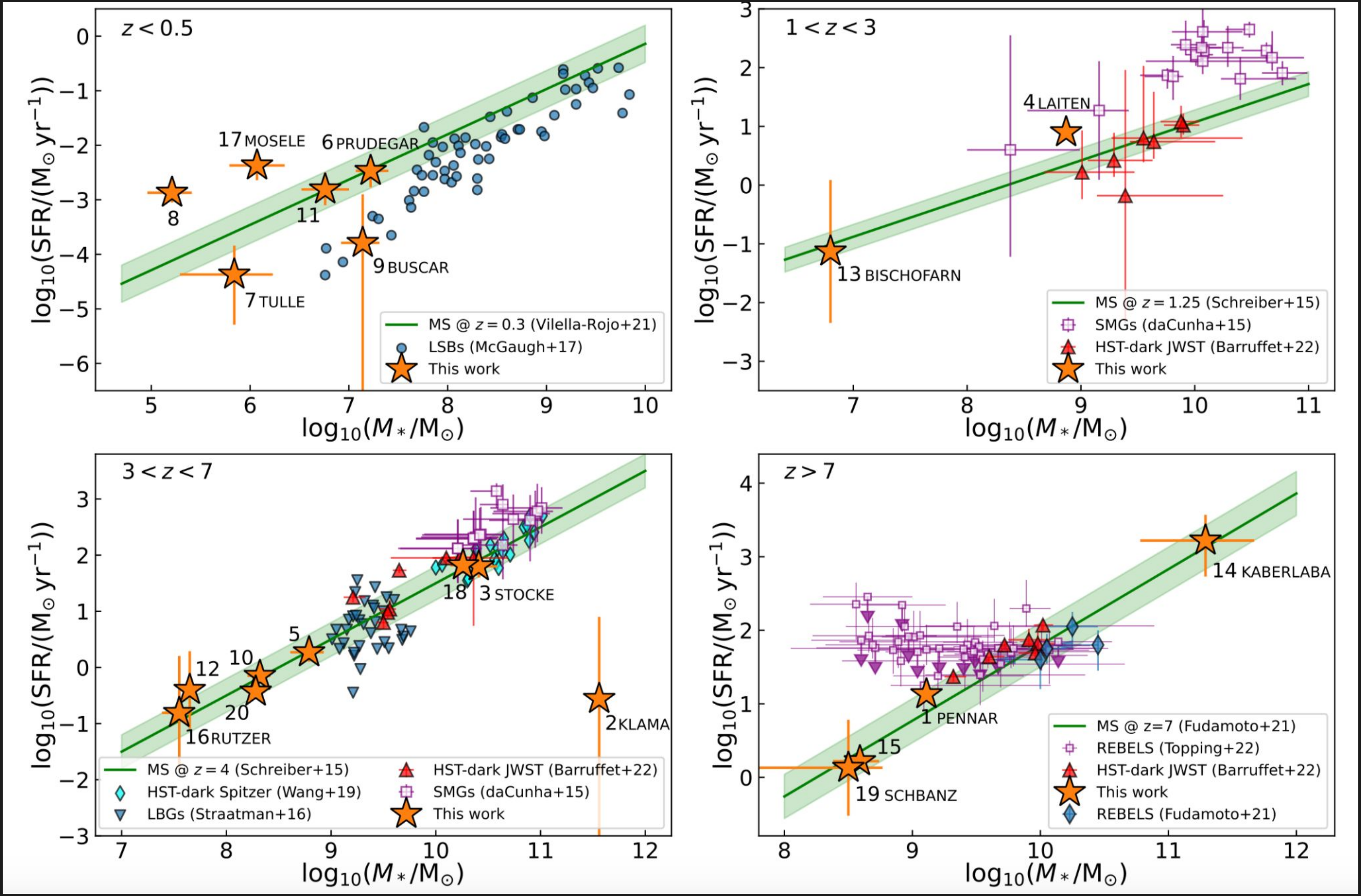
1) **The SFRD at z > 3 contributed by massive OFGs (log(M*/Msun) > 10.3) is at least two orders of magnitude higher than the one contributed by equivalently massive LBGs.**

1) **The combined contribution of OFGs and LBGs to the cosmic SFRD at z = 4 − 5 is about 0.15 dex (43%) higher than the SFRD derived from UV-selected samples alone (Madau & Dickinson 2014) at the same redshift.**

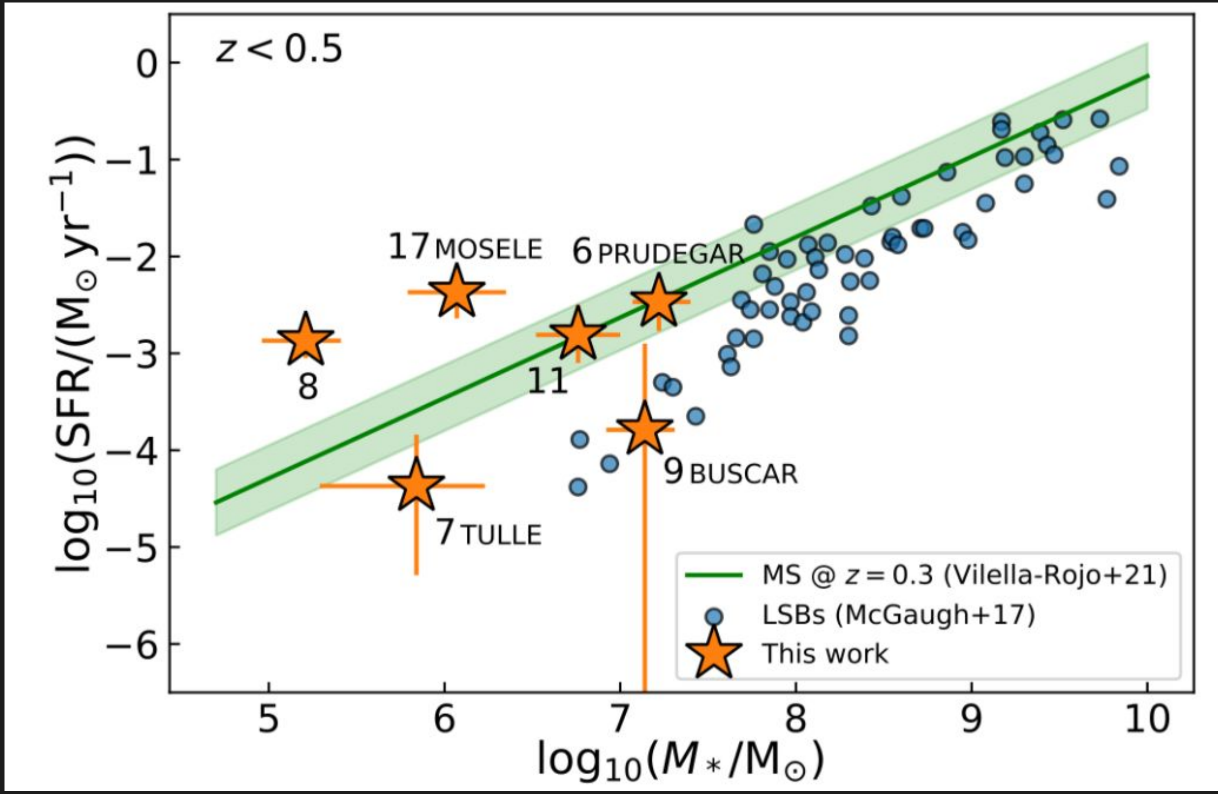
COLOR-mag diagram: comparison to Wang+19 ~HIERO selection



Comparison to the MS at different redshifts



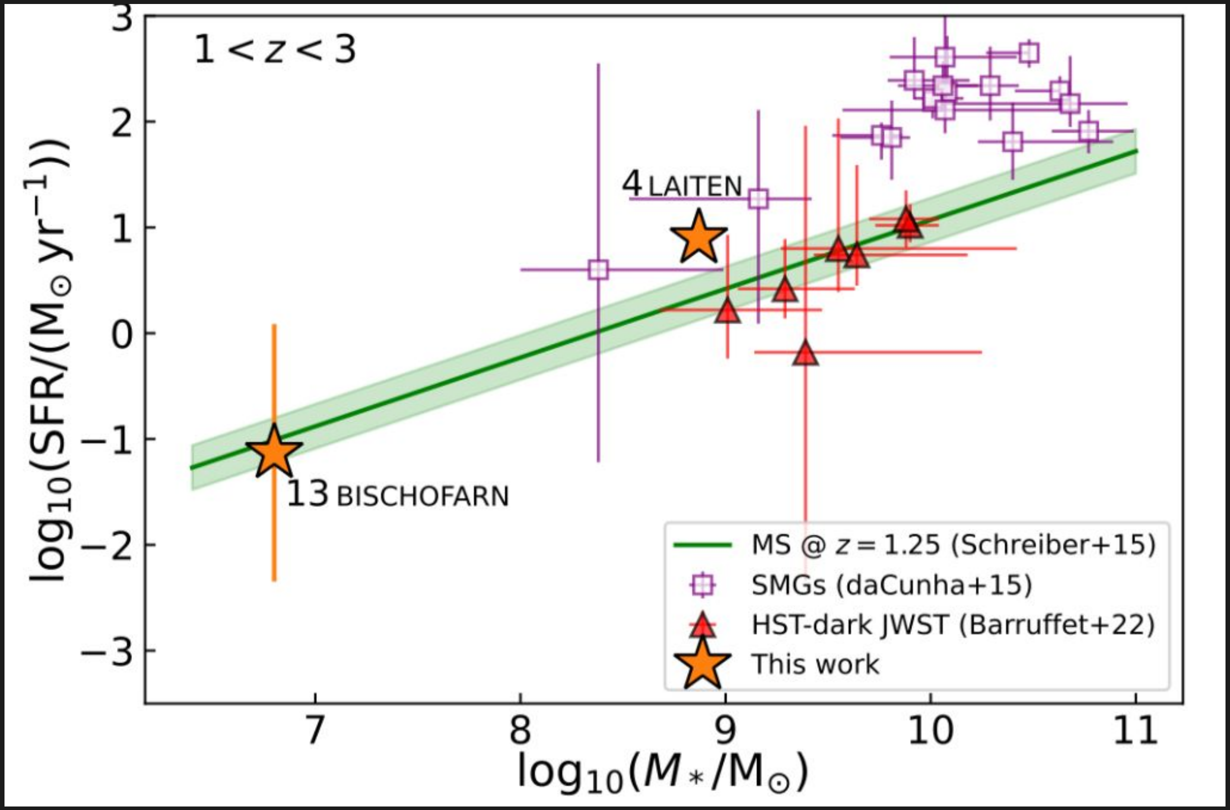
z<0.5



Red and dusty low-z dwarf galaxies:

JWST dwarves are much more extinguished than traditional UV selection, with AV up to ~5.5 mag

Cosmic Noon



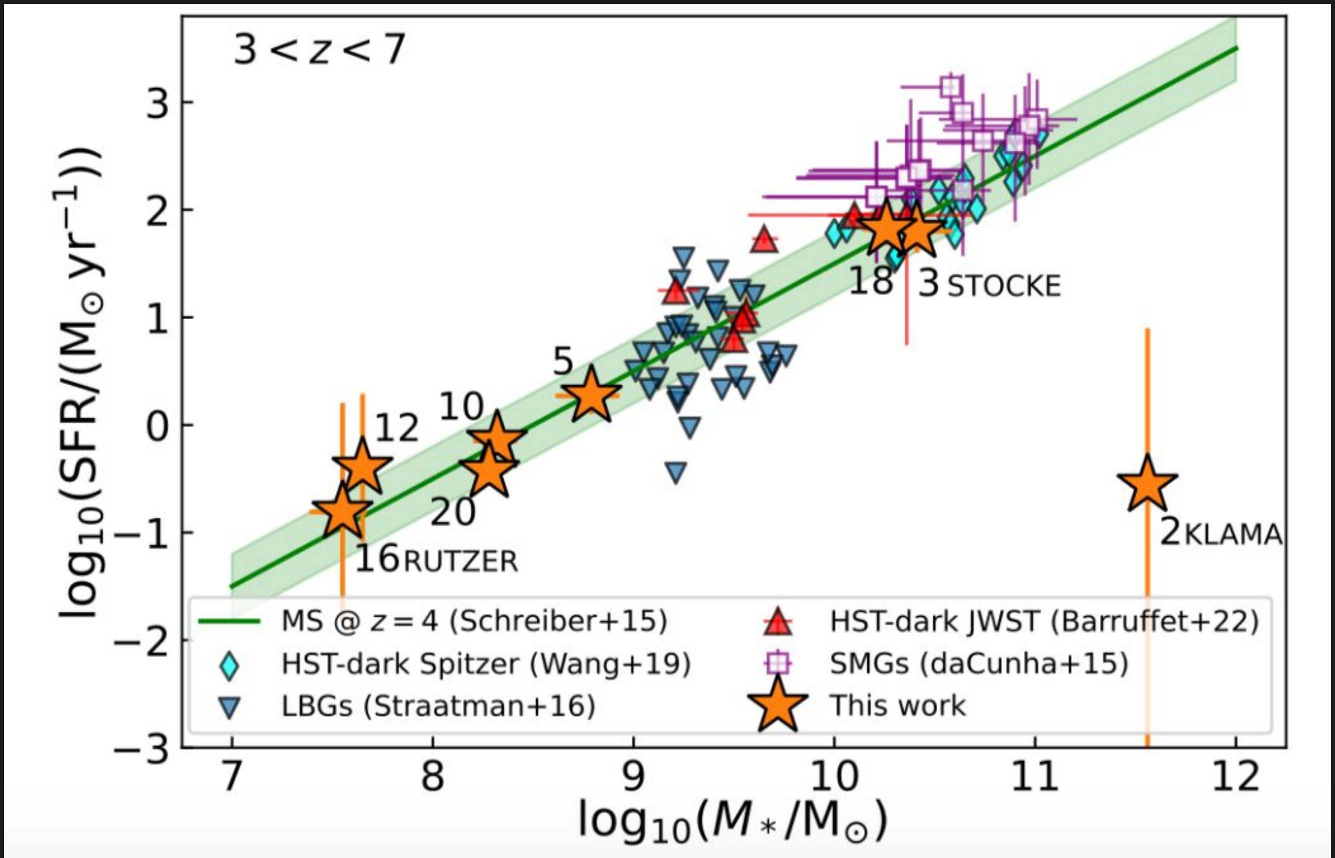
Not much statistics

faint end

Av~2-4

3<z<7

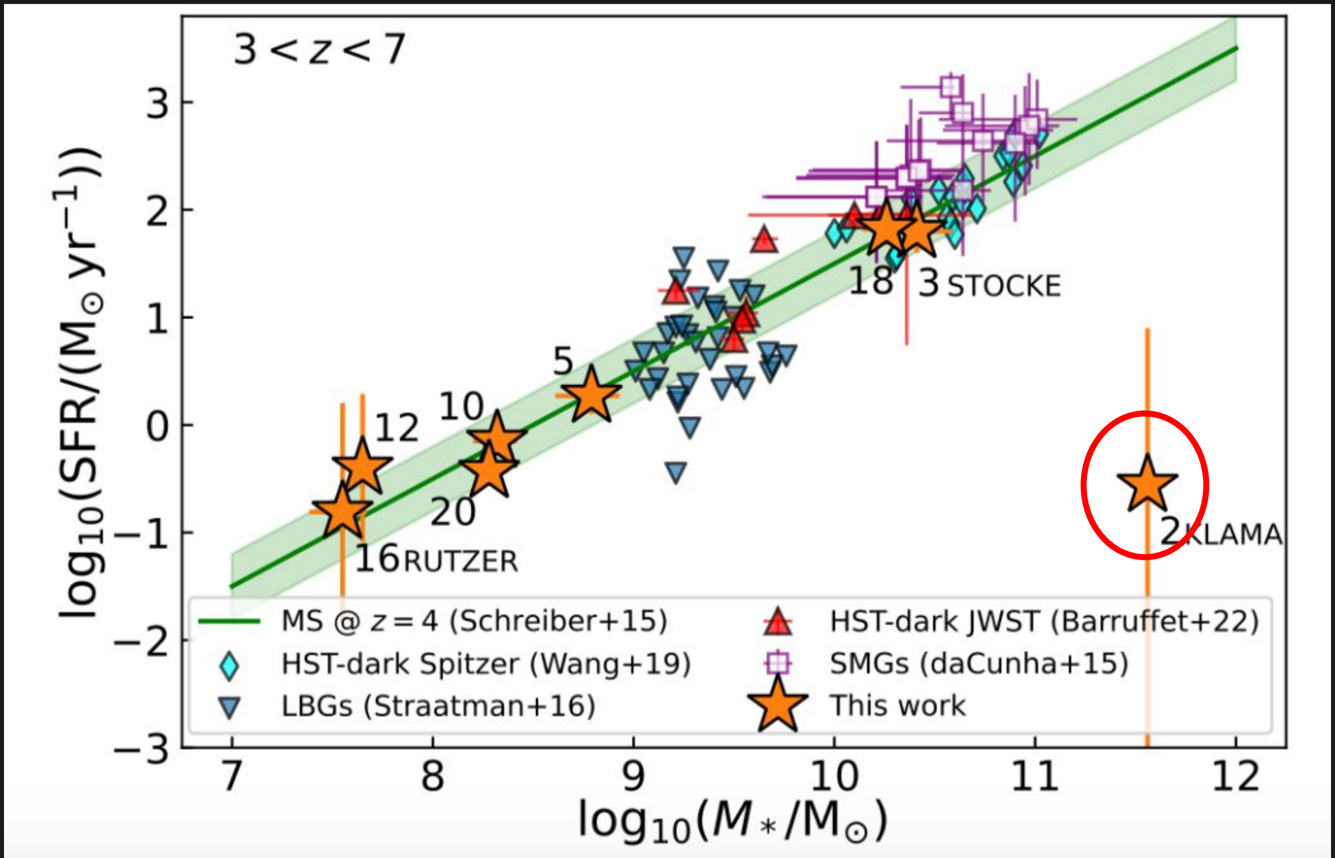
The HST-dark territory:



⇒ dusty star forming sources consistent with HIERO properties (highly extinguished)

⇒JWST probes a much lower stellar mass range!

3<z<7

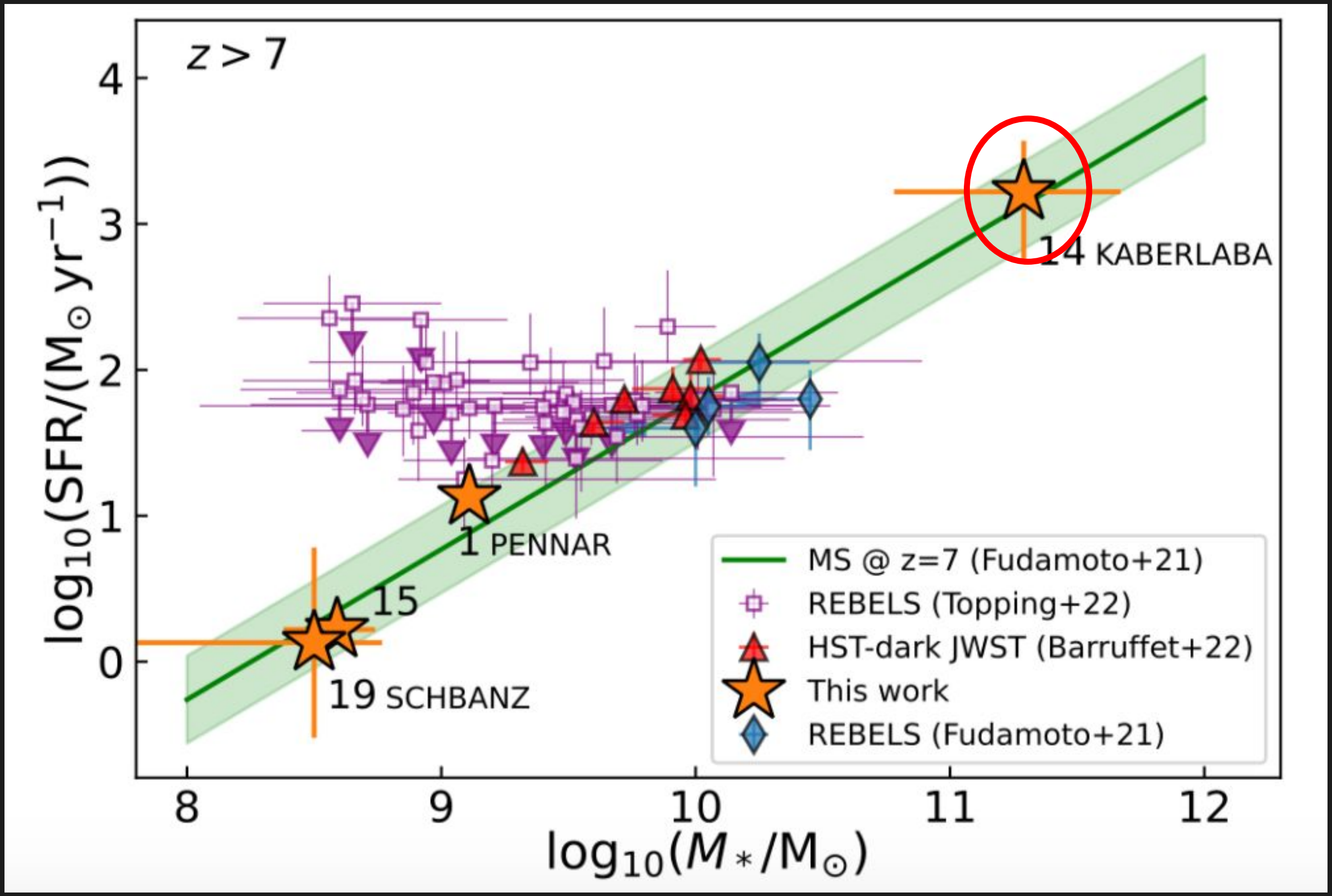


The HST-dark territory:

A quenched, dusty and massive galaxy at z~5? Av~4.7mag

⇒A quiescent galaxy whose dust content has yet to be destroyed, a possible indicator of recent quenching??

and finally, high-z! (really????)



tension with LambdaCDM? (Lovell+22)

Extinguished high-z star-forming sources:

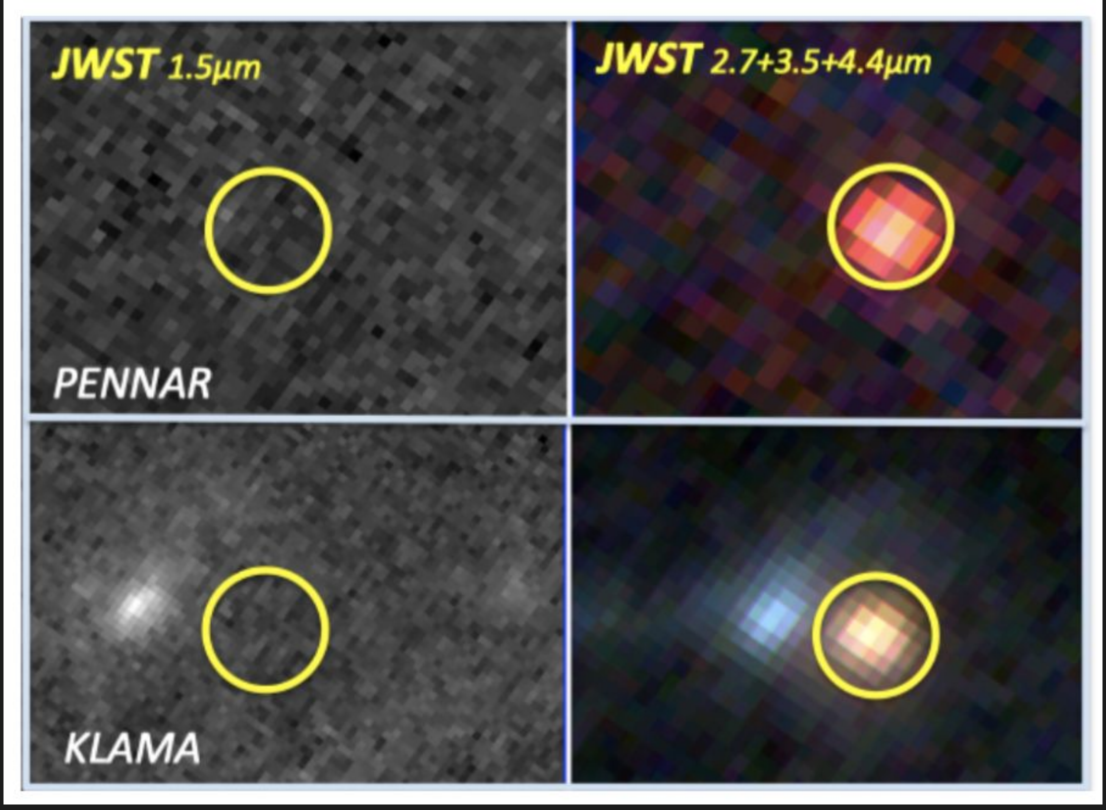
We classify four objects at $z > 8$, with mature stellar populations, log(M∗/M☉) ~9-11, that differ from already detected JWST sources at similar cosmic epochs for their extreme dust content (Av = 0.4−5.8mag)

⇒

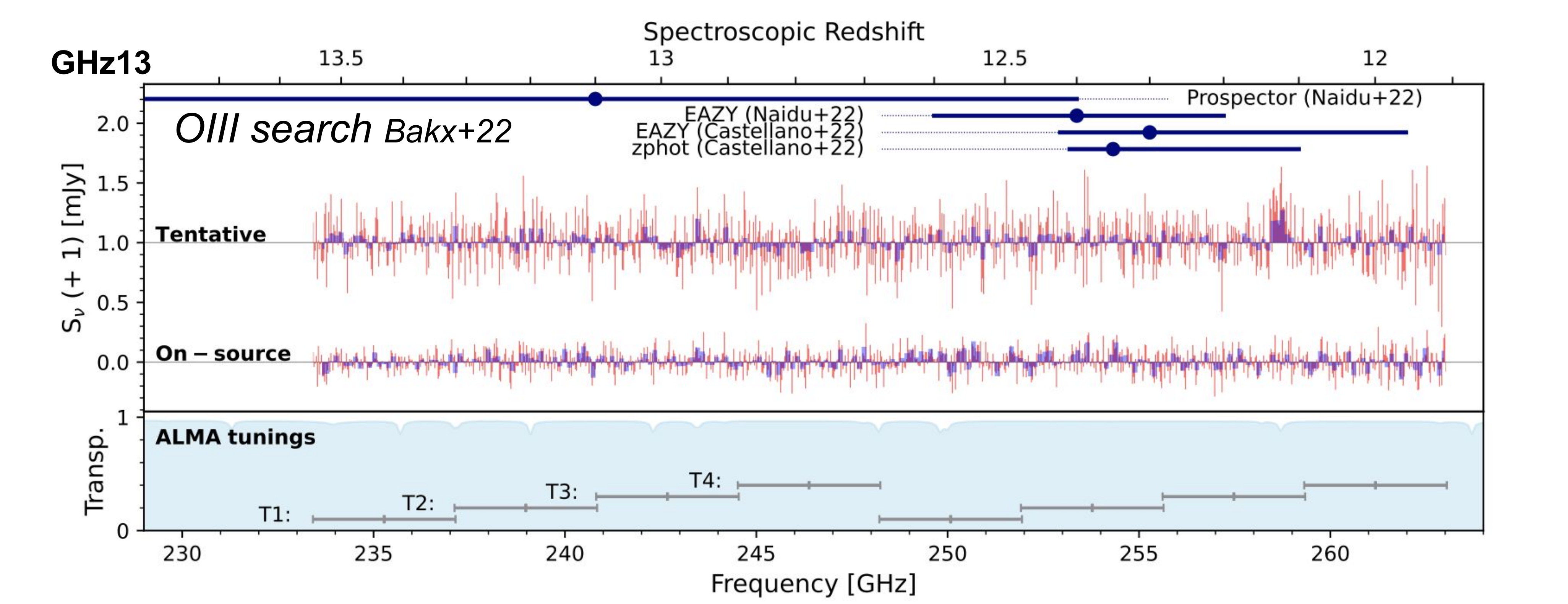
unexplained by current theoretical model

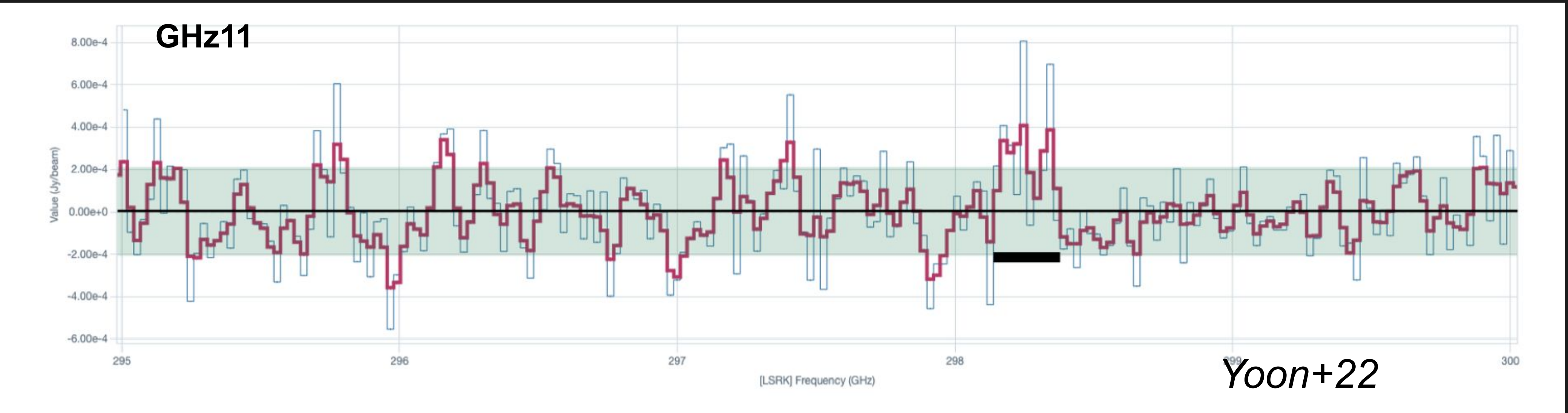
CONCLUSIONS

- All the preliminary JWST photometric candidates require an urgent confirmation
- However, it is clear that the dark and extremely dark sources detected by Webb should include at least a few very high-z objects
- LBG only technique looses the dustier side of dropouts
- Our results suggests that JWST very red sources represent a dust rich population at different redshifts, previously missed even by HST and Spitzer
- New parameters spaces are being filled (low mass, high-z, obscuration)



What's next? Need for spectroscopic information .. ALMA → tentative detections up to now



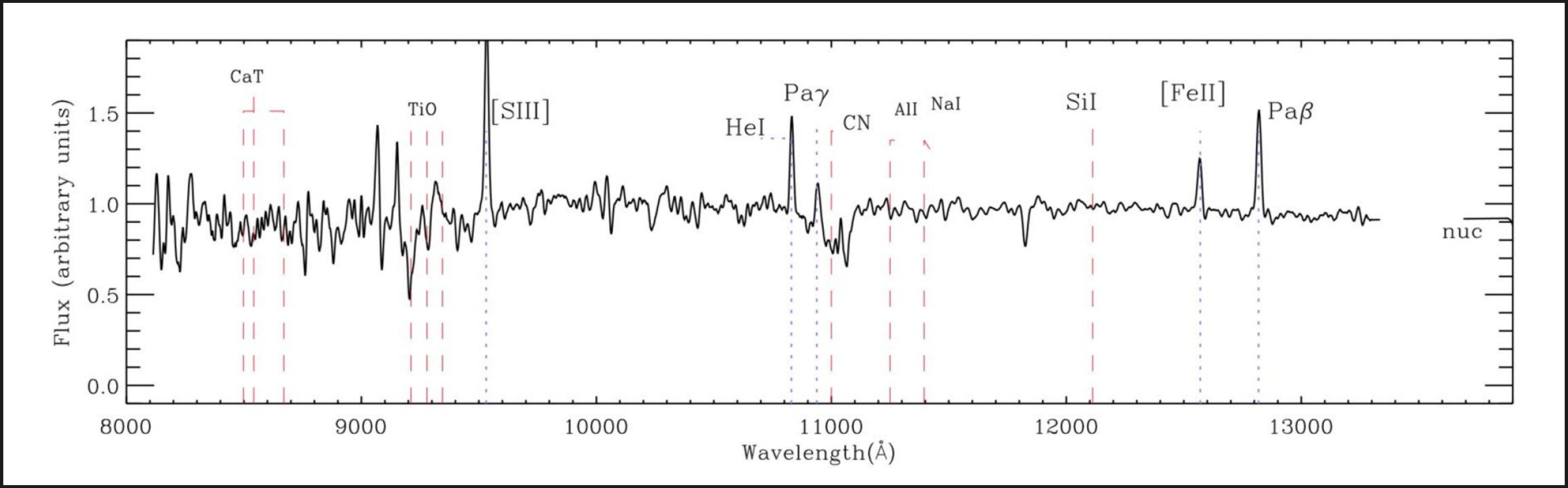


z = 10.38, tentative [OIII] at <4sigma

JWST is the right redshift machine for these obscured massive sources!

For the most obscured and massive galaxies Hα can indeed be optically thick. But for example Paβ can be 10–30 times less attenuated than even Hα.

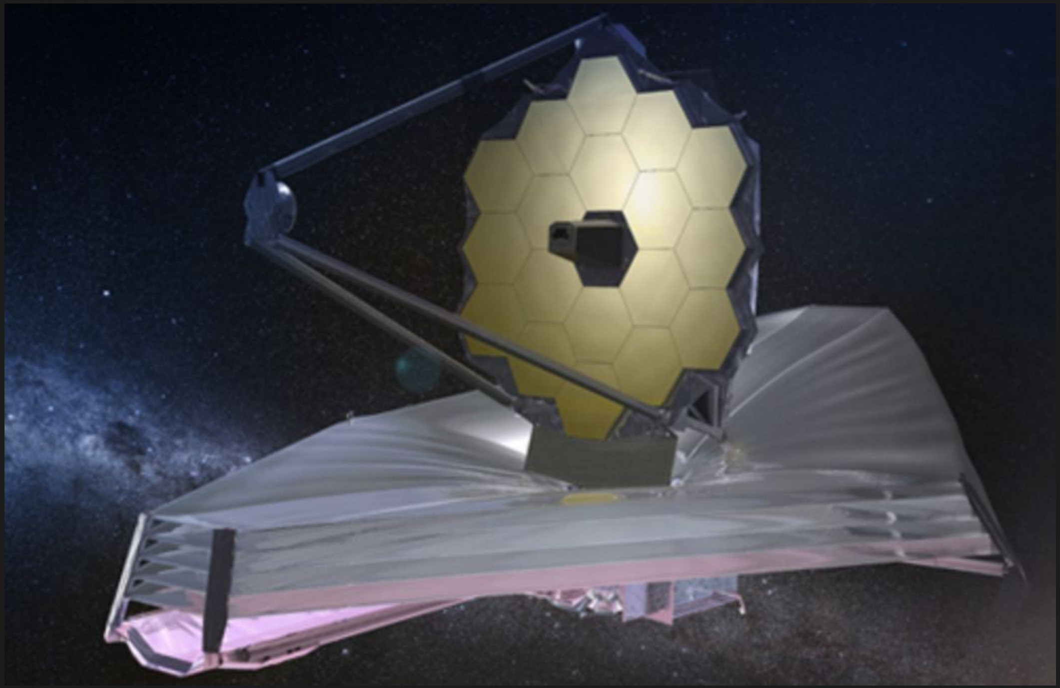
The *MIRI* spectrograph (slit or IFU modes) covers the spectral range between 5 and 25um ideal to detect near-IR lines at z≈4 and measure spectroscopic redshift.



Rest-frame near-IR spectrum of a typical local star forming galaxy, showing strong [SIII], HeI, Paγ, Paβ and FeII lines.

COMPLEMENTARY TO SPECTRAL SCAN WITH MILLIMETRIC INTERFEROMETERS!

Conclusions



We have just entered the golden age that will witness the understanding of the interplay of the physical processes that have assembled and shaped today's massive galaxies.

The synergy of EUCLID, JWST with the major past, current and upcoming facilites (in particular ALMA and ELT…) will be the stronger player of this exciting game.