Optimizing Roman Photometric Redshifts for HLIS

Brett Andrews

In collaboration with (PhD students in blue):

- **Finian Ashmead**
- **Ashod Khederlarian**
- **Yoki Salcedo**
- **TQ Zhang**
- **Jeff Newman**
- **Biprateep Dey** (UToronto/CITA)
- **Chun-Hao To** (UChicago)
- **Emma Moran** (Pitt undergrad applying to grad school in Fall)

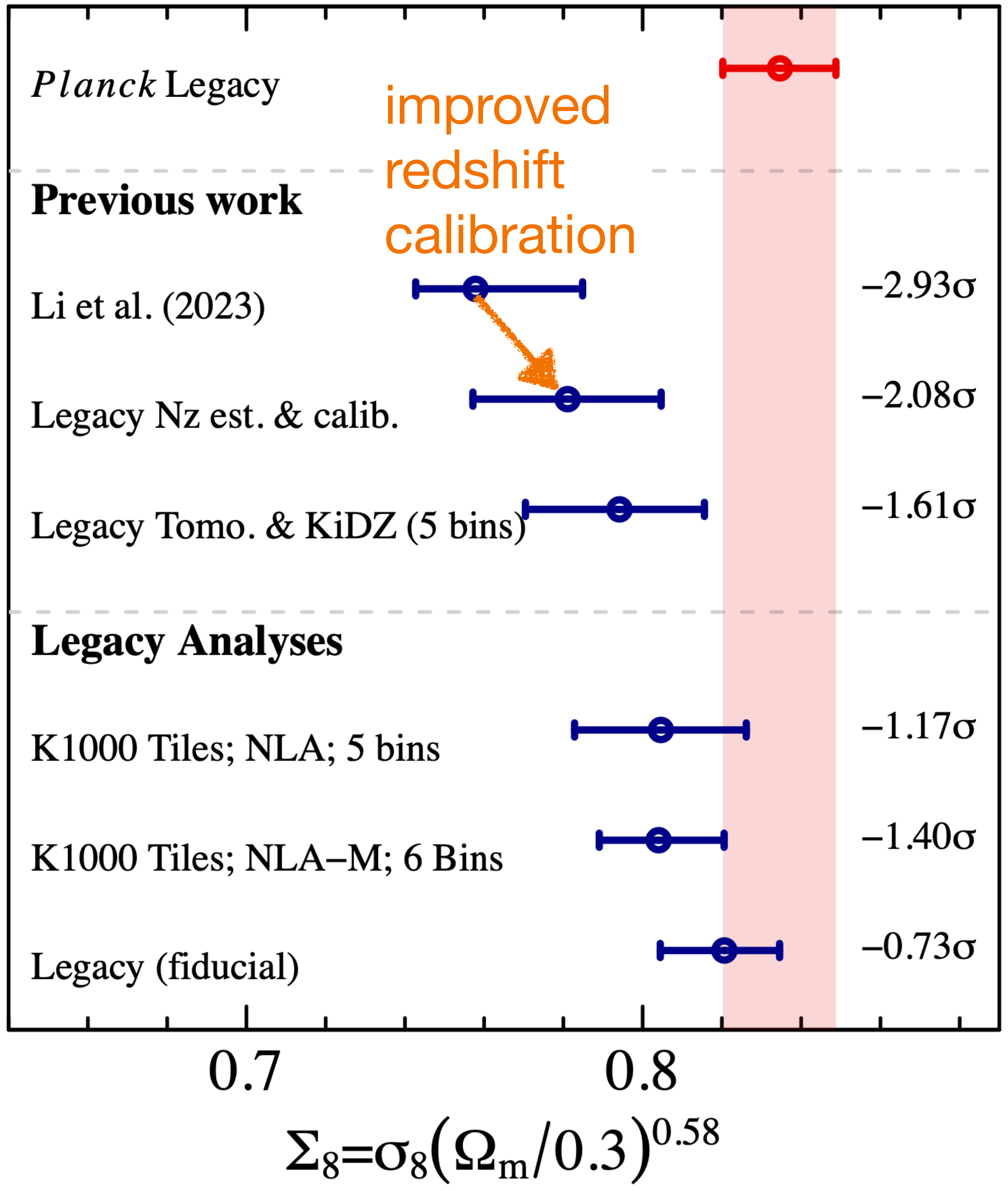
- **Roman HLIS Cosmology PIT**

University of Pittsburgh

Cosmic Cartography with Roman
7.7.2025

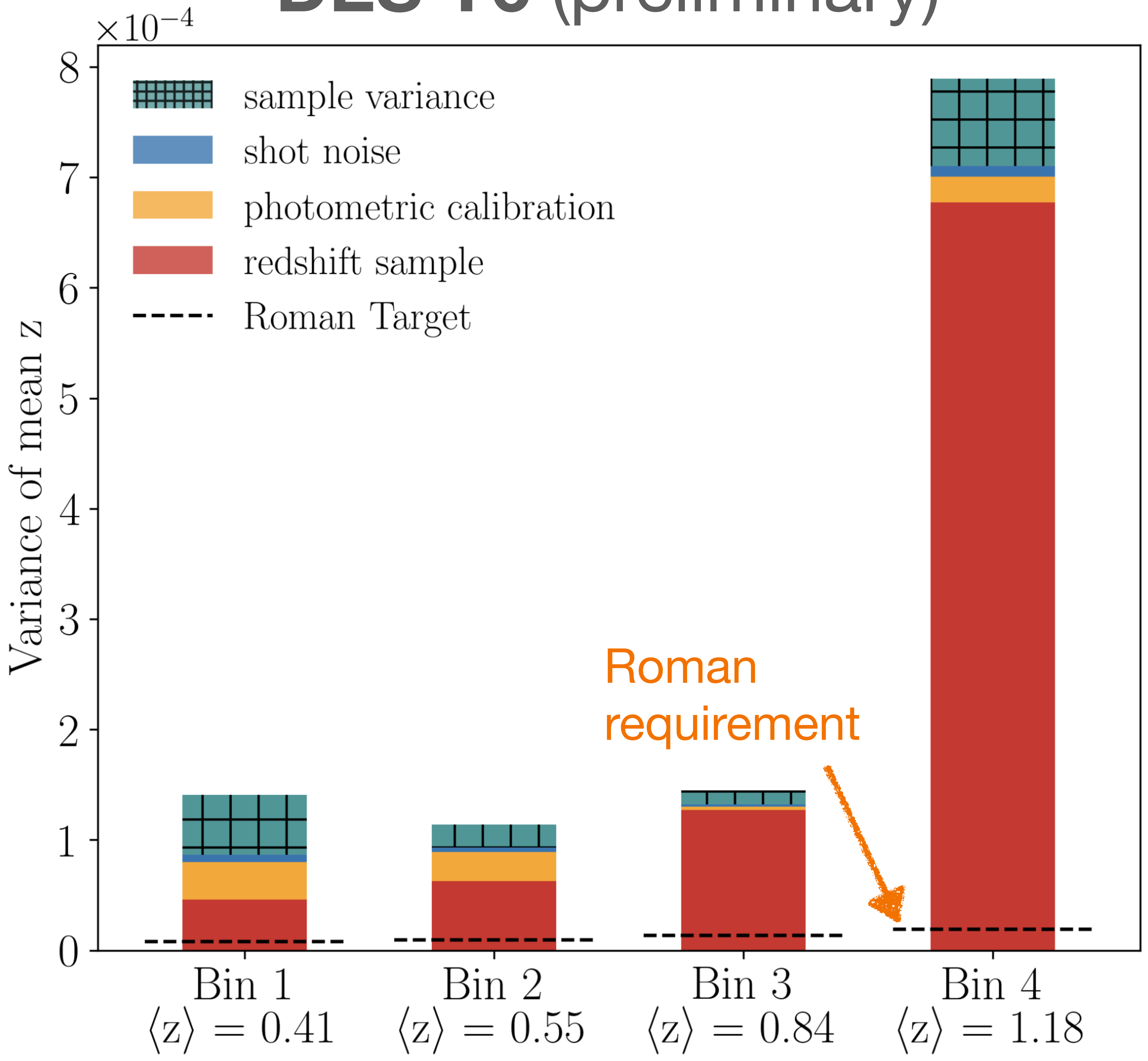
Photo-z's Crucial for Roman Weak Lensing and 3x2 pt Analyses

KiDS-Legacy



Wright et al. (2025)

DES Y6 (preliminary)



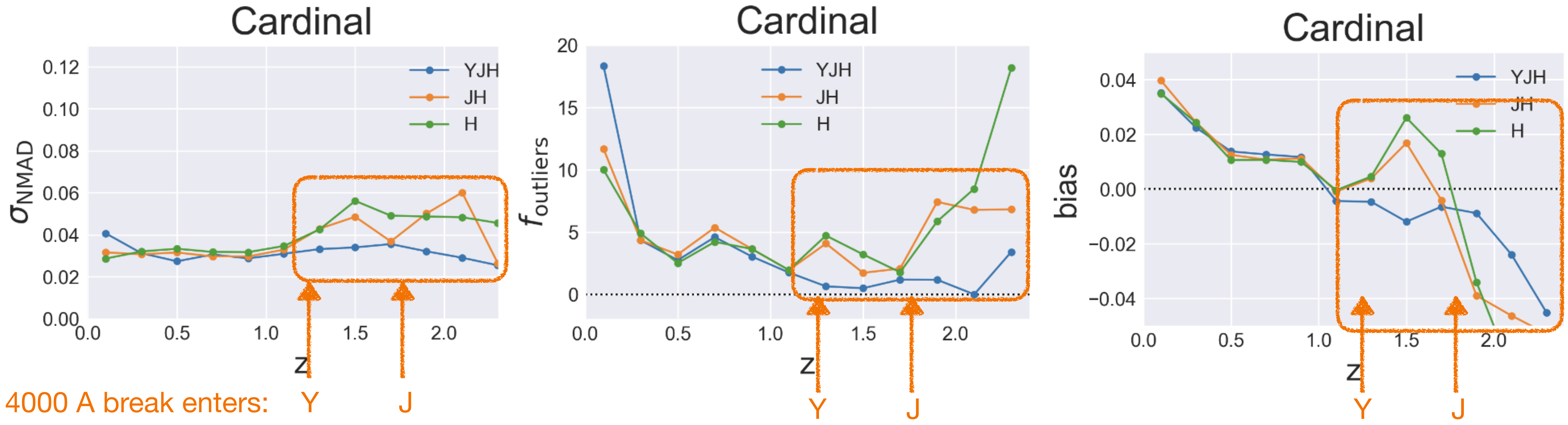
Boyan Yin et al. (in prep.) ➔ check out her talk

HLIS Survey Design: Photo-z Forecasts

- **Random Forest** (decision-tree-based ML method)
  - photometry: LSST *ugrizy* (Y4 depth) + Roman bands
  - spec-z's: representative training set w/ 20k objects *(strong assumption!)*

- Simulated and Observational Data:
  - **Cardinal** simulation (Chun-Hao To et al. 2024)
  - **OpenUniverse** simulation (aka Roman-Rubin)
  - **COSMOS2020** catalog (Weaver et al. 2022)
  - caveats: none perfect but provide sense of range of outcomes

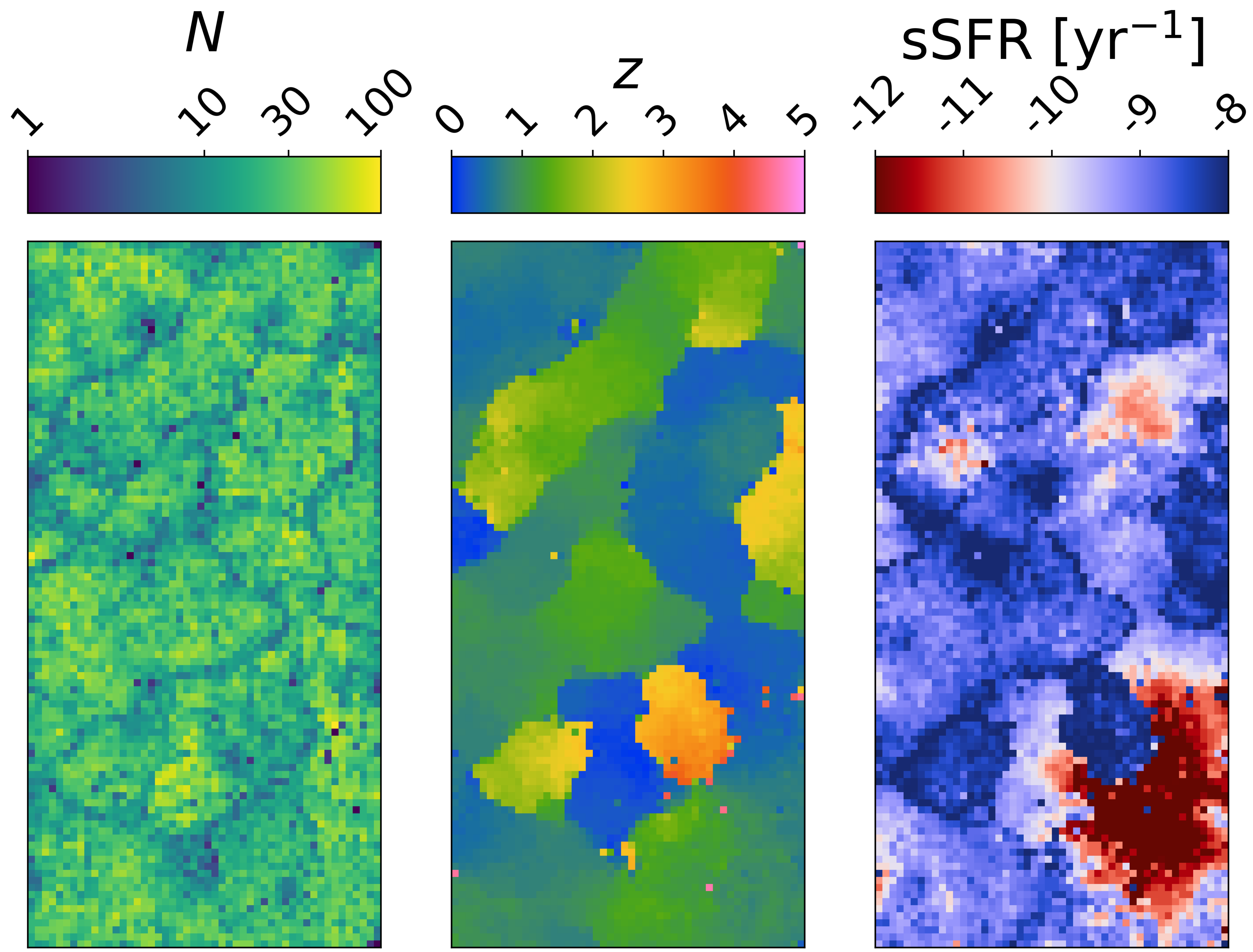
Photo-z Metrics vs. Redshift



HLIS Cosmology PIT and ROTAC Recommendations

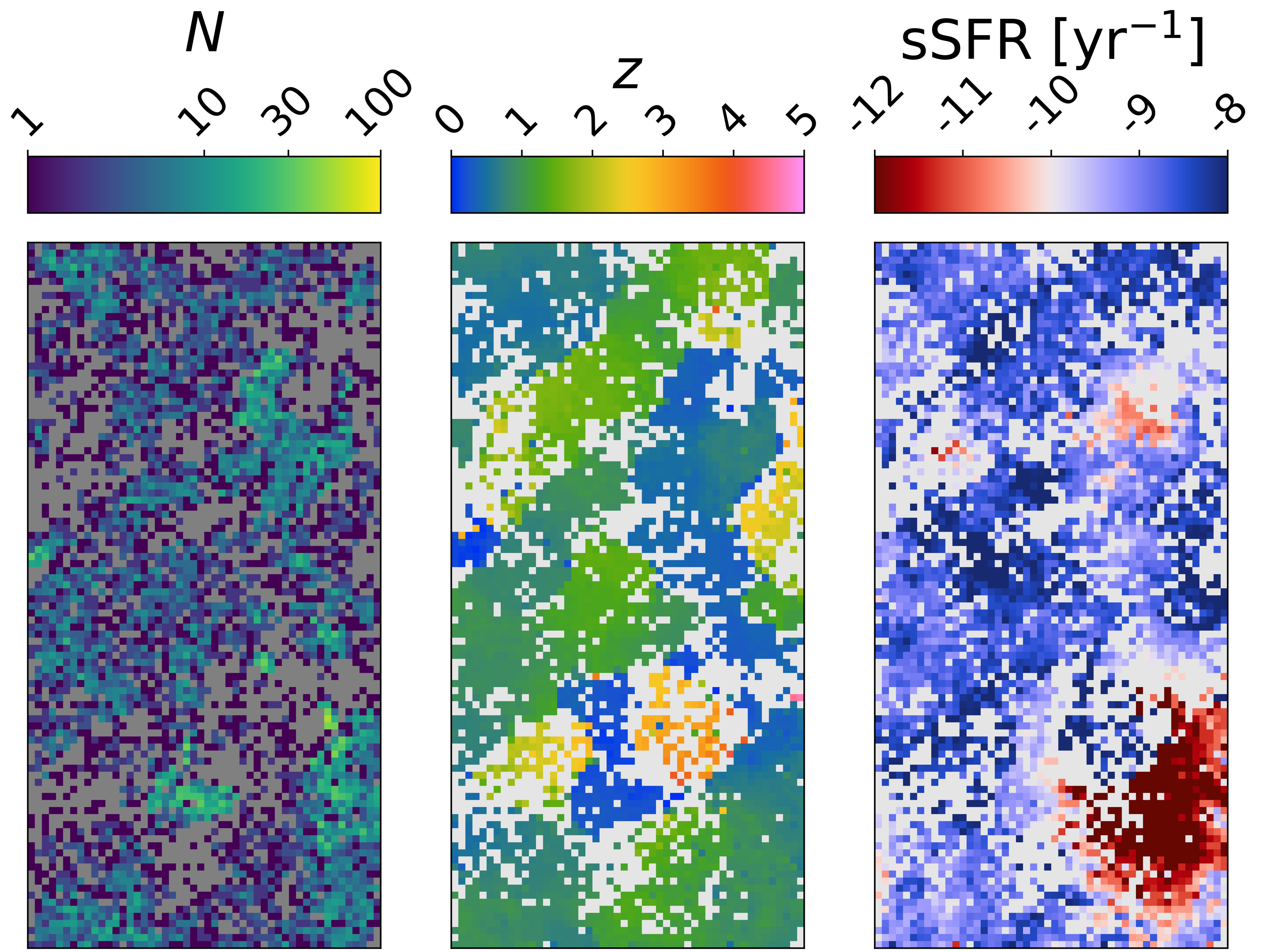
- **YJH Medium Tier**
  - dropped F184 (mostly helpful at $z > 3$, beyond lensing sample)
- **H-only Wide Tier**
  - 2x increase in area (vs. JH) with only slightly worse photo-z point estimates
  - but broader tomographic bins (**Chun-Hao To's** analysis) ➔ controlling for systematics will be essential (*will require highly complete spec-z sample*)

Spectroscopic Incompleteness: Key Photo-z Calibration Challenge



Finian Ashmead et al. (in prep.) ➔ check out his poster
COSMOS2020 *u*grizyJH* + LePHARE sSFR
LePHARE many-band z_{phot}

Spectroscopic Incompleteness: Key Photo-z Calibration Challenge

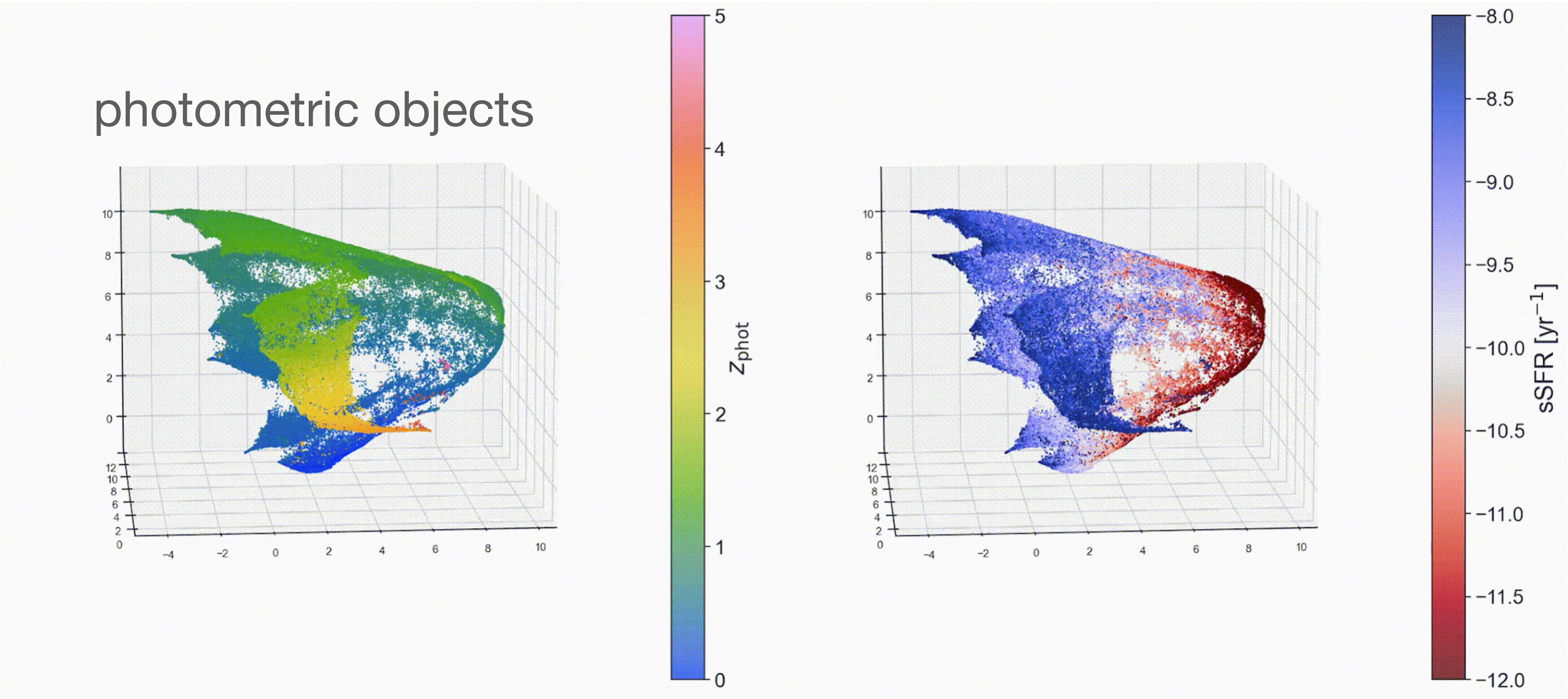


Finian Ashmead et al. (in prep.) ➔ check out his poster
COSMOS2020 *u*grizyJH* + LePHARE sSFR
z_{spec} (confidence > 95%) from Khostovan et al. (2025)

Creating a Representative Spec-z Training Set

- **UMAP** as a SOM-alternative for dimensionality reduction of *u*grizyJH* color space

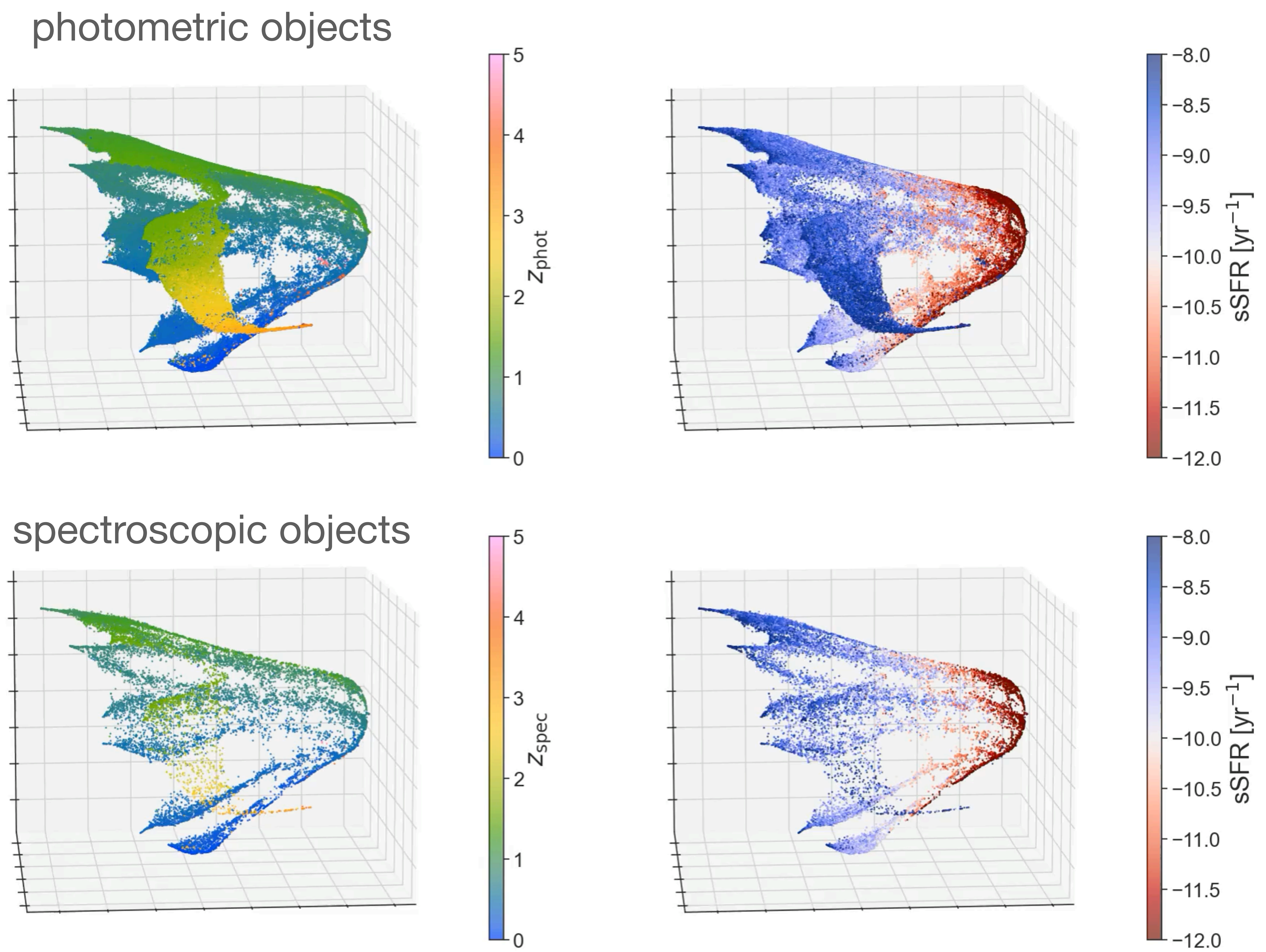
- Produces thin (almost 2-D) manifold that is monotonic in redshift and specific SFR



Finian Ashmead et al. (in prep.) ➔ check out his poster

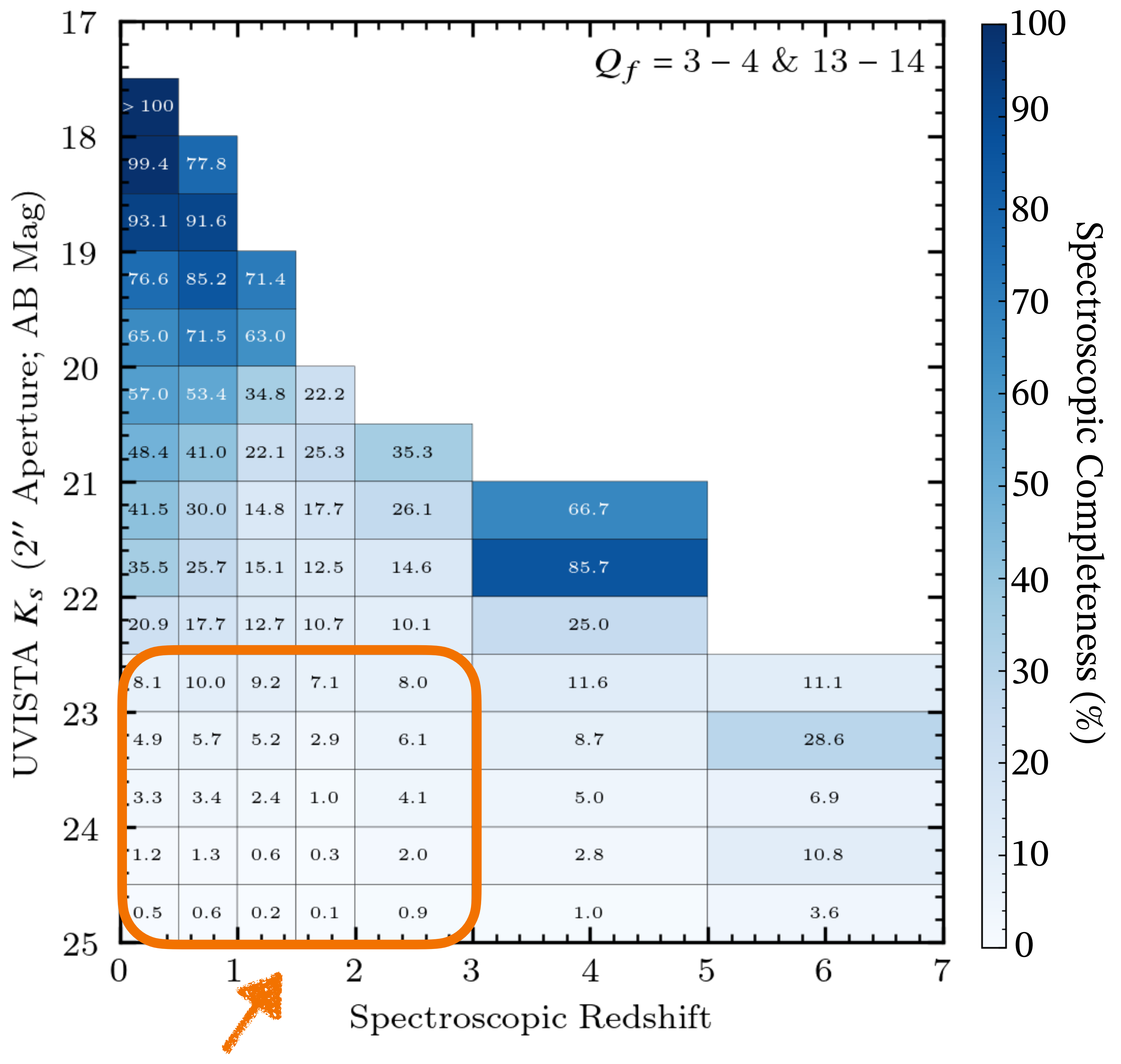
Creating a Representative Spec-z Training Set

- **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!
- Next step: re-weighting and interpolating spec-z datasets ➔ e.g., as input for SOMPZ



Finian Ashmead et al. (in prep.) ➔ check out his poster

Need Deep 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

- Want **spec-z's down to $H_{AB} \sim 24.5$** (depth of weak lensing sample) with representative colors
- Requested **50 dark nights**
- 5-15 pointings w/ **60-20 hour exposure** times for **10k-30k objects**
- **COSMOS and XMM-LSS** fields (HLIS and LSST equatorial deep fields)

DESI Deep Survey

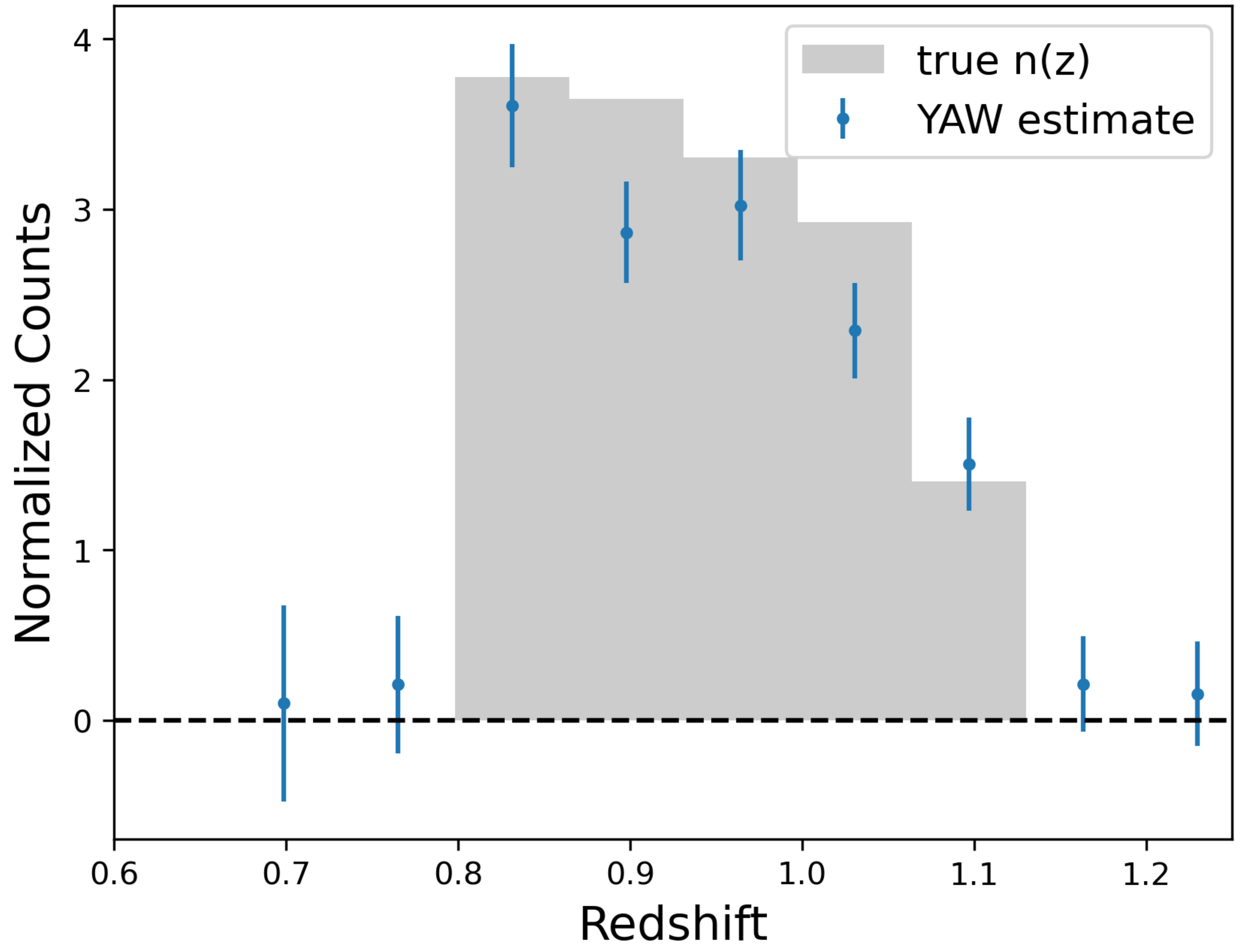
- complements SuPR Deep Survey at $z < 1.6$ (see Biprateep Dey talk)

Clustering-z: Calibrating Redshift Distributions

- Clustering-z provides an independent cross-check on photo-z distributions

- Use cross-correlations with DESI and (hopefully) Roman HLSS grism samples

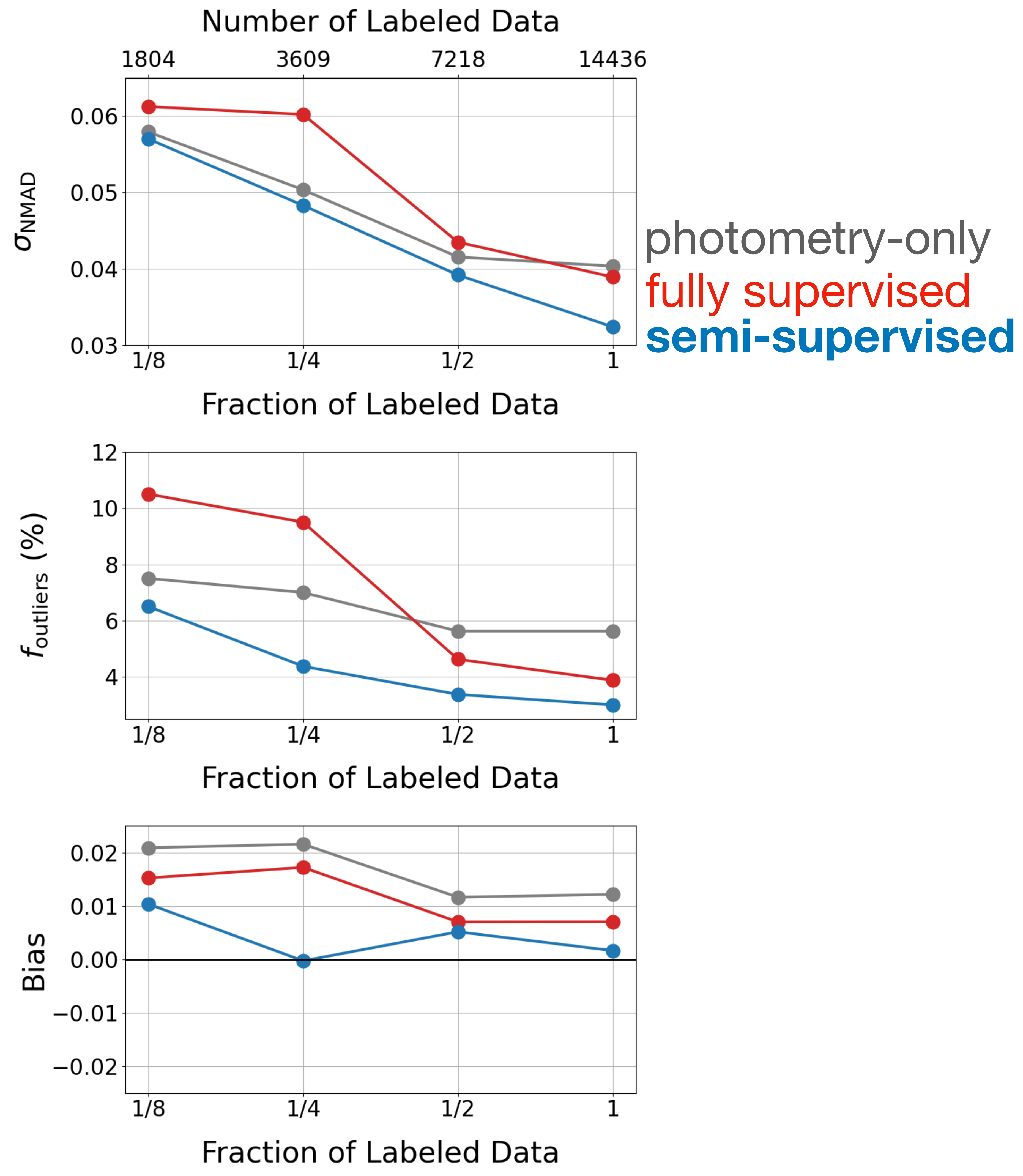
- Currently testing and validating clustering-z code (RAIL YAW) with mock DESI catalogs from simulations



Yoki Salcedo et al. (in prep.) ➔ check out his talk

Better Photo-z *Performance* w/ Deep Learning

- HST/CANDELS: key Roman precursor dataset
- Deep learning outperforms photometry-only photo-z's
- Semi-supervised training (inc. colors and spec-z's) beats fully-supervised and fine-tuned foundation models
  - **Key is low-dimensional space with continuous trends in redshift and sSFR/ color (similar to UMAP insights)**.
- WHY? Deep learning more optimally weights redshift information in pixels than photometry (**Emma Moran** et al. 2025 ➔ arxiv:2507.06299)



Ashod Khederlarian et al. (in prep.) ➔ check out his talk

Summary

- Photo-z calibration crucial for weak lensing and 3x2 pt analyses, but must mitigate spectroscopic incompleteness
- Photo-z forecasts influential for HLIS survey design
- UMAP to optimally leverage spec-z datasets (**Finian Ashmead poster**)
- Working to obtain new spec-z training sets:
  - Subaru-PFS/Roman (SuPR) Deep Survey (**Jeff Newman poster**)
  - DESI-Deep Survey (**Biprateep Dey talk**)
- Testing/validating clustering-z code, which will provide an independent cross-check on redshift distributions (**Yoki Salcedo talk**)
- Deep learning improves individual object photo-z's for HST/CANDELS, a key Roman precursor dataset (**Ashod Khederlarian talk**)

Bonus Slides

Roman-only photo-z's will be unreliable: need LSST photometry!

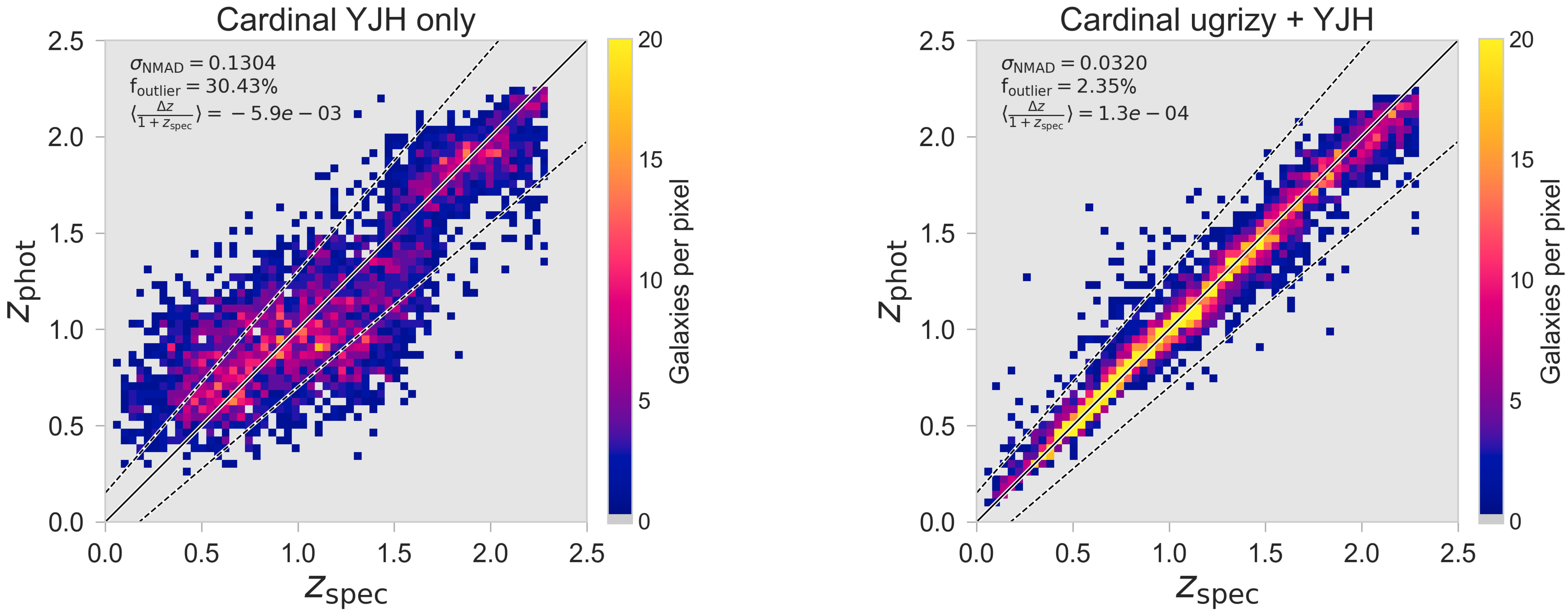


Photo-z Performance: YJH vs. JH vs. H

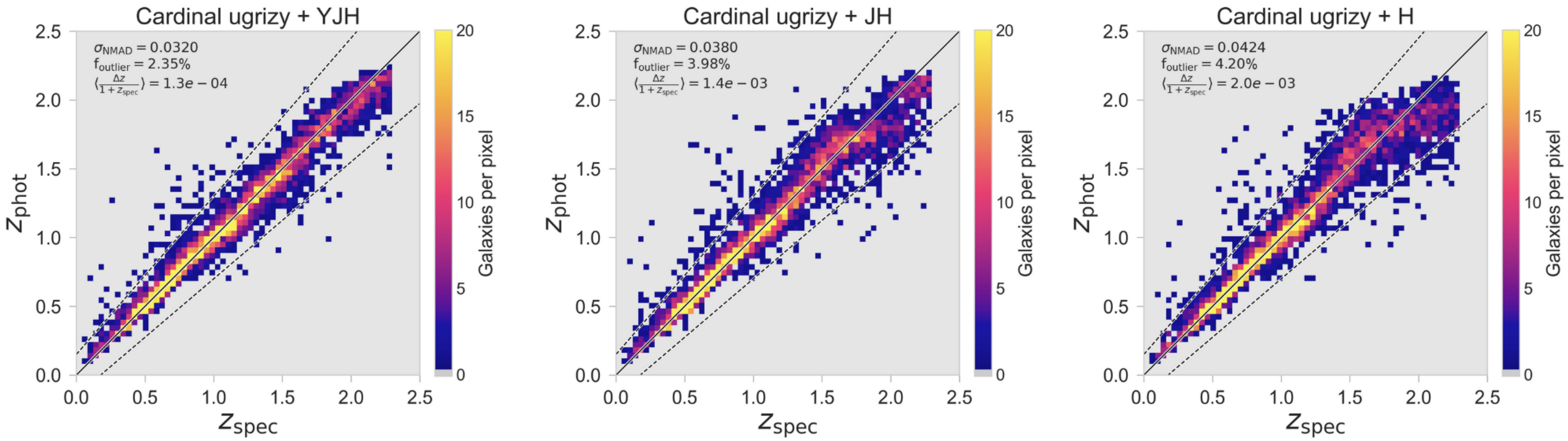
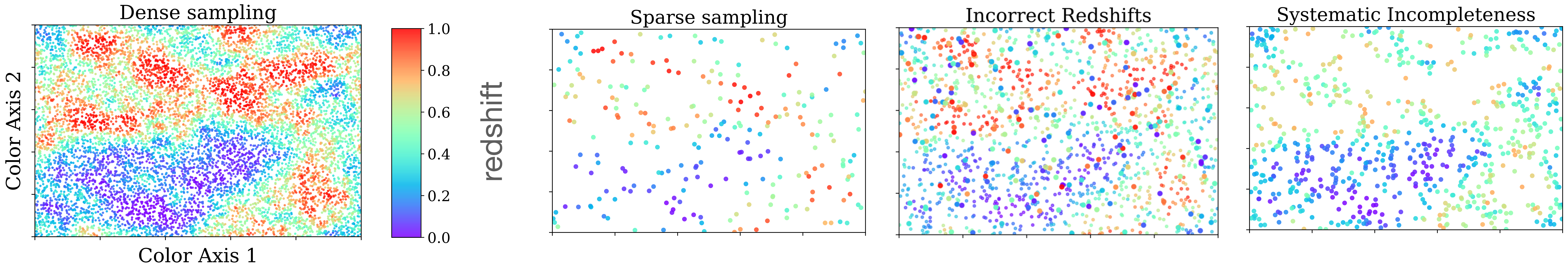
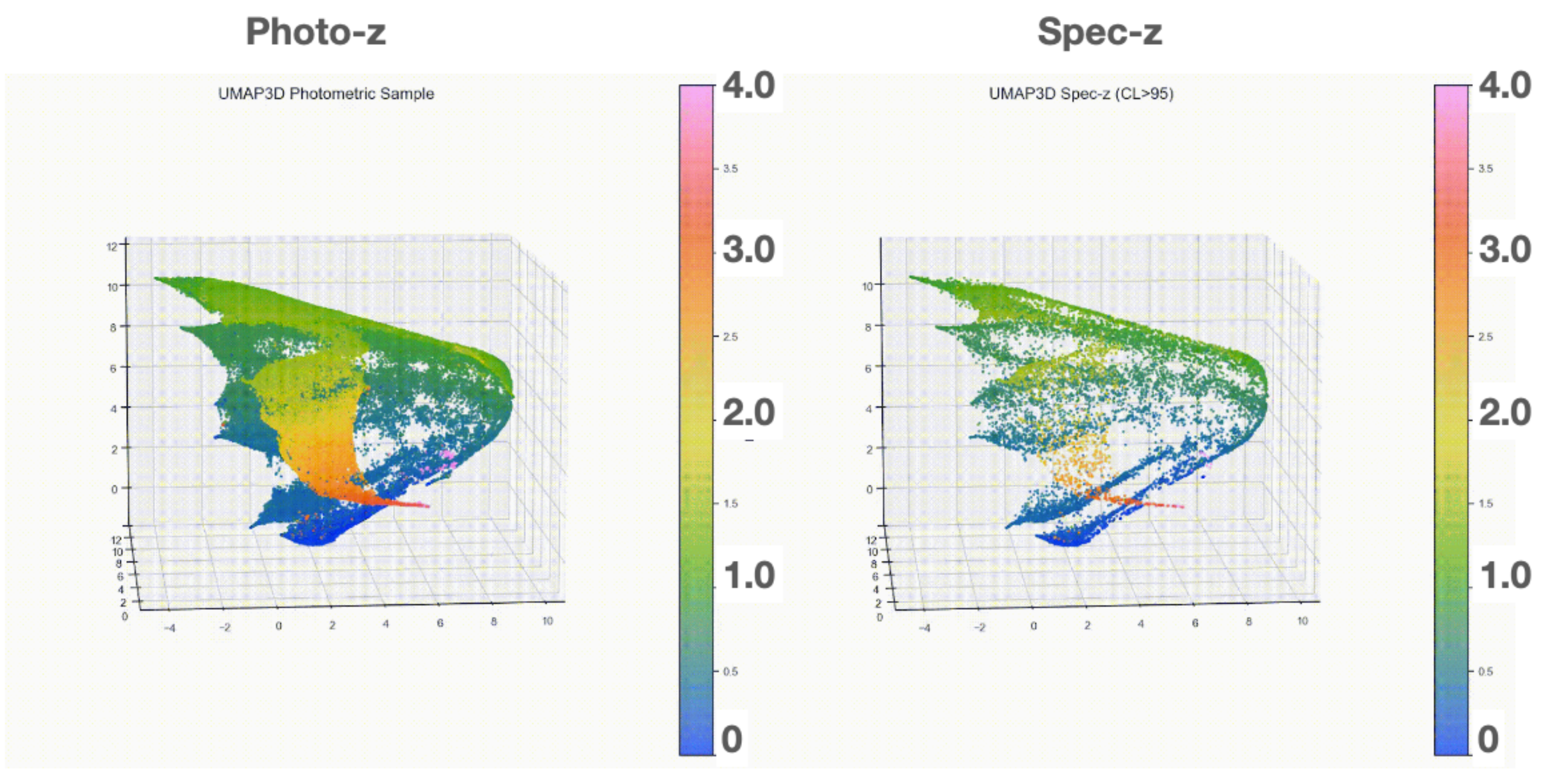


Photo-z Calibration Challenges



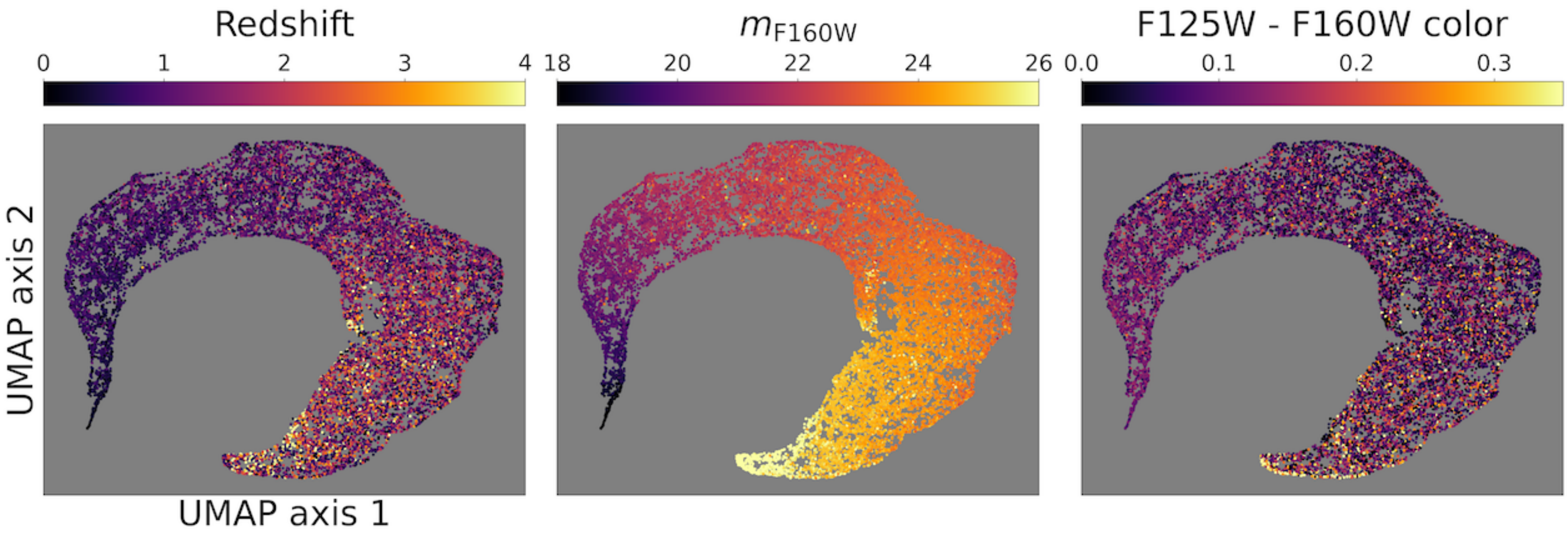
Newman & Gruen (2022)



credit: Finian Ashmead

- Compared to photometric objects in color-redshift space, existing spec-z training sets suffer from
  - **sparse sampling → interpolation**
  - **incorrect spec-z's → robust regression**
  - **systematic incompleteness → rebalance training set to match photometric objects**

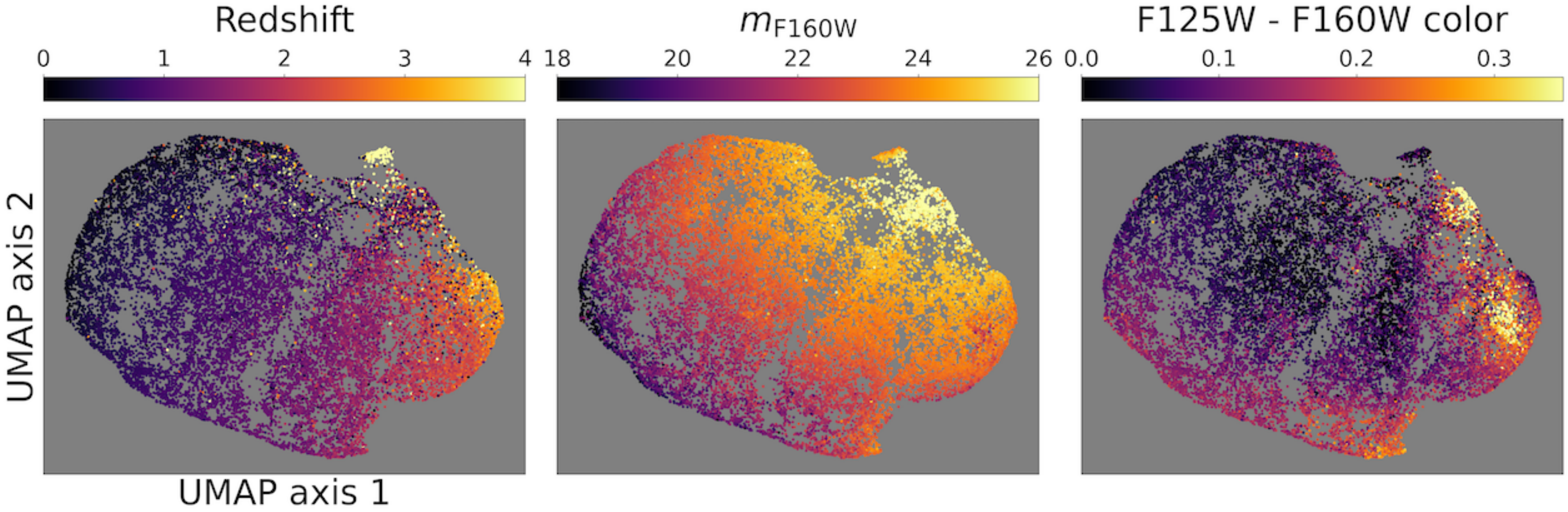
Self-Supervised Latent Space



- *discontinuous* redshift and color trends

Ashod Khederlarian et al. (in prep.) ➔ check out his talk

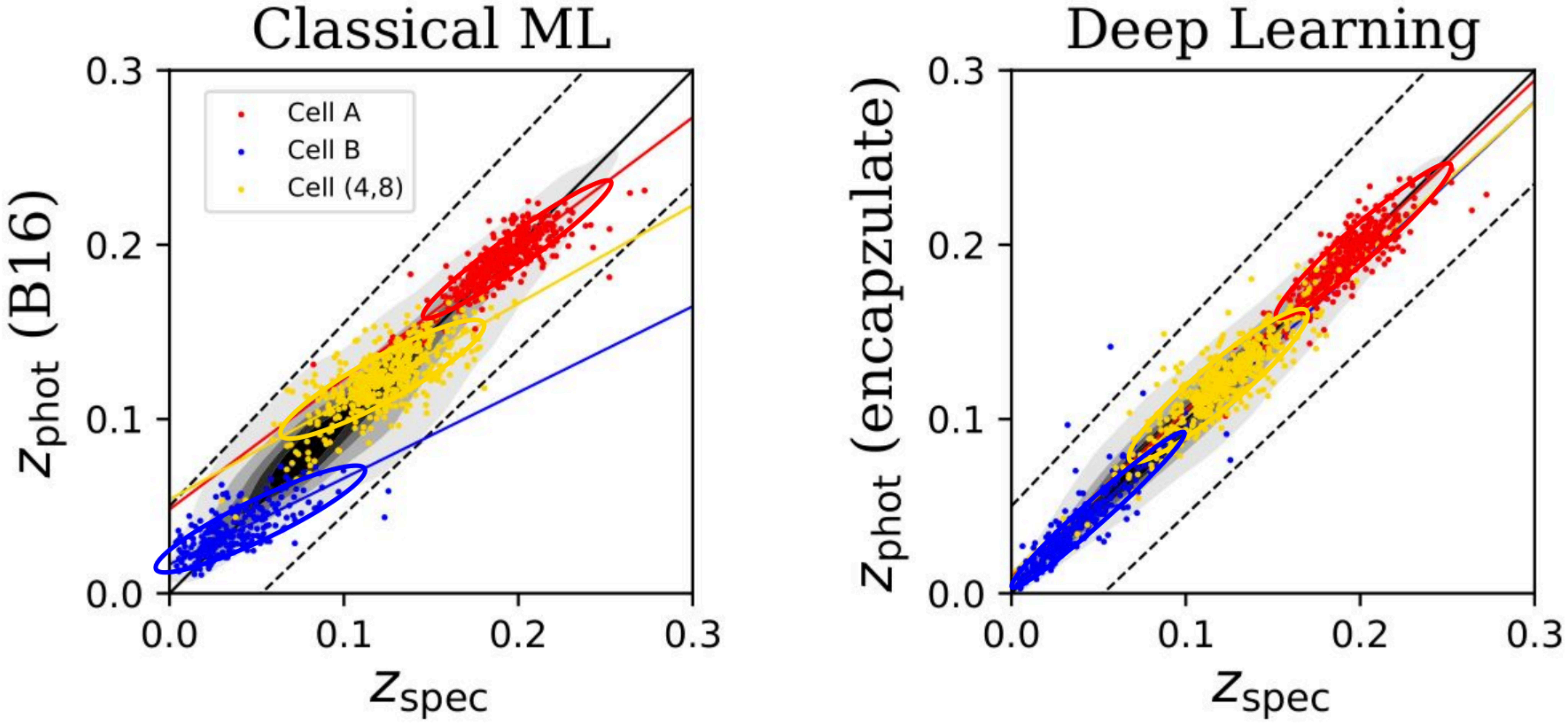
Semi-Supervised Latent Space



- *continuous* redshift and color trends

Ashod Khederlarian et al. (in prep.) ➔ check out his talk

Deep Learning: reduced color-dependent attenuation bias



- **Emma Moran** et al. 2025 ➔ arxiv:2507.06299