

Photometric Redshifts for Roman HLIS

Brett Andrews

In collaboration with:

- **Finian Ashmead**
- **Ashod Khederlarian**
- **Yoki Salcedo Hernandez**
- **Marcos Tamargo-Arizmendi**
- **Emma Moran**

- **TQ Zhang**
- **Jeff Newman**
- **Biprateep Dey (UToronto/CITA)**



Finian Ashmead



Ashod Khederlarian

Roman Quarterly
6.23.2026

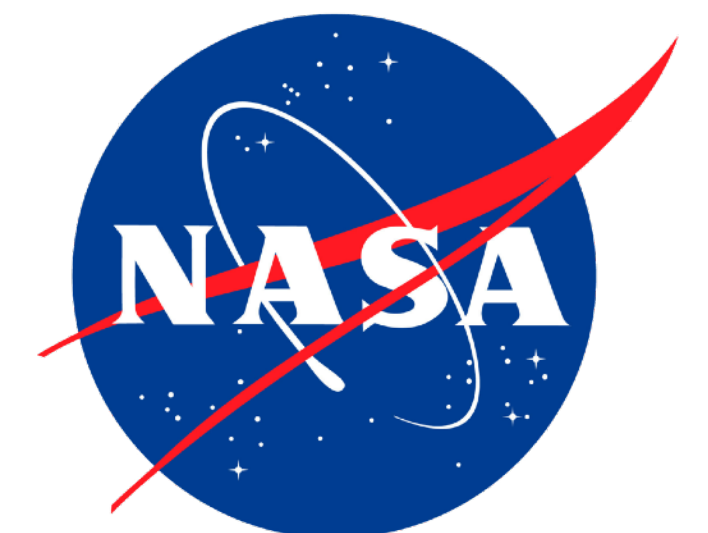


Photo-z's will be crucial for Roman HLIS science

Roman HLIS will rely on photo-z's for extragalactic analyses...

- e.g., cosmology, galaxies, galaxy clusters, supernovae, AGN, identification of transient hosts

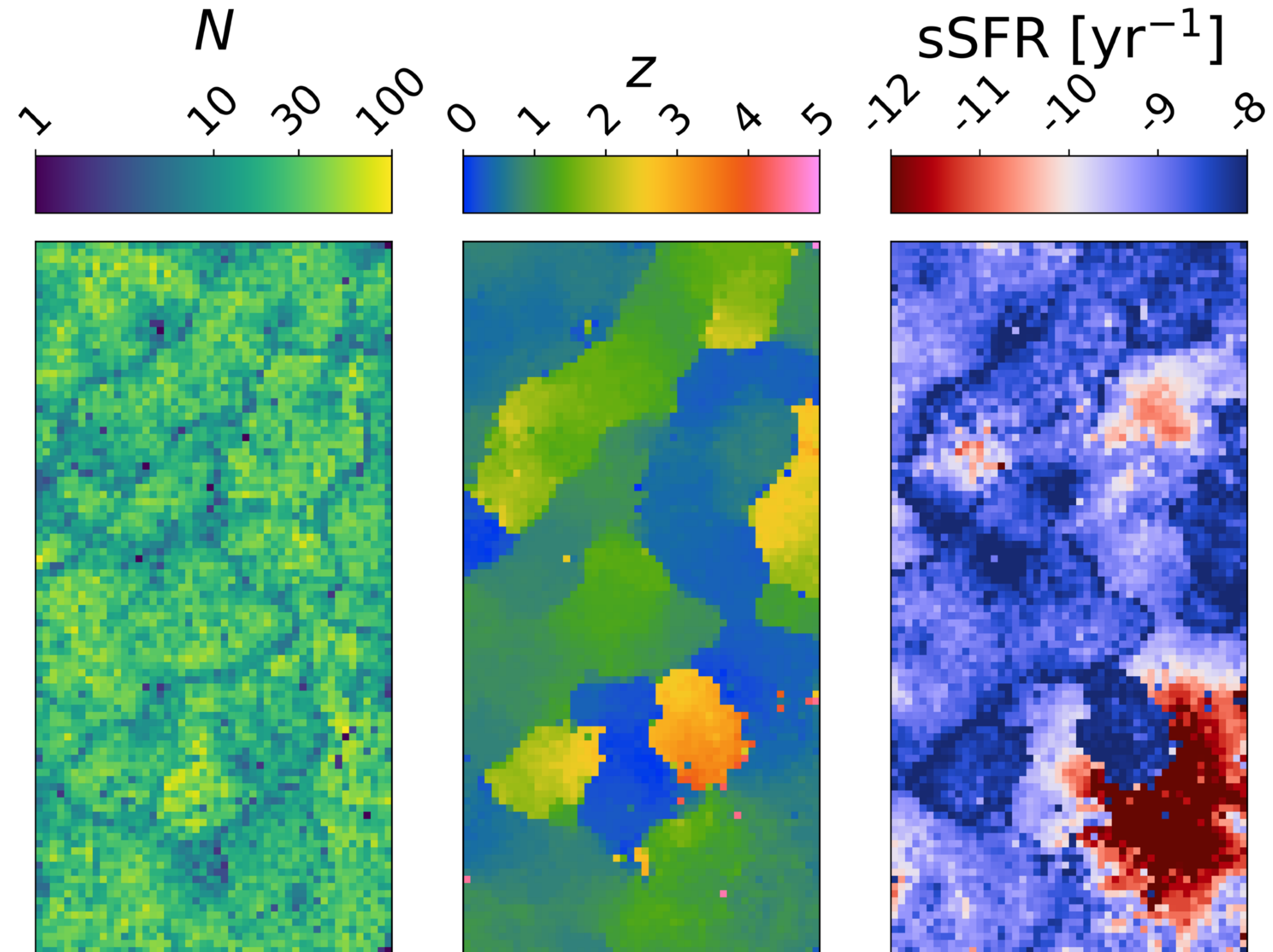
...but its revolutionary combination of breadth, depth, spatial resolution, and wavelength coverage present new challenges...

- incomplete spectroscopic training sets

...and opportunities

- deep learning approaches

Spectroscopic Incompleteness: Key Photo-z Calibration Challenge

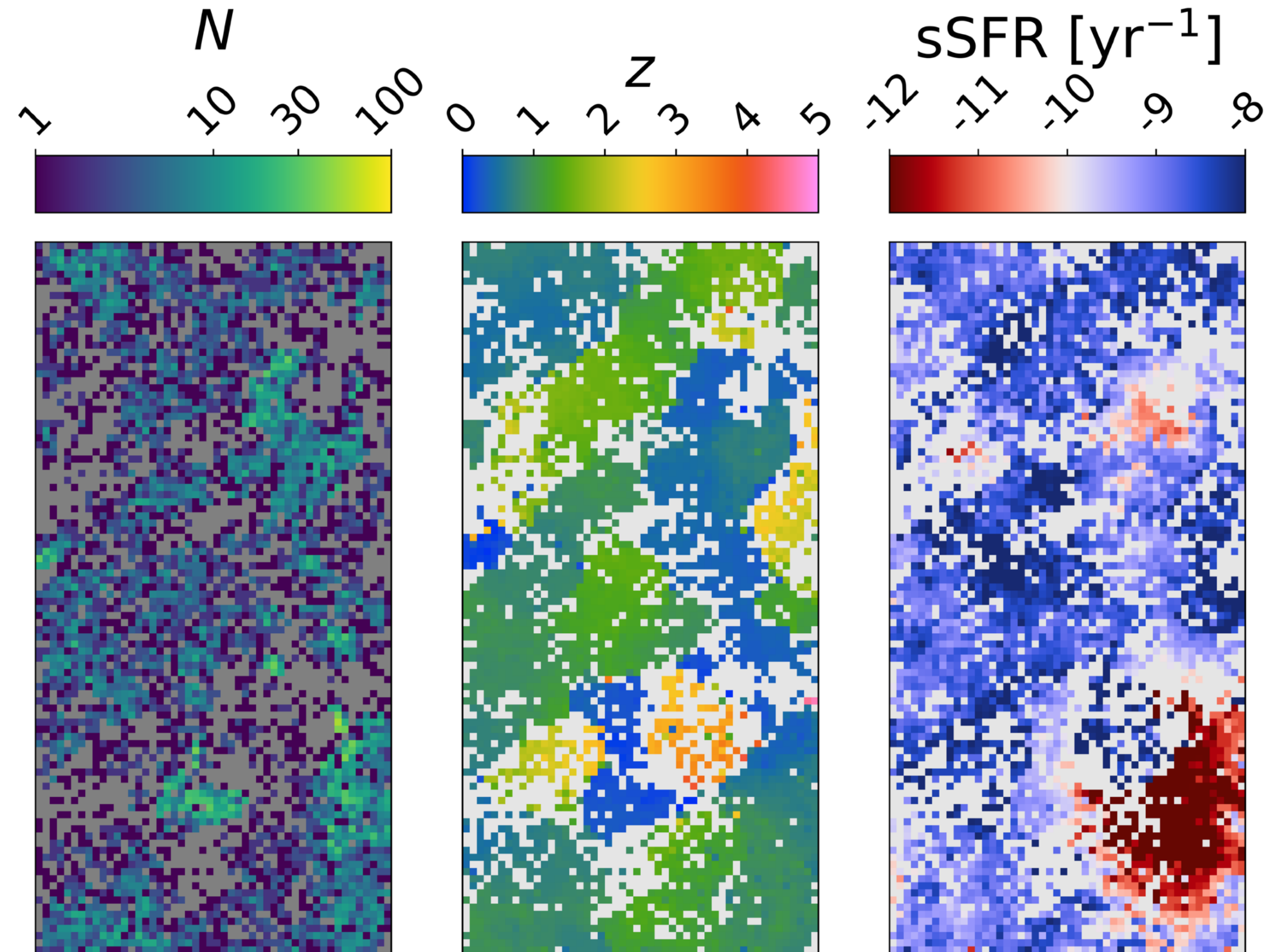


Photometric Objects

LePHARE many-band z_{phot}

COSMOS2020 $u^*grizyJH$ + LePHARE $sSFR$

Spectroscopic Incompleteness: Key Photo-z Calibration Challenge



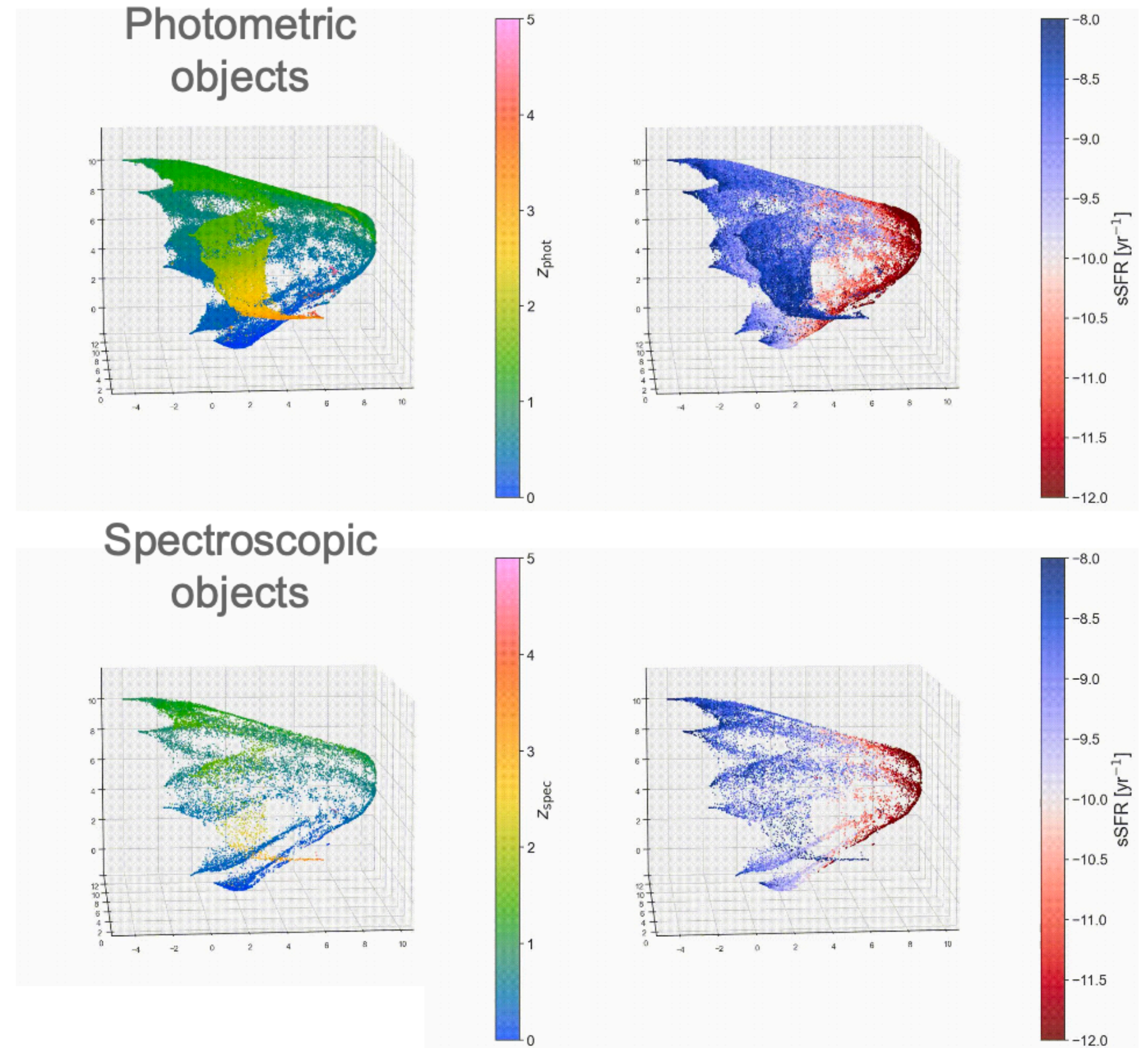
Spectroscopic Objects

z_{spec} (confidence > 95%) from Khostovan et al. (2025)

COSMOS2020 $u^*grizyJH$ + LePHARE sSFR

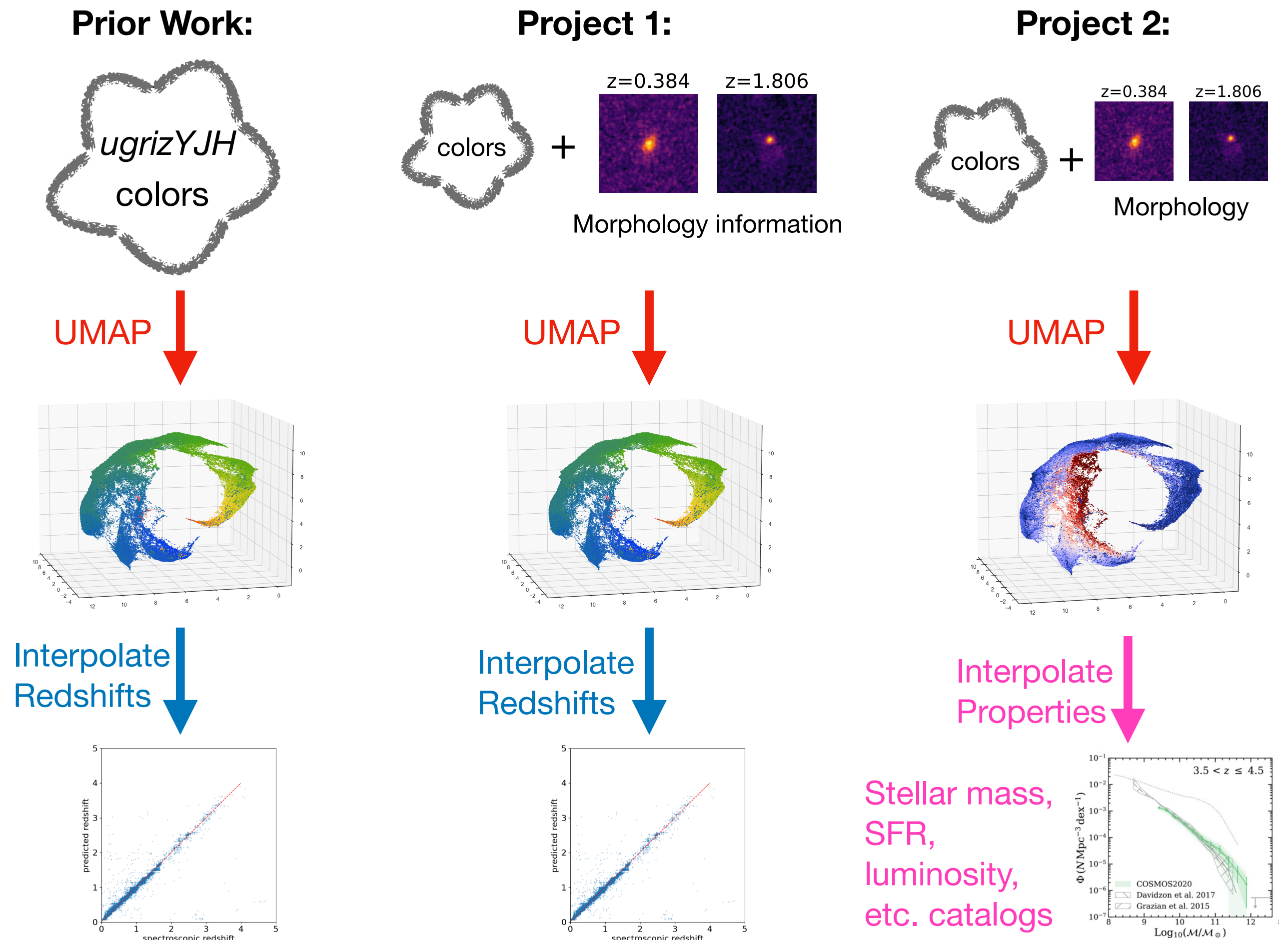
Creating a Representative Spec-z Training Set

- **UMAP** as a SOM-alternative for dimensionality reduction of $u^*grizyJH$ color space
- Produces thin (almost 2-D) manifold that is monotonic in redshift and specific SFR
- Non-representative spec-z datasets sparsely populate the manifold, but in a physically-meaningful and well-behaved way!
- Application: re-weight and interpolate spec-z datasets \rightarrow as input for standard methods of estimating redshift distributions (e.g., SOMPZ)



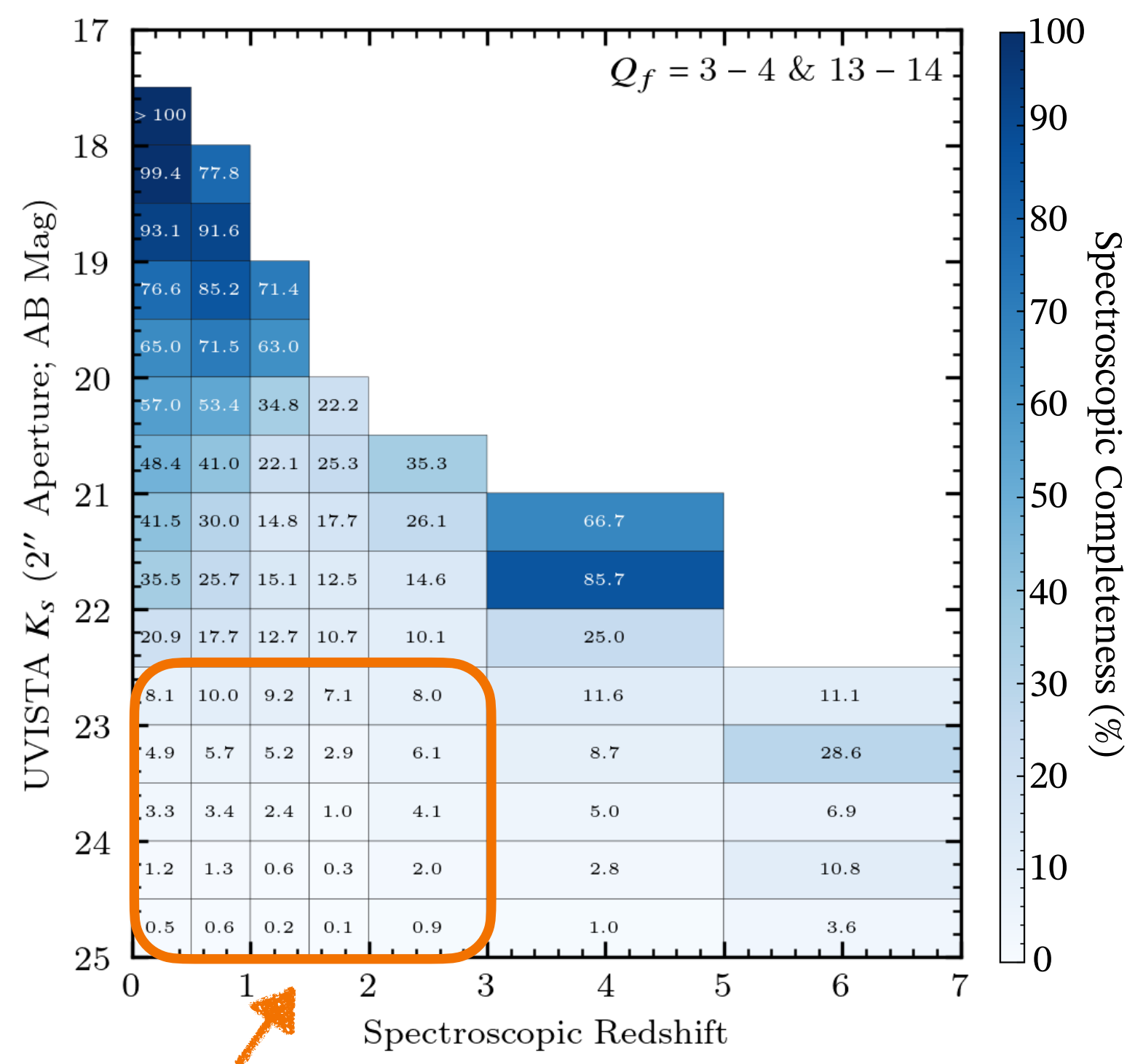
UMAP Next Steps

- implement Gaussian process regression for redshift (and uncertainty!) predictions across UMAP manifold
- fold-in morphology information to improve low- z vs. high- z disambiguation
- interpolate galaxy properties (Mstar, SFR, luminosity) across UMAP manifold



WFS 2025 (PI: Jeff Newman) "Exploiting the Intrinsic Correlations Between Galaxy Properties"

Need Deep NIR-selected Spec-z Training Sets



<10% completeness Khostovan et al. (2025)

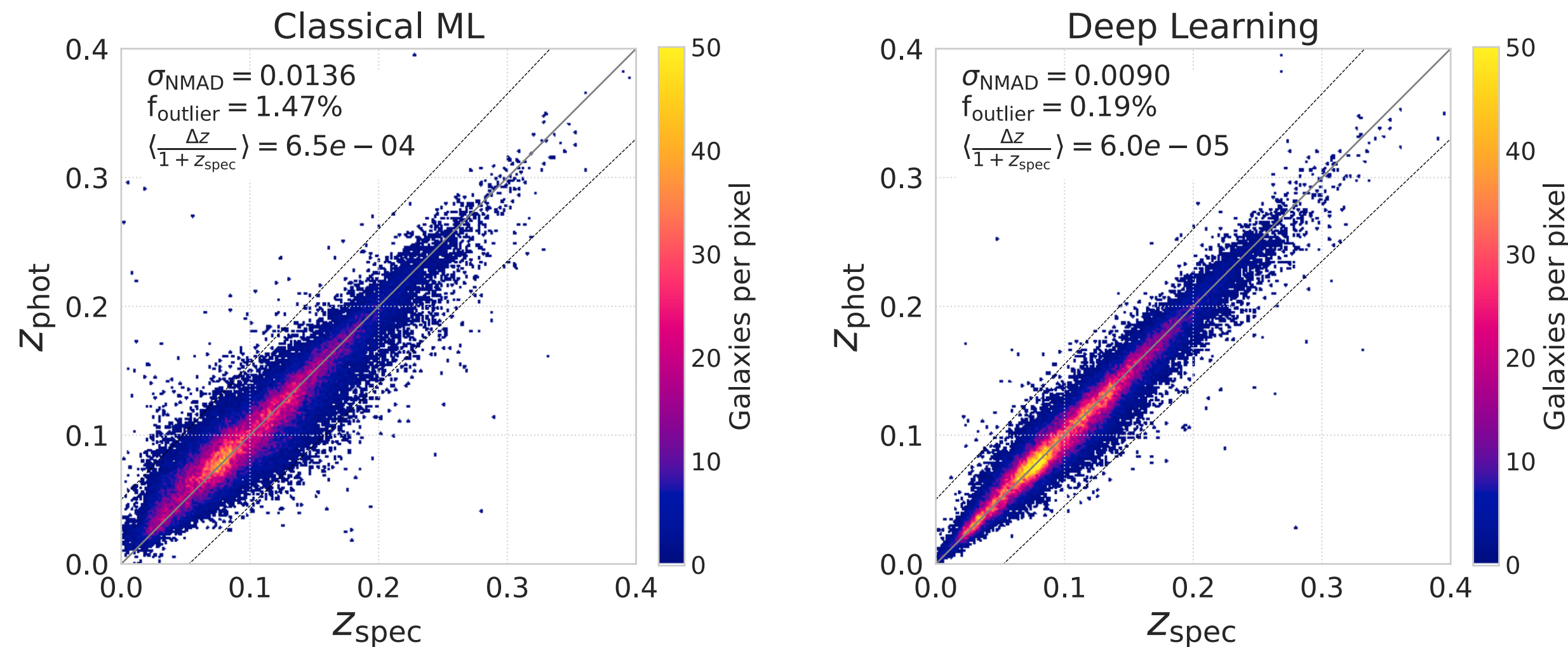
- Existing spec-z datasets sparsely cover color–magnitude–redshift space, especially at faint NIR magnitudes and $z > 1$

Subaru-PFS/Roman (SuPR) Deep Survey

- Goal: **spec-z's down to $H_{AB} \sim 24.5$** (depth of weak lensing sample) with representative colors
- Theme assigned **40 +/- 10 nights** for Roman-Subaru Synergistic Observations
- 5-15 pointings w/ **60-20 hour exposure** times for **10k-30k objects**
- COSMOS and XMM-LSS** fields (HLIS and LSST equatorial deep fields)
- Email Jeff Newman jnewman@pitt.edu by July 1 if you're interested in joining the proposal (either for the main survey or complementary observations)

Image-based Deep Learning Photo-z's

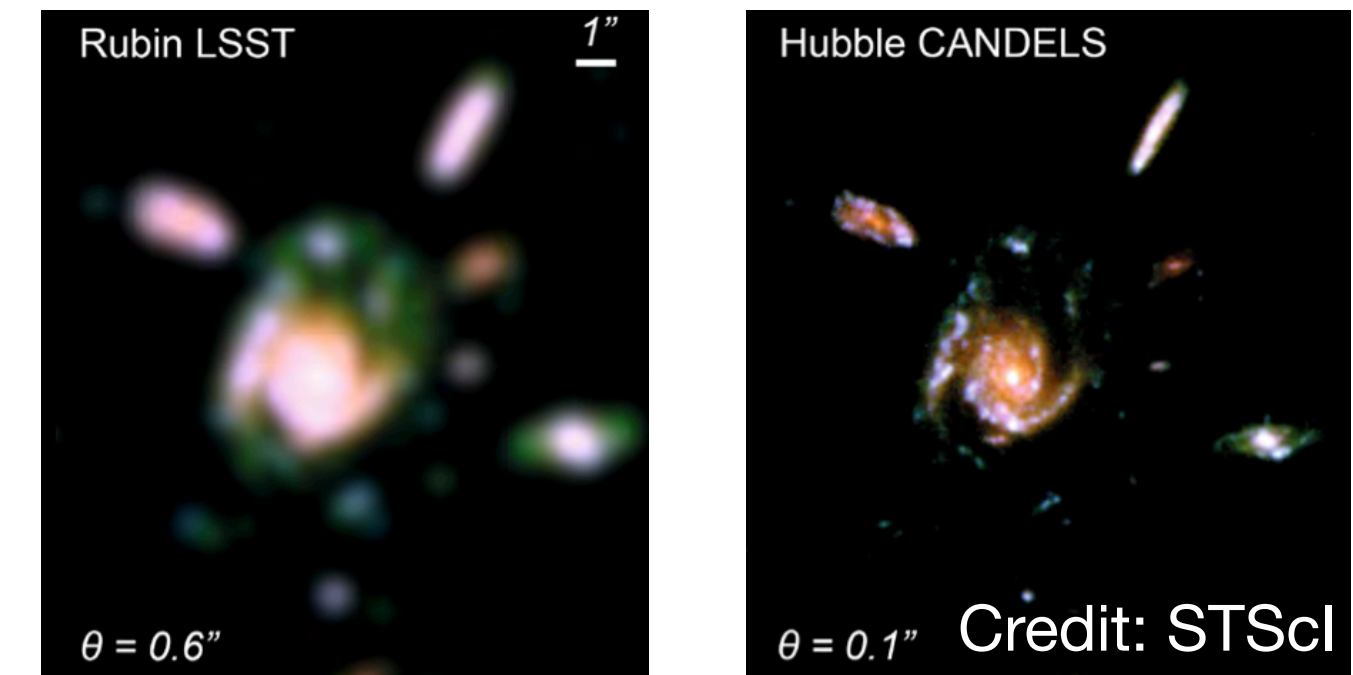
SDSS Main Galaxy Sample: deep learning models produce best photo-z's regardless of metric



photometry-based

image-based

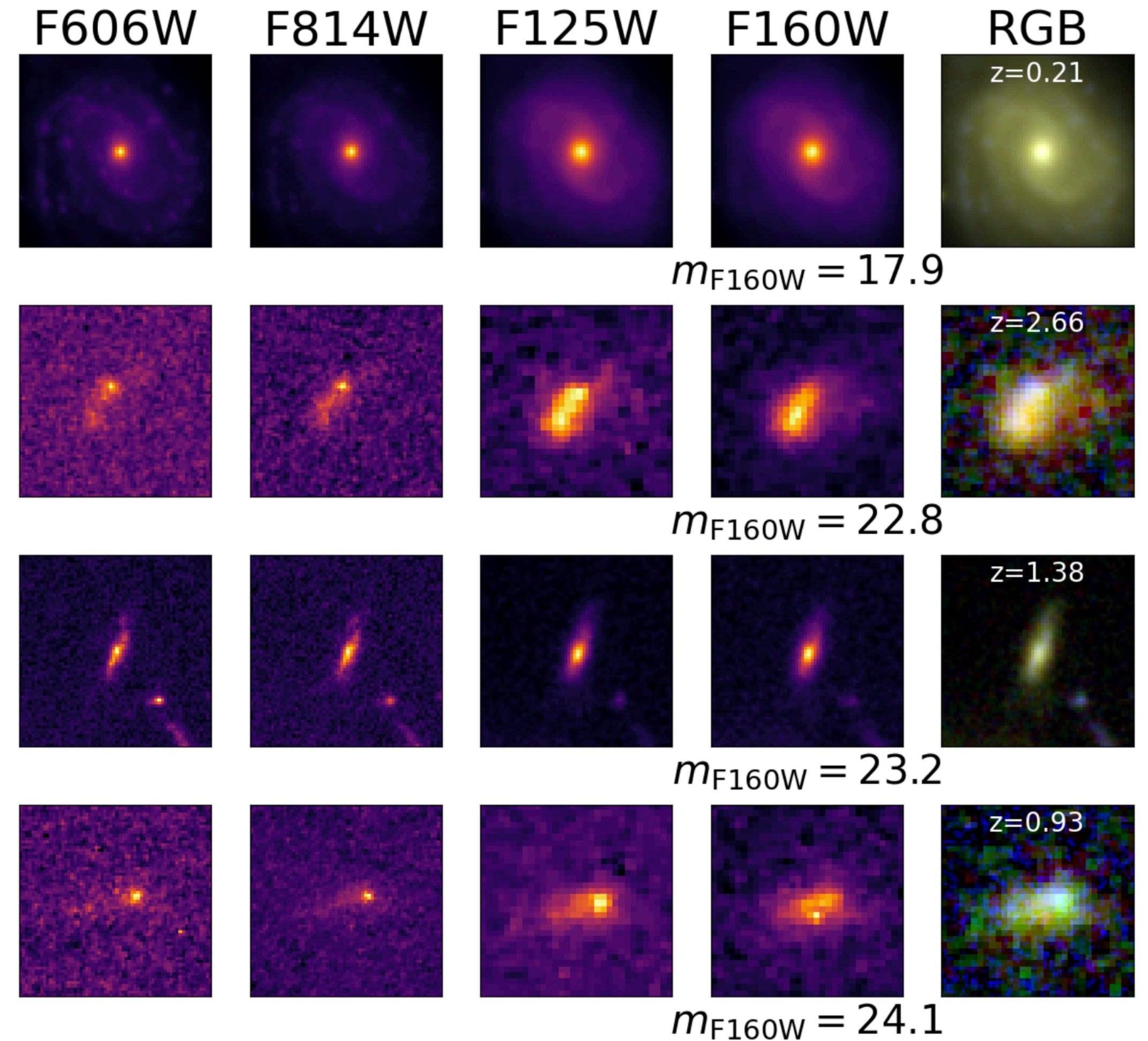
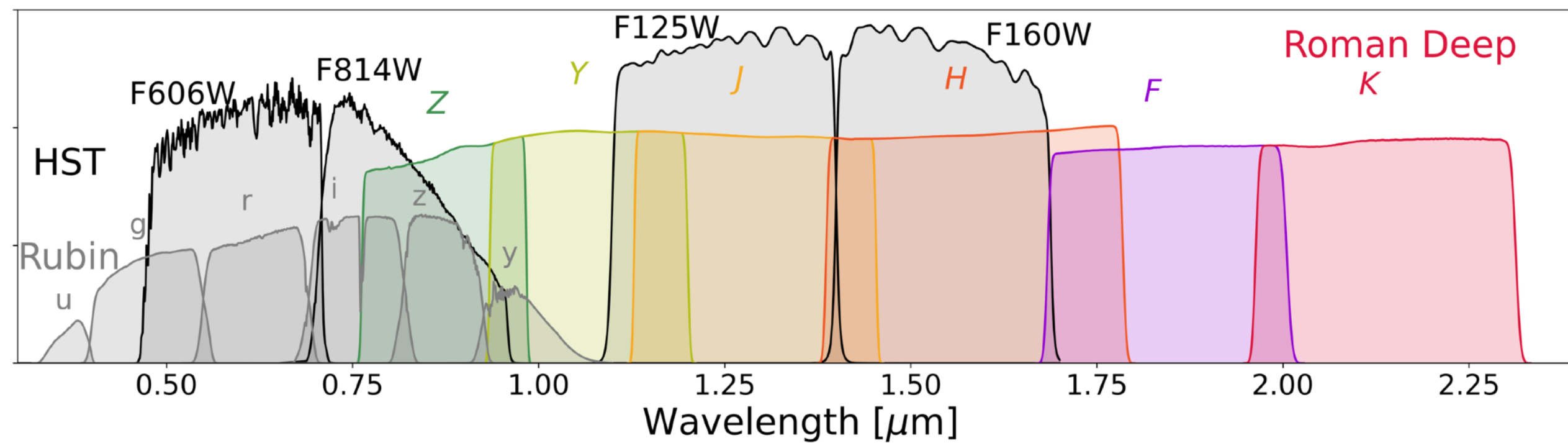
Roman+Rubin: new challenges and opportunities at high-z



- Roman: high spatial resolution in NIR, so it spans 4000 Å break to $z \sim 2.5$...
- but spec-z training sets are small and incomplete.
- Will deep learning outperform photometry-based methods?
- If so, what is the optimal deep learning approach?

Emma Moran, *BHA*, Newman, & Dey (2026) [arXiv:2507.06299](https://arxiv.org/abs/2507.06299)
Dey, *BHA*, Newman, et al. (2022)

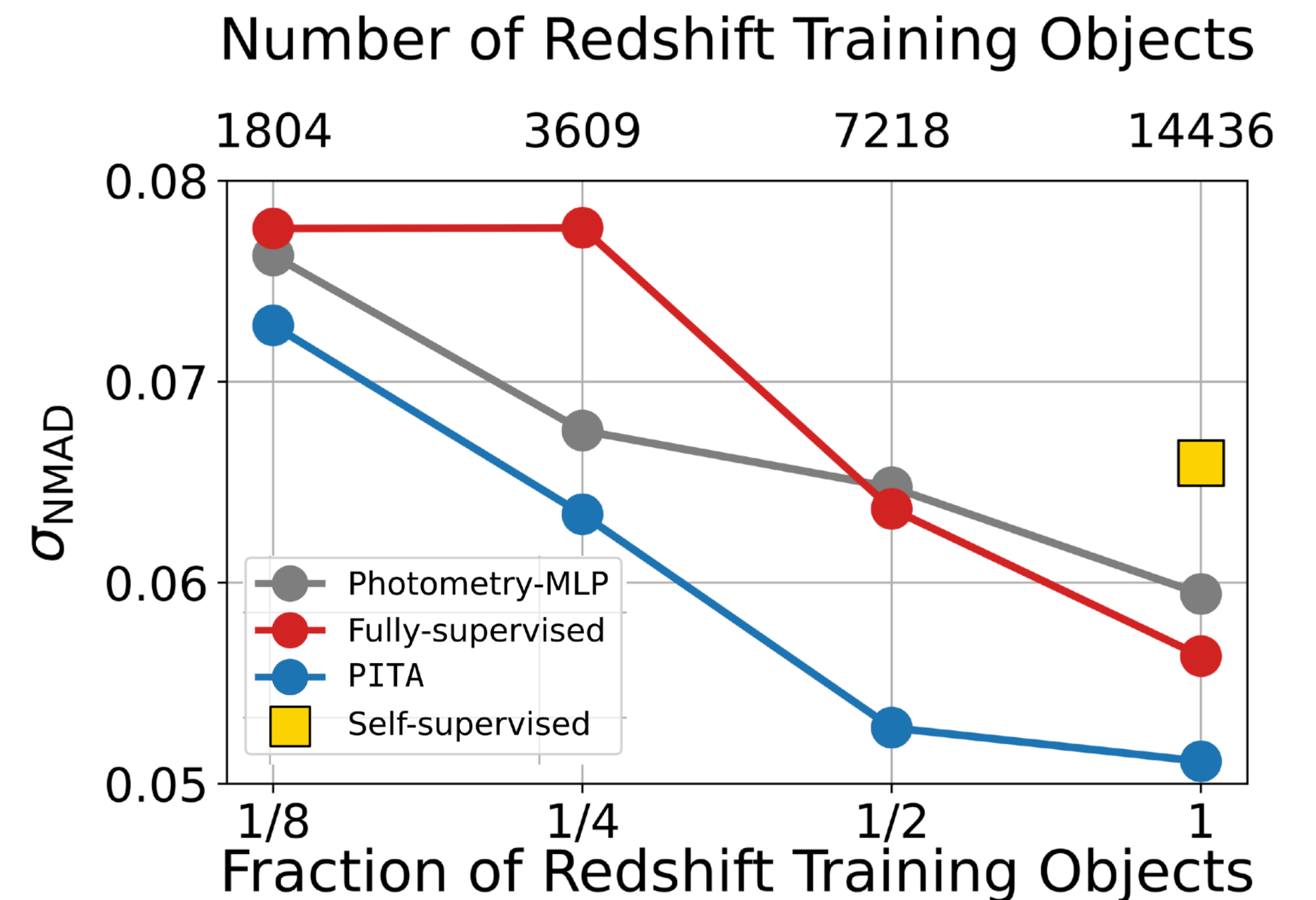
HST CANDELS: Key Roman Precursor Dataset



- Roman-like resolution in NIR
- 100k galaxy images
- 20k galaxies w/ redshifts (spec-z's, grism-z's, many-band photo-z's)

Leveraging Galaxies without Spec-z's

- **Photometry only**
- **Fully-supervised deep learning**: trained using subset of galaxies with spec-z's
- **Self-supervised deep learning**: First trained on galaxy images then fine-tuned for redshift prediction with subset of galaxies with spec-z's
- **Semi-supervised deep learning**: *simultaneously* trained on galaxy images and colors (available for all objects) and spec-z's (available for a subset of objects)



Deep Learning Next Steps

- COSMOS2025 HSC + JWST images (anticipating Rubin+Roman)
 - optical ground-based + NIR space-based imaging
- **Cal-PITA**: combines our image-based deep learning photo-z model with Cal-PIT (statistical method for producing locally-calibrated probability distribution functions)

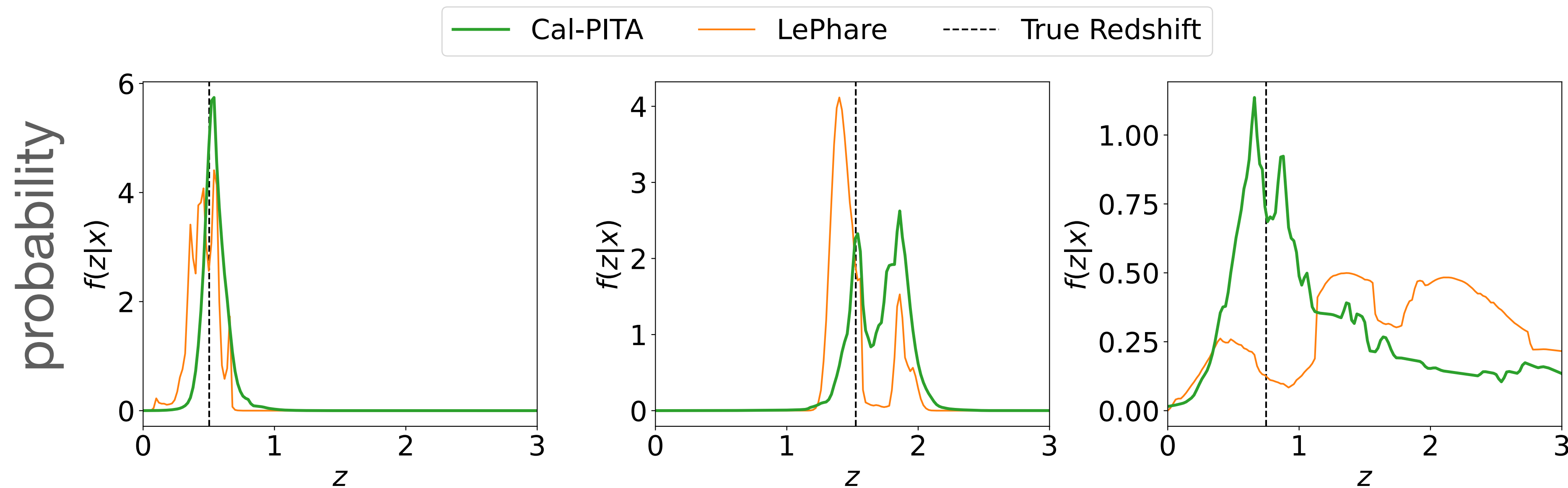
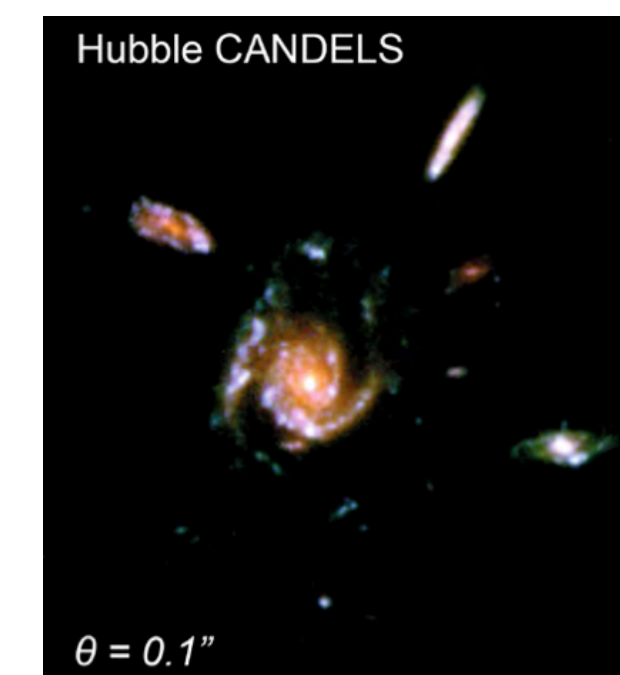
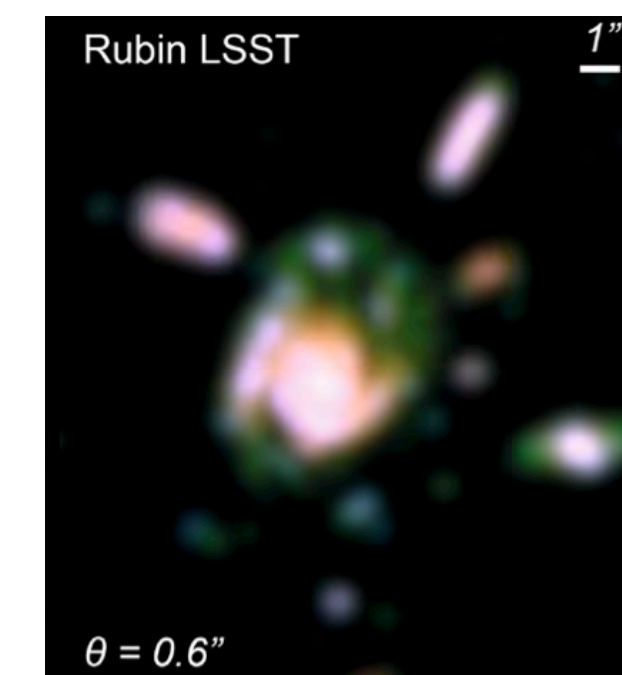
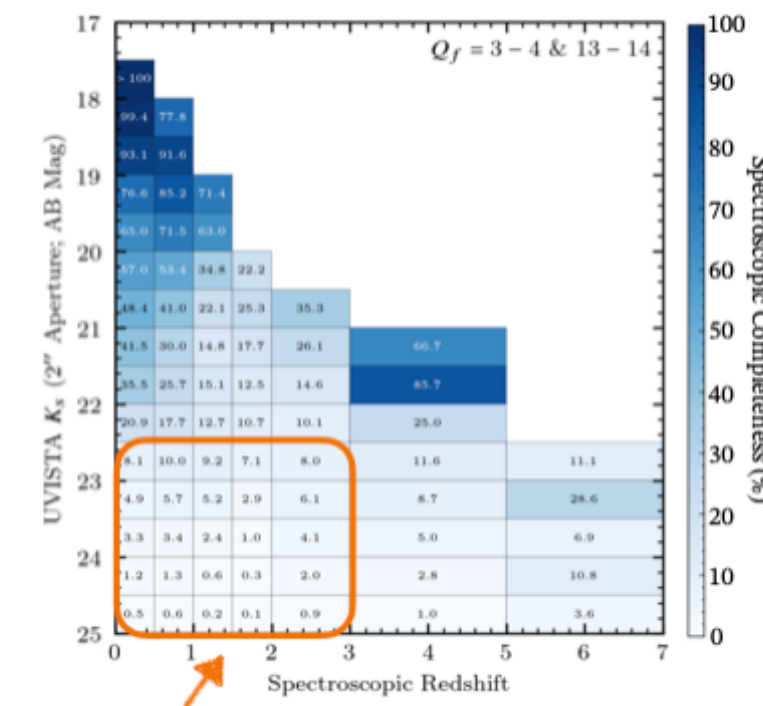
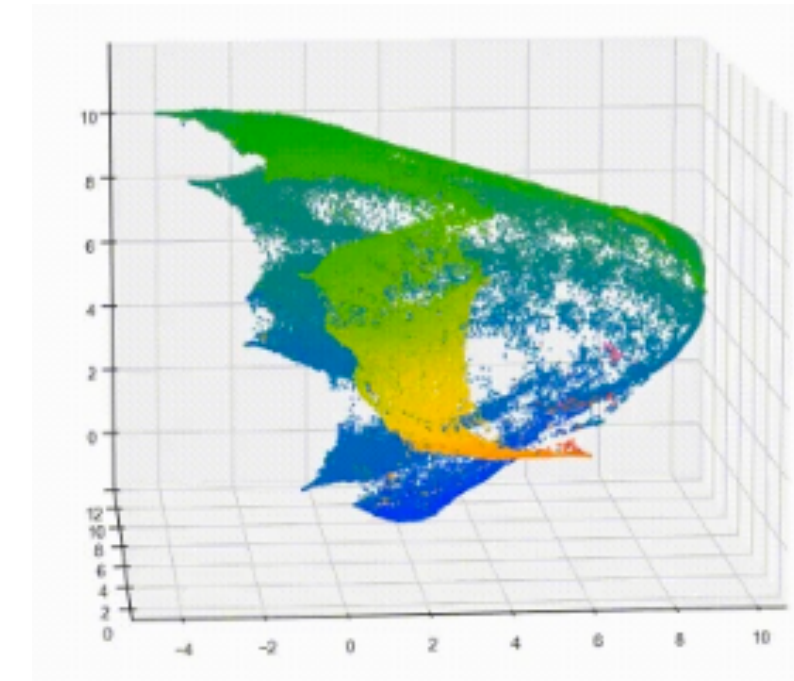


fig credit: Ashod Khederlarian

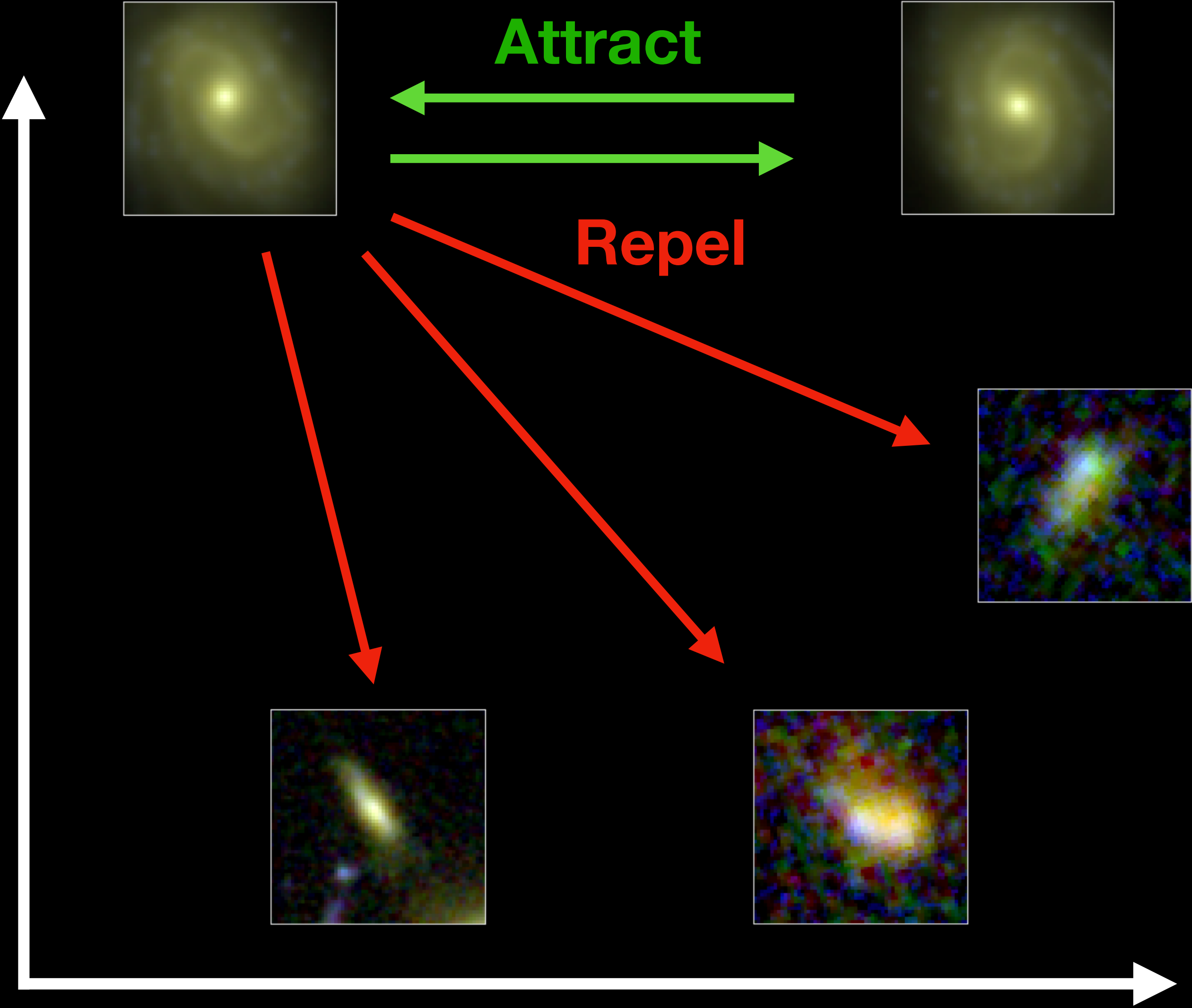
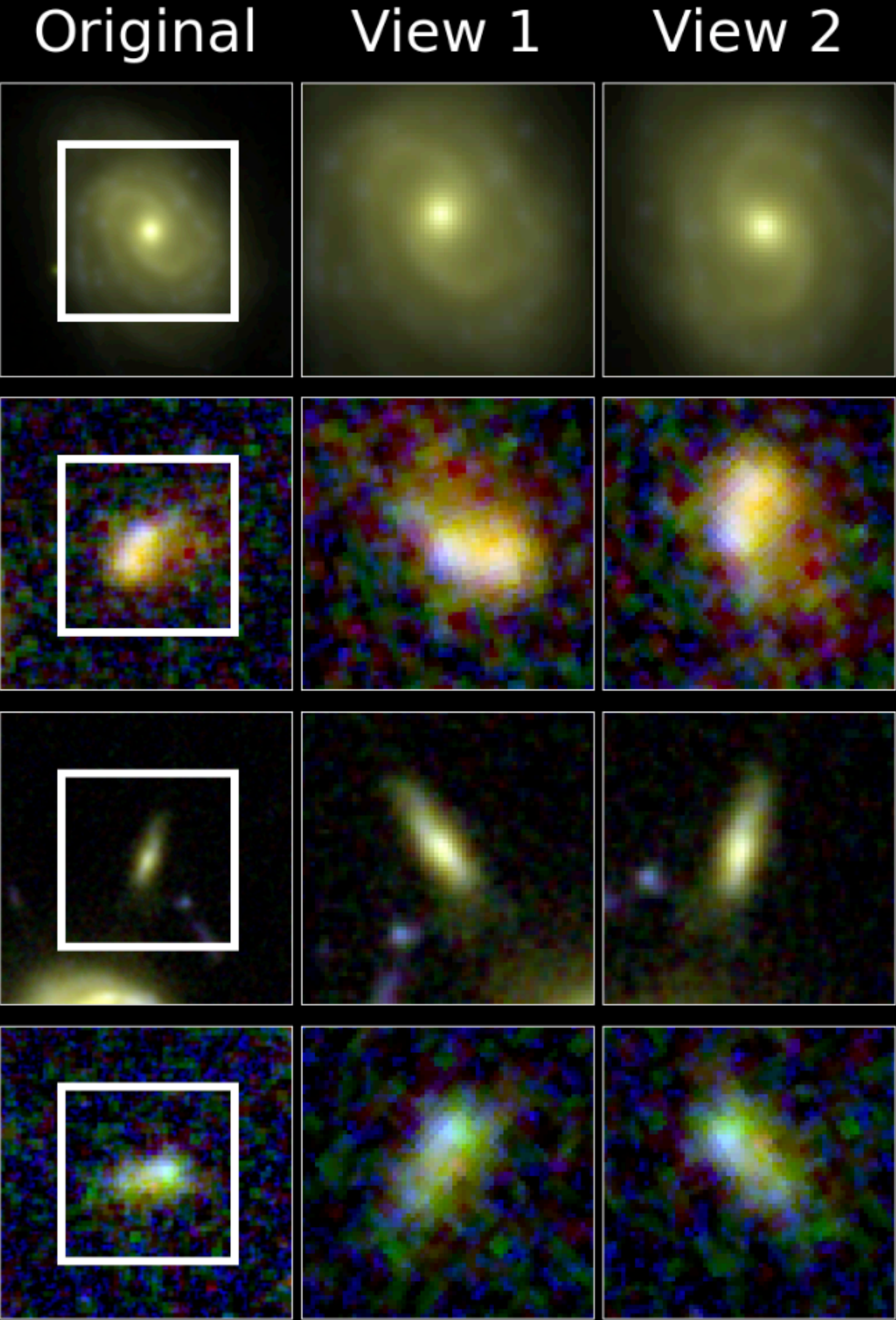
Summary

- UMAP as a SOM-alternative to optimally leverage spec-z datasets (**Finian Ashmead** et al. 2025)
- Need new deep NIR-selected spec-z training sets: **Subaru-PFS/Roman (SuPR) Deep Survey**
- Deep learning improves photo-z's for Roman-like images (**Ashod Khederlarian** et al. 2026)



BONUS

Contrastive Learning: Leveraging Galaxies w/o Redshifts

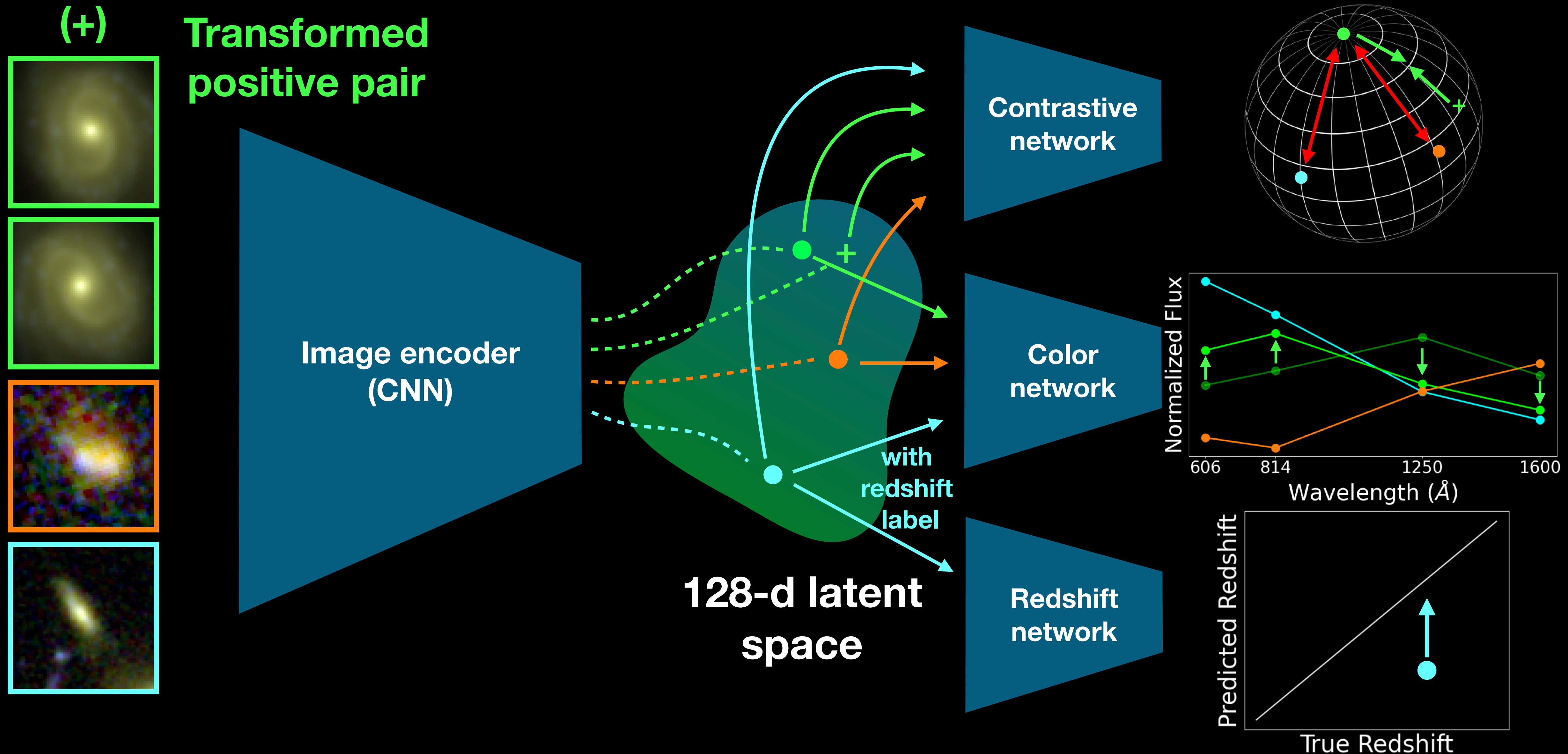


MoCo from He et al. (2020)

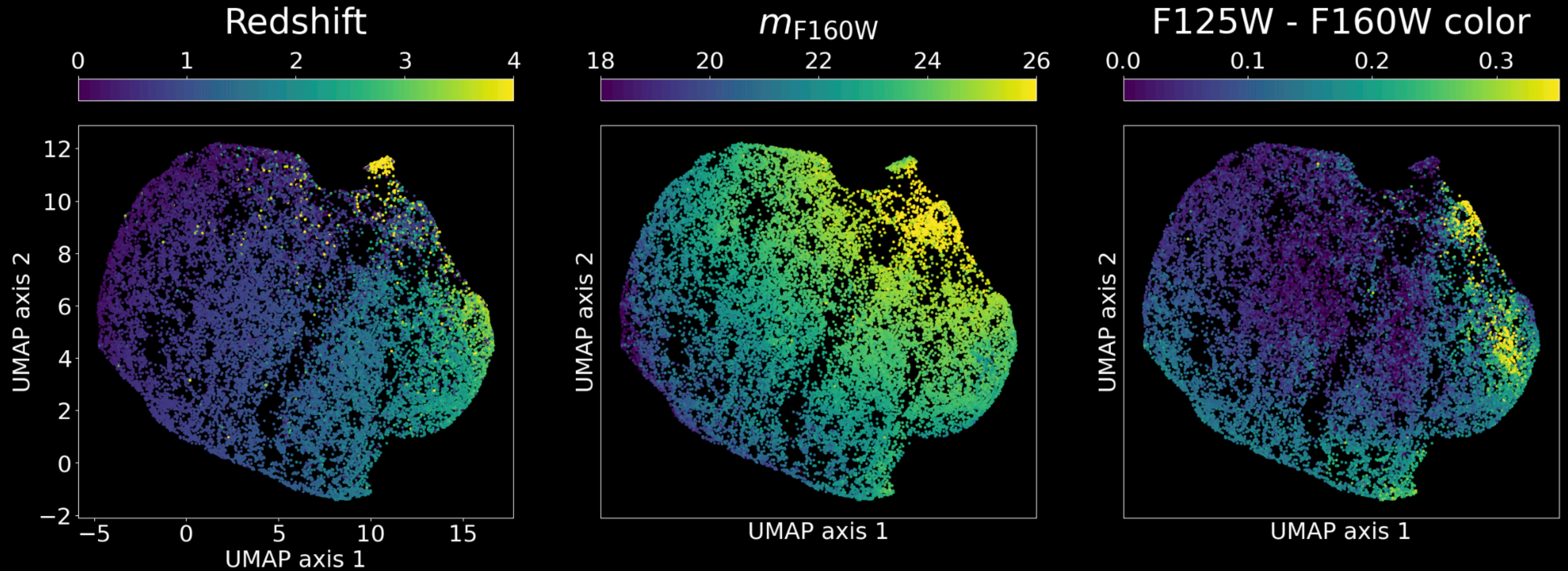
Latent Space

slide adapted from Ashod Khederlarian

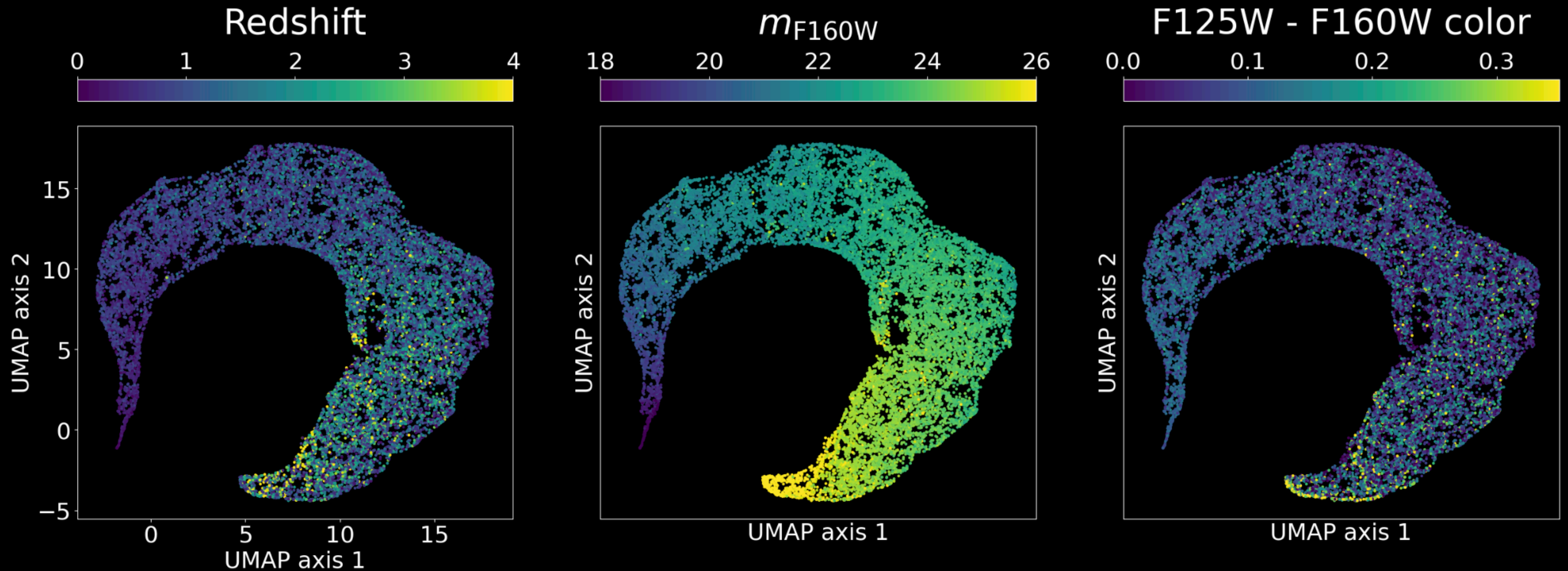
Aligning the Semi-supervised Latent Space



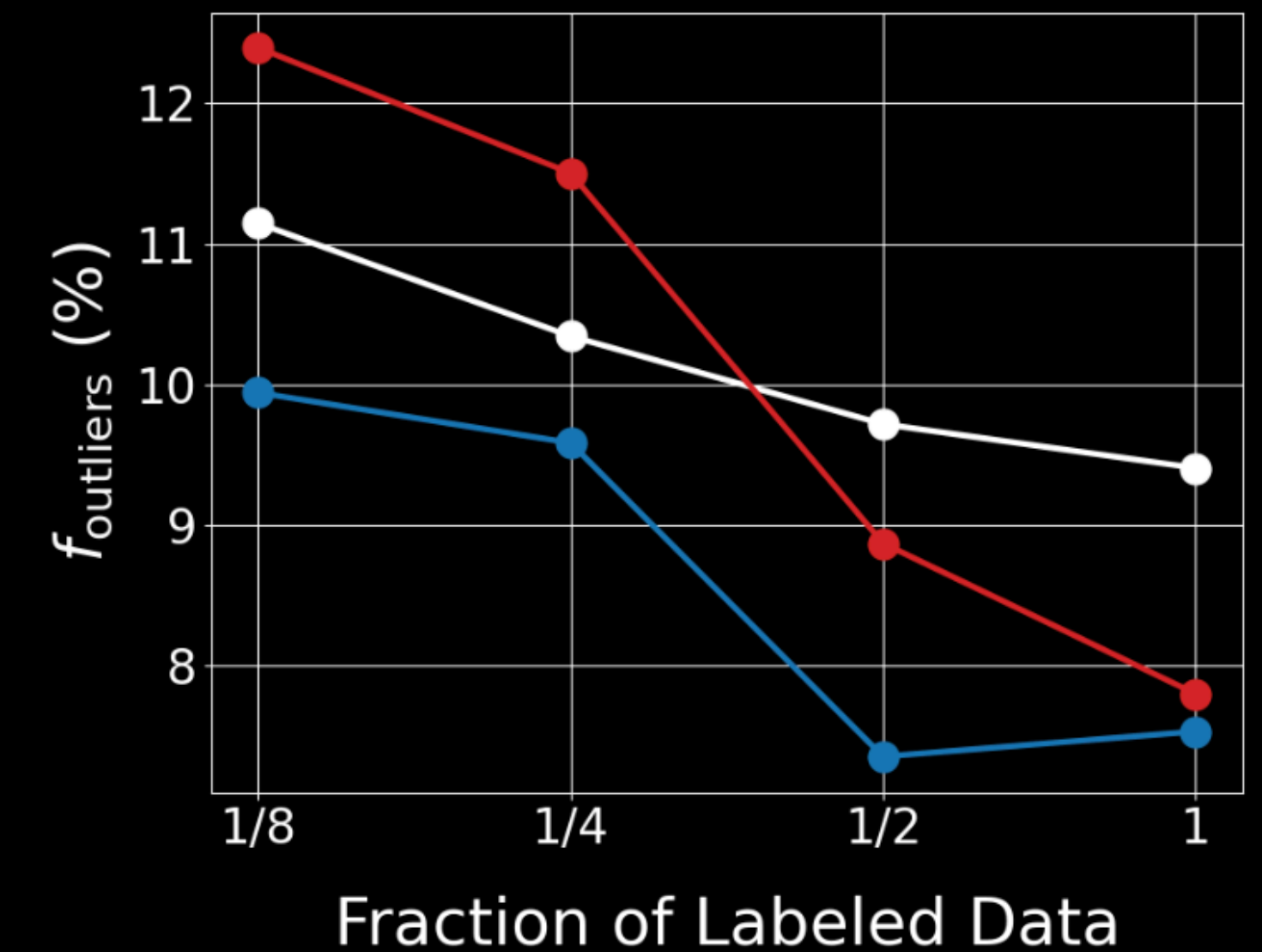
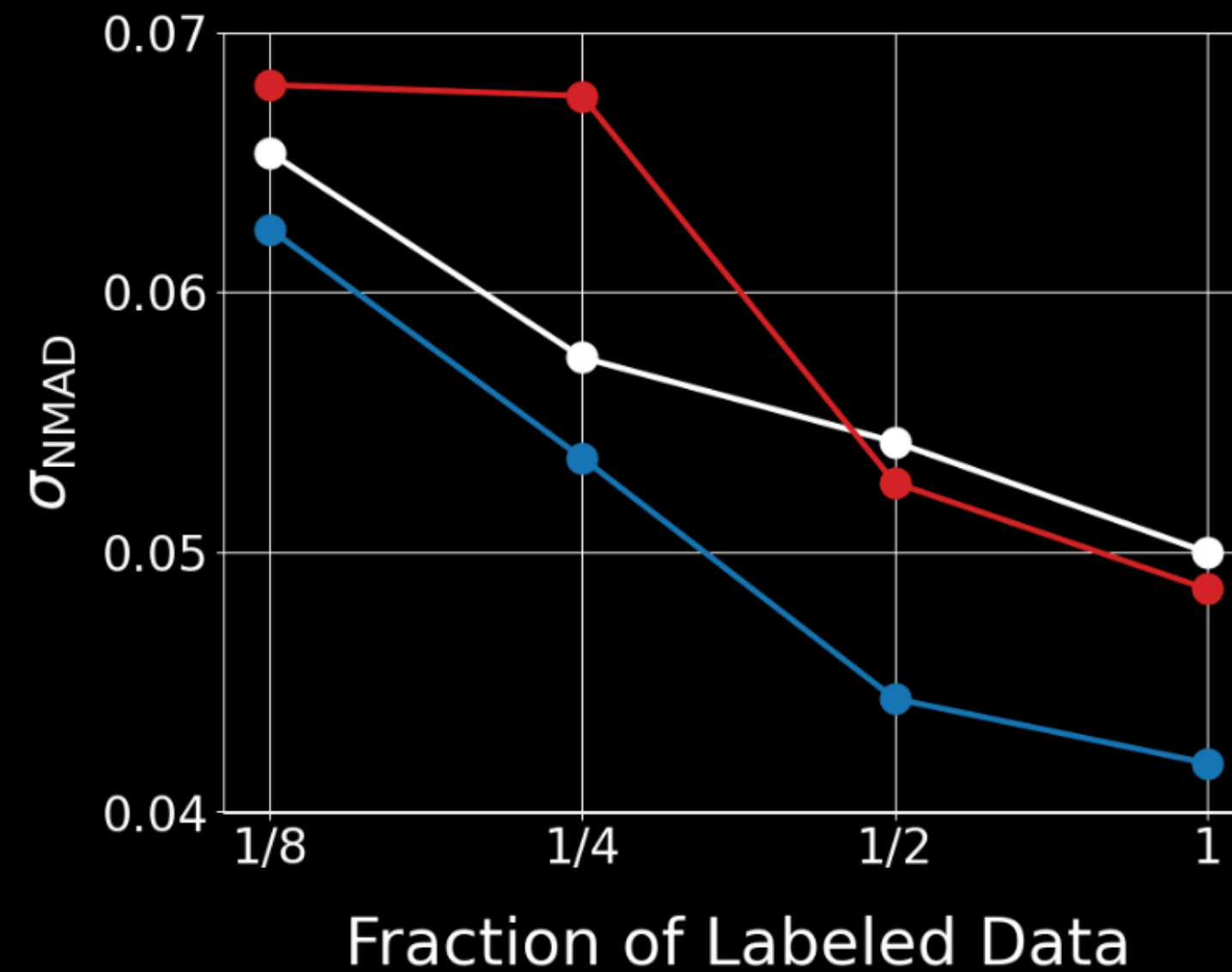
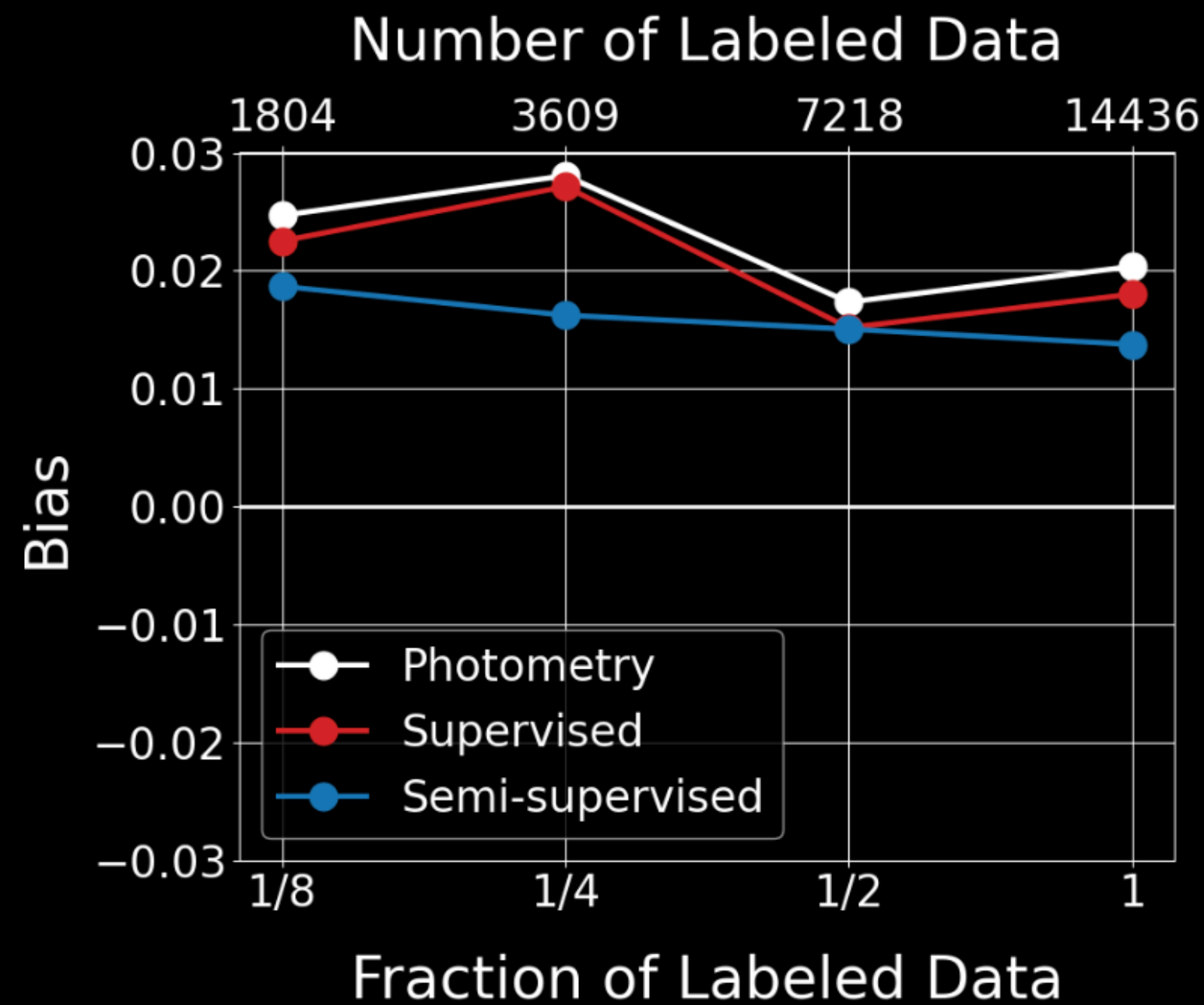
Semi-supervised Latent Space



Self-supervised Latent Space (no color alignment)



Semi-supervised Deep Learning Approach: Best Performance for ALL Redshift Training Set Sizes



Photometry-only performance

$$\Delta z = \frac{z_{\text{pred}} - z_{\text{true}}}{1 + z_{\text{true}}}$$

bias = $\langle \Delta z \rangle$

$\sigma_{\text{NMAD}} = 1.48 \times \text{median}[\Delta z - \text{median}(\Delta z)]$

fraction of $\Delta z > 0.15$

