

Galaxy Evolution meets Cosmological Inference

Hiranya V. Peiris



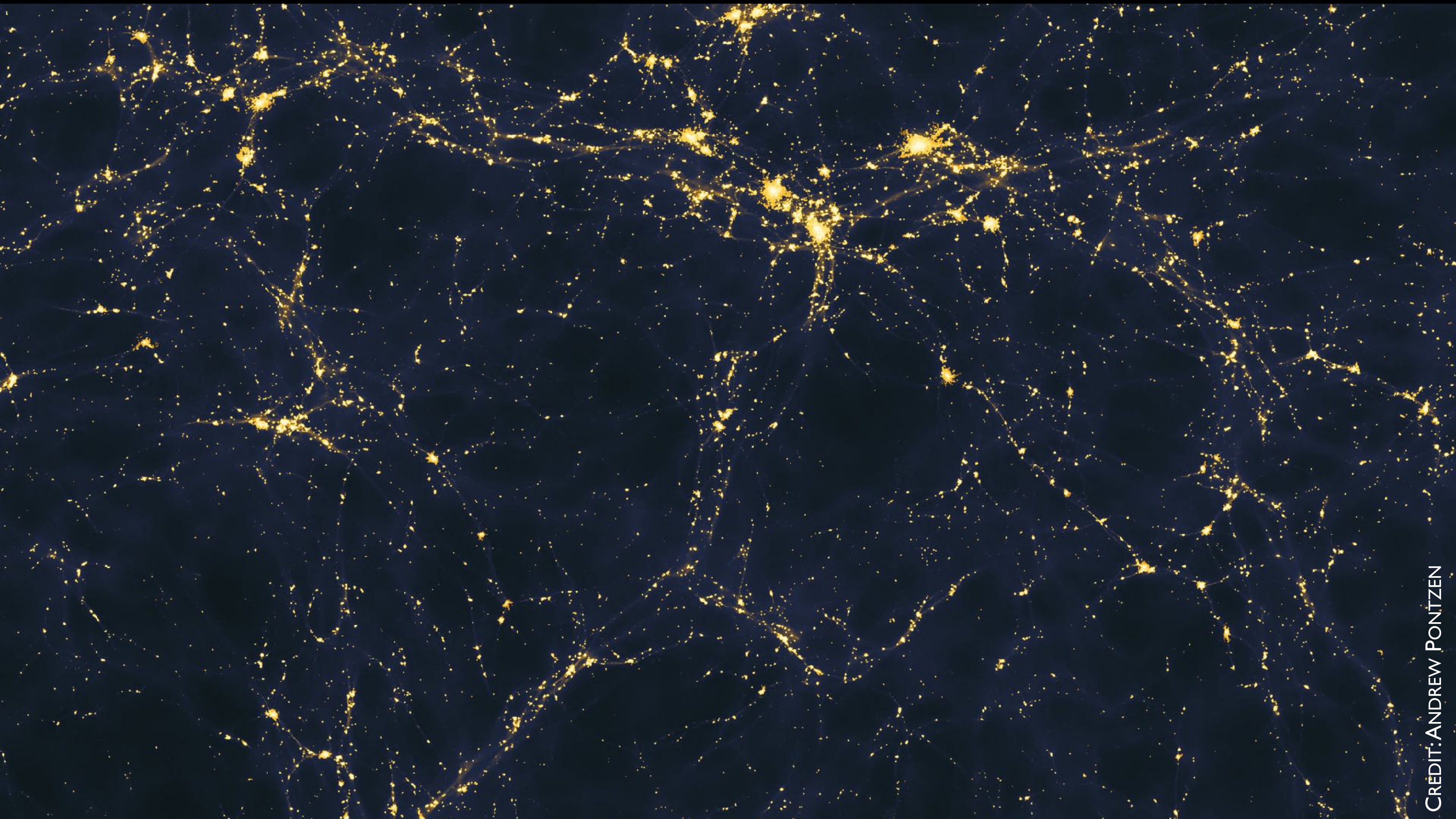
European Research Council
Established by the European Commission



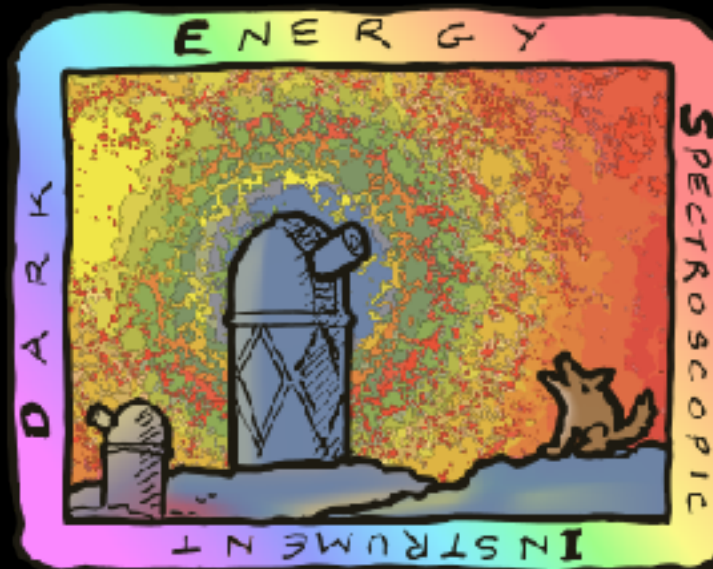
UNIVERSITY OF
CAMBRIDGE



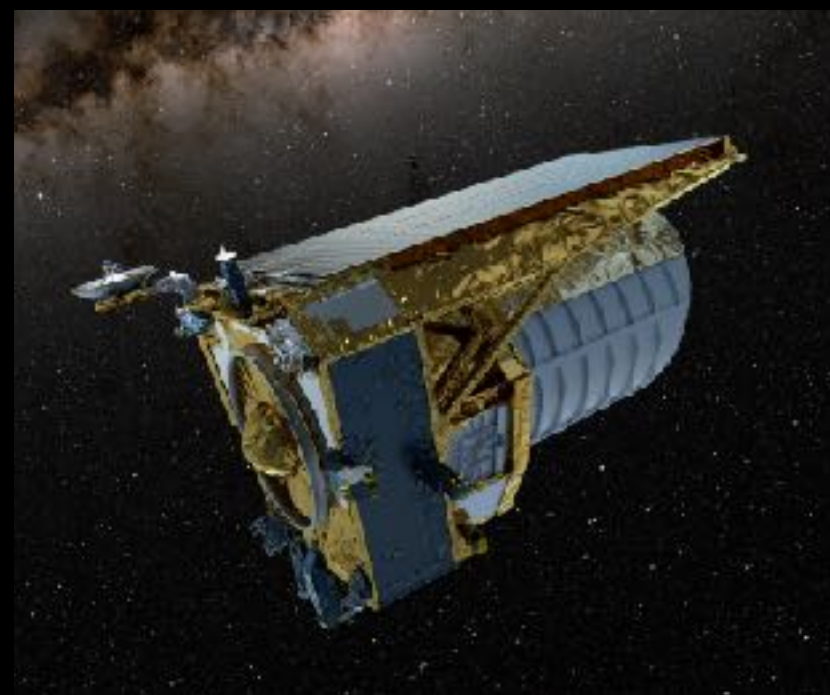
UK Research
and Innovation



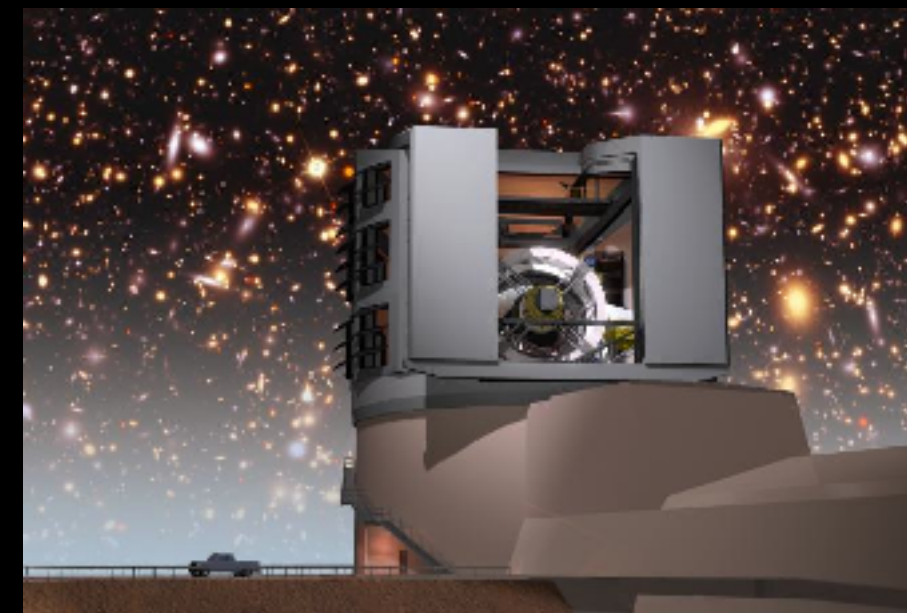
New eyes on the Universe



Started 2021



Started 2024

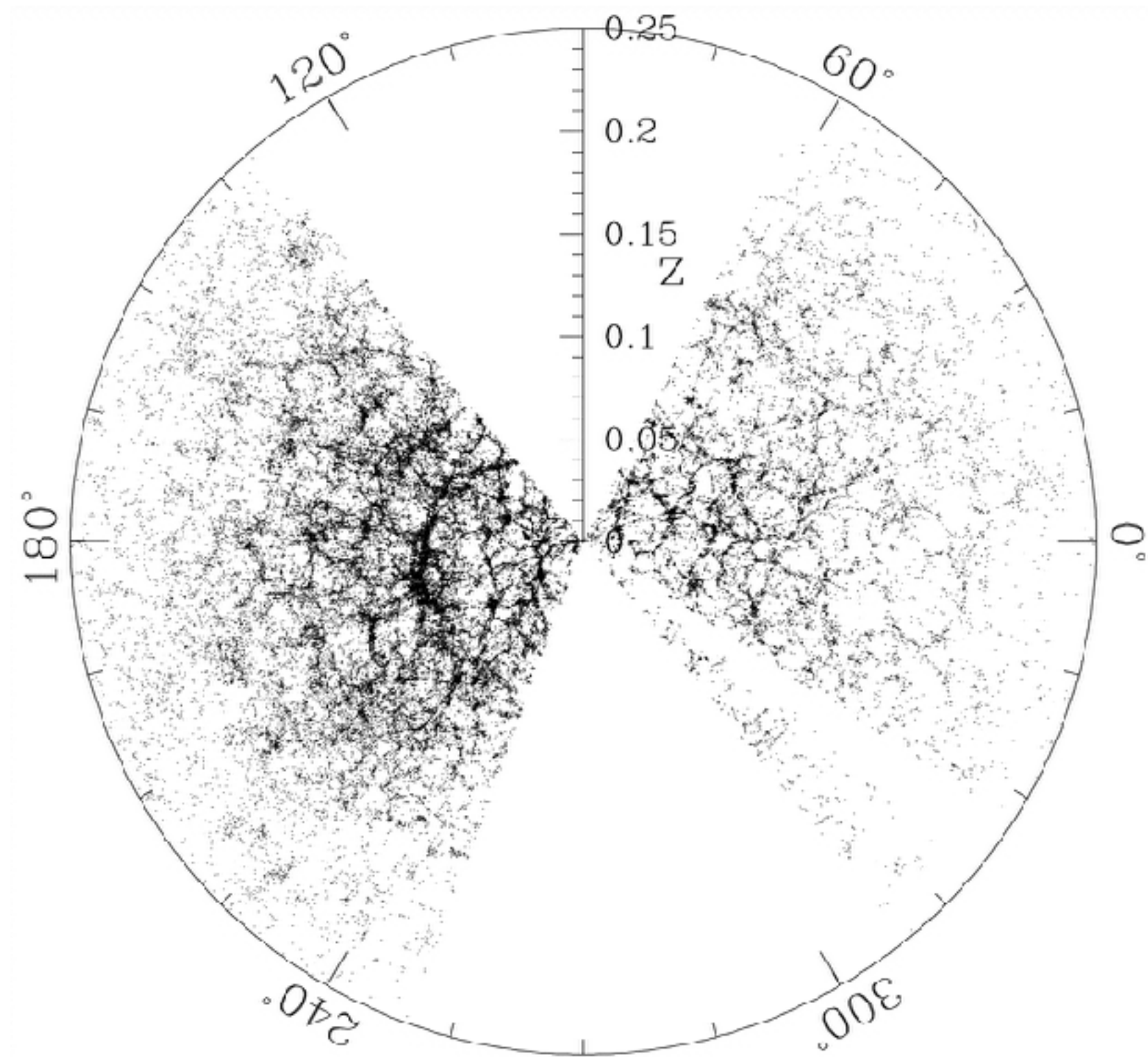


Started 2025

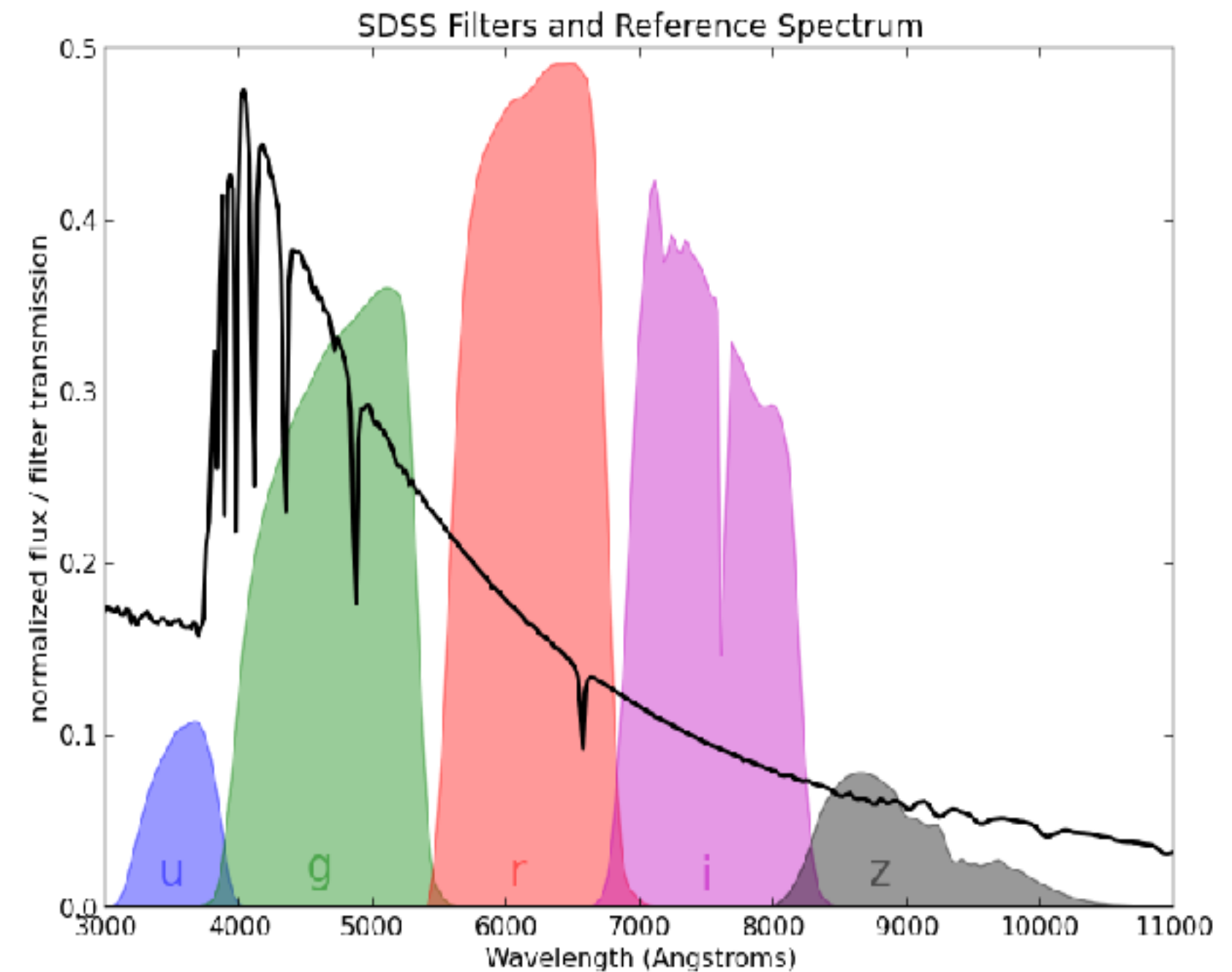


Launching 2026

Large scale structure cosmology



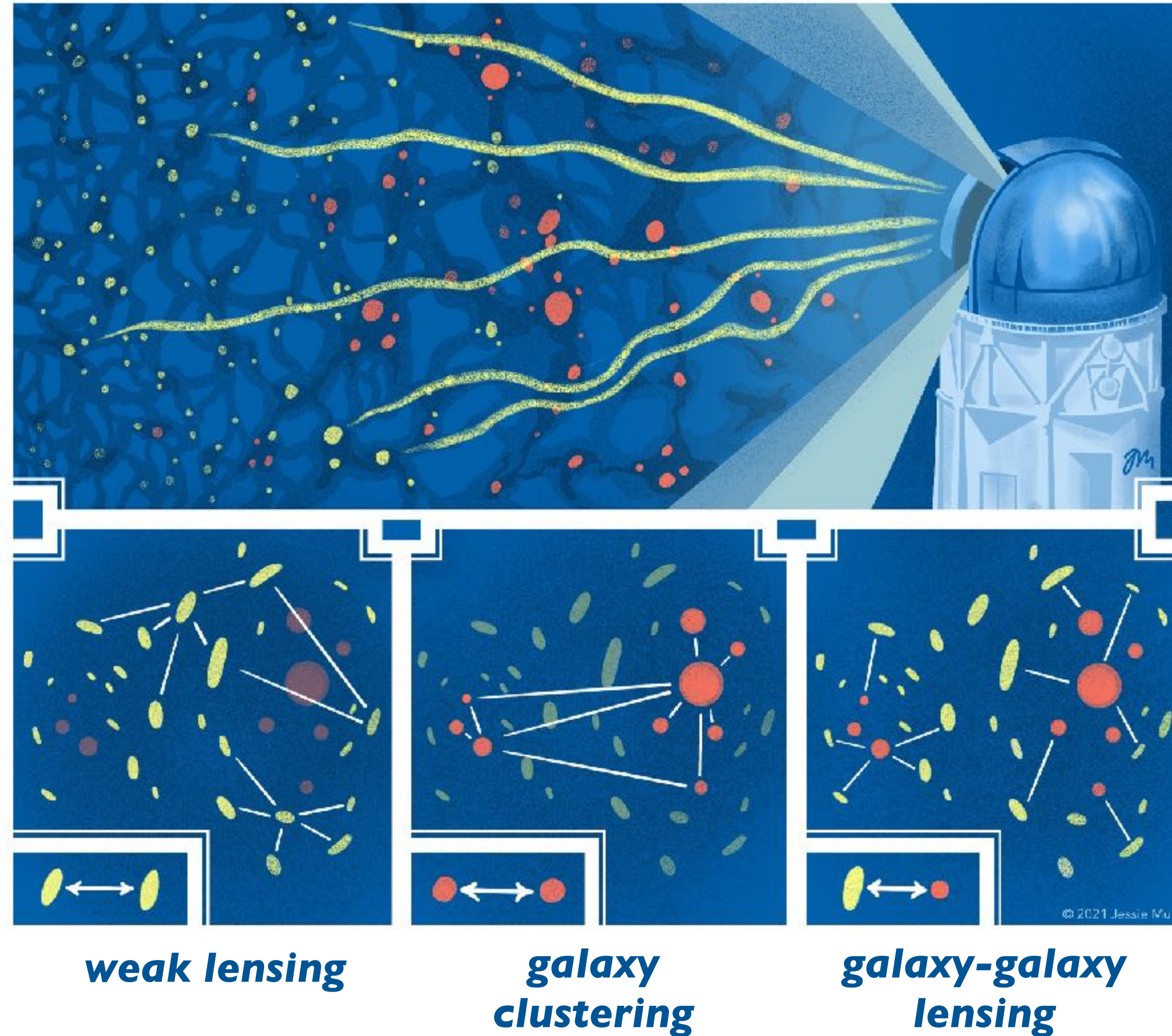
Spectra: precise redshift,
precise 3D positions of galaxies



Photometry: precise 2D positions, requires
redshift estimation from broadband fluxes

Map matter distribution in the Universe using galaxies as tracers

Large scale structure cosmology



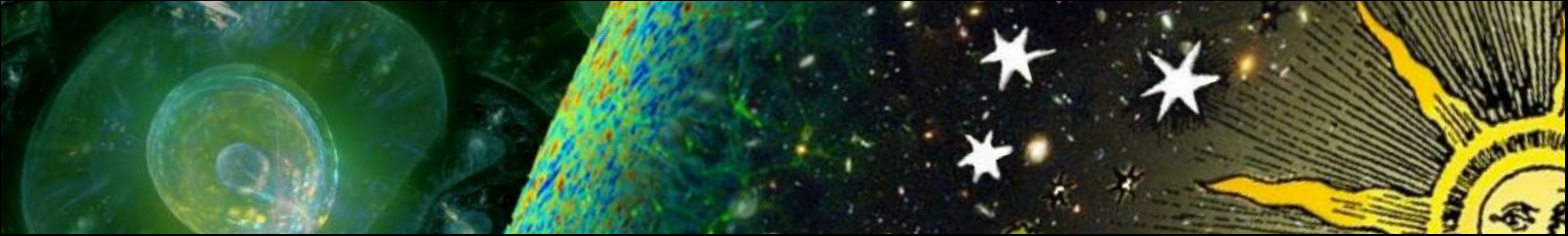
Information from galaxy positions and shapes help us to map dark matter across cosmic time



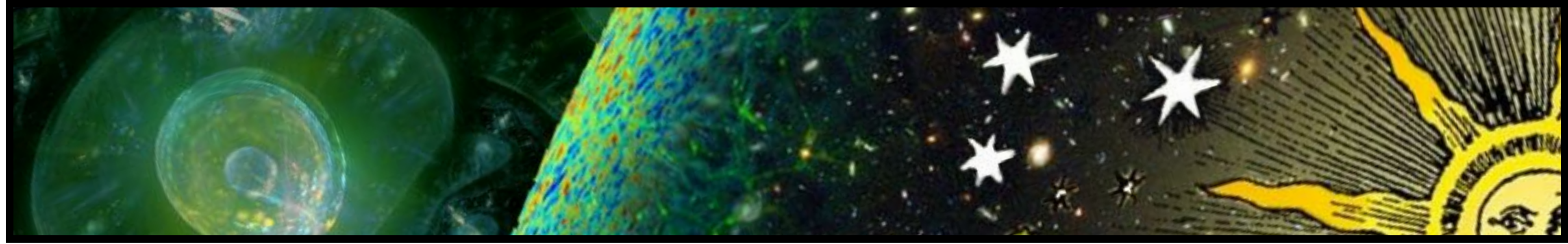
Highest resolution dark matter map ever made

COSMOS-Web / Scognamiglio et al (2026)





***To decode the cosmos,
we need to understand galaxies***



Simulation-based calibration of high-dimensional generative models



Stephen Thorp



Sinan Deger



Anik Halder



Madalina Tudorache



Benedict Van den Bussche



Joy Gong



Justin Alsing



Boris Leistedt



Gurjeet Jagwani



Daniel Mortlock



Joel Leja



Hiranya Peiris



erc

European Research Council
Established by the European Commission

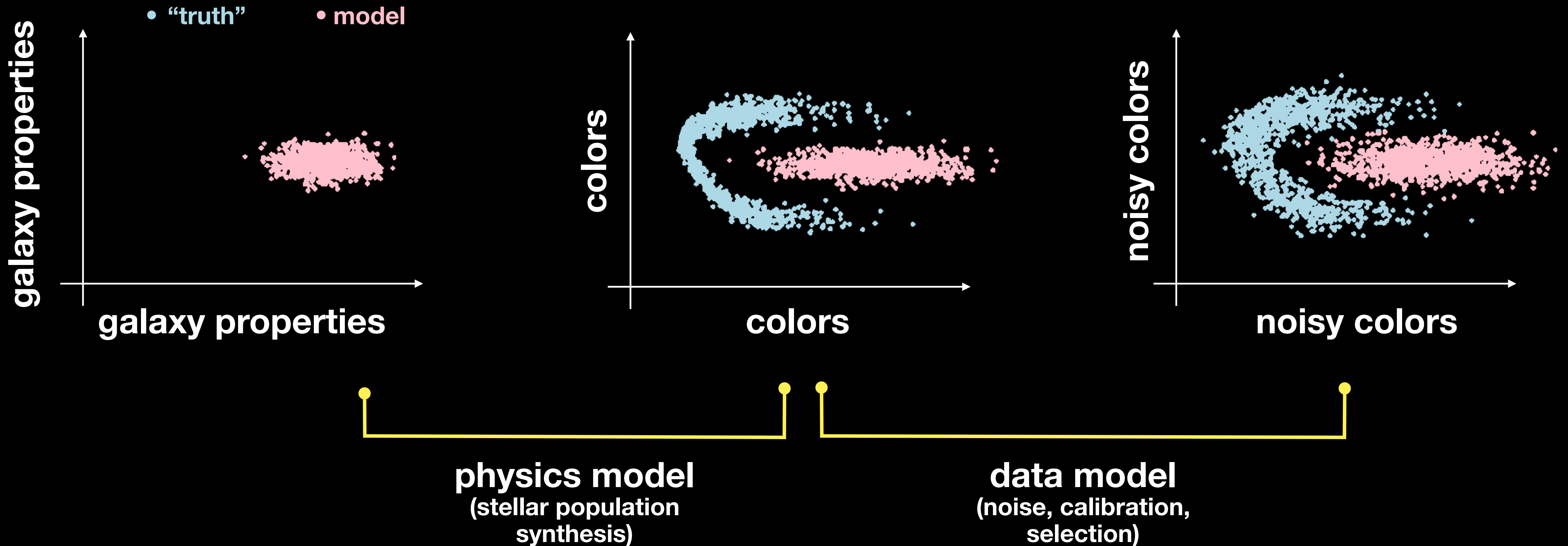
AI for data-driven discovery

What is the distribution of galaxy properties in the Universe?
This is governed by **galaxy evolution**.

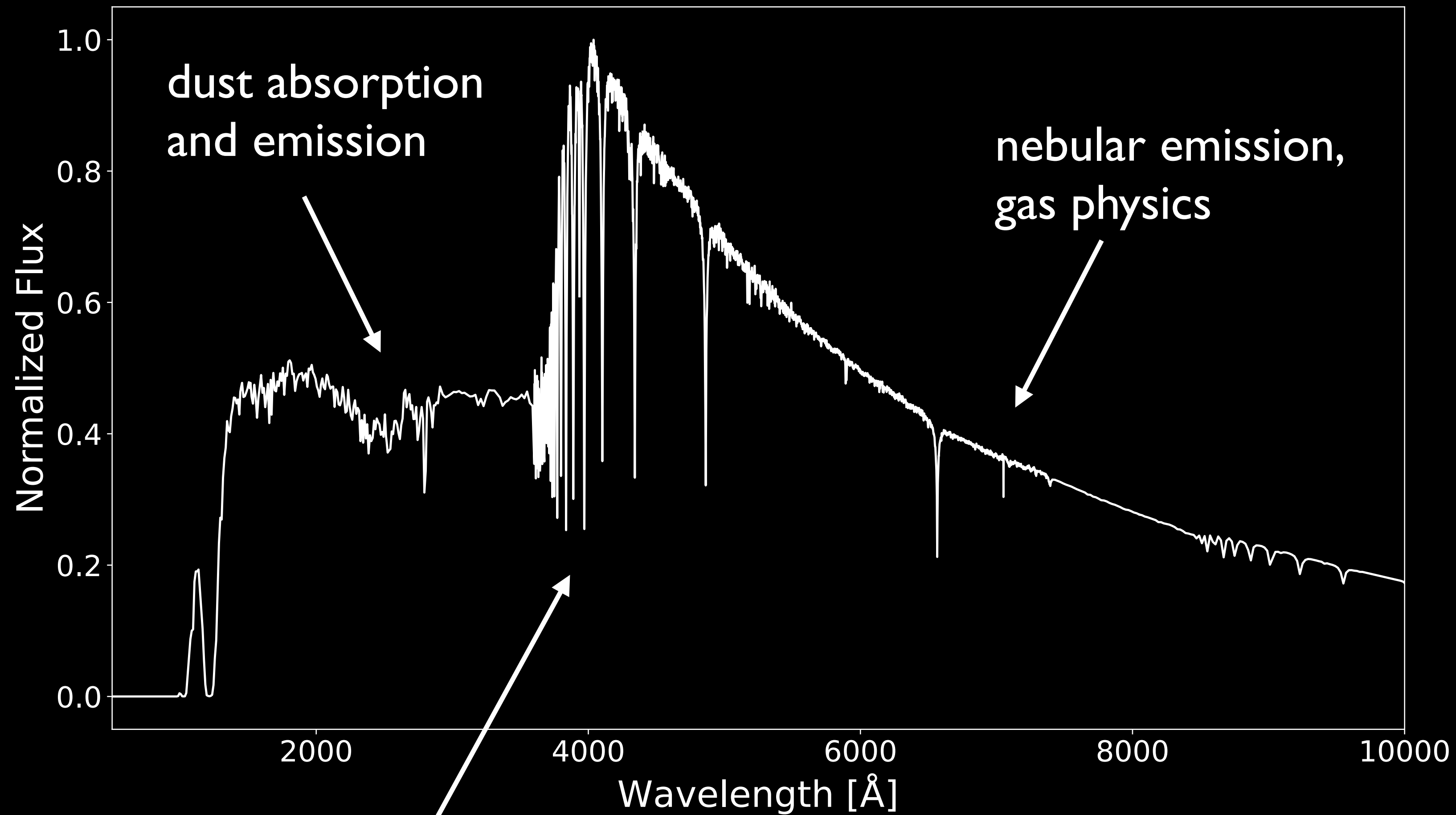


AI for data-driven discovery

Generative AI galaxy population model, coupled with rigorous physics and data models.

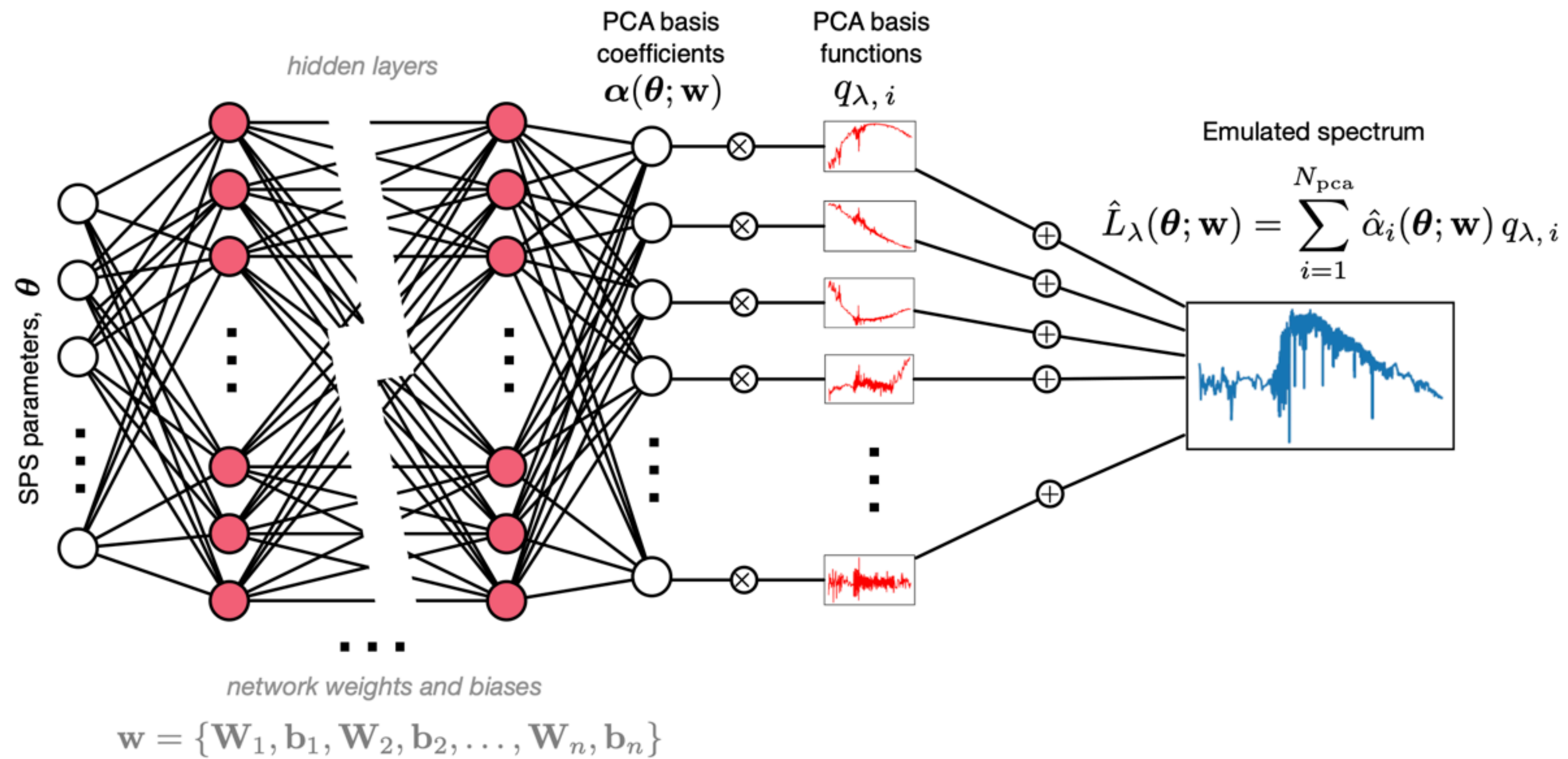


Model galaxy spectra using stellar population synthesis

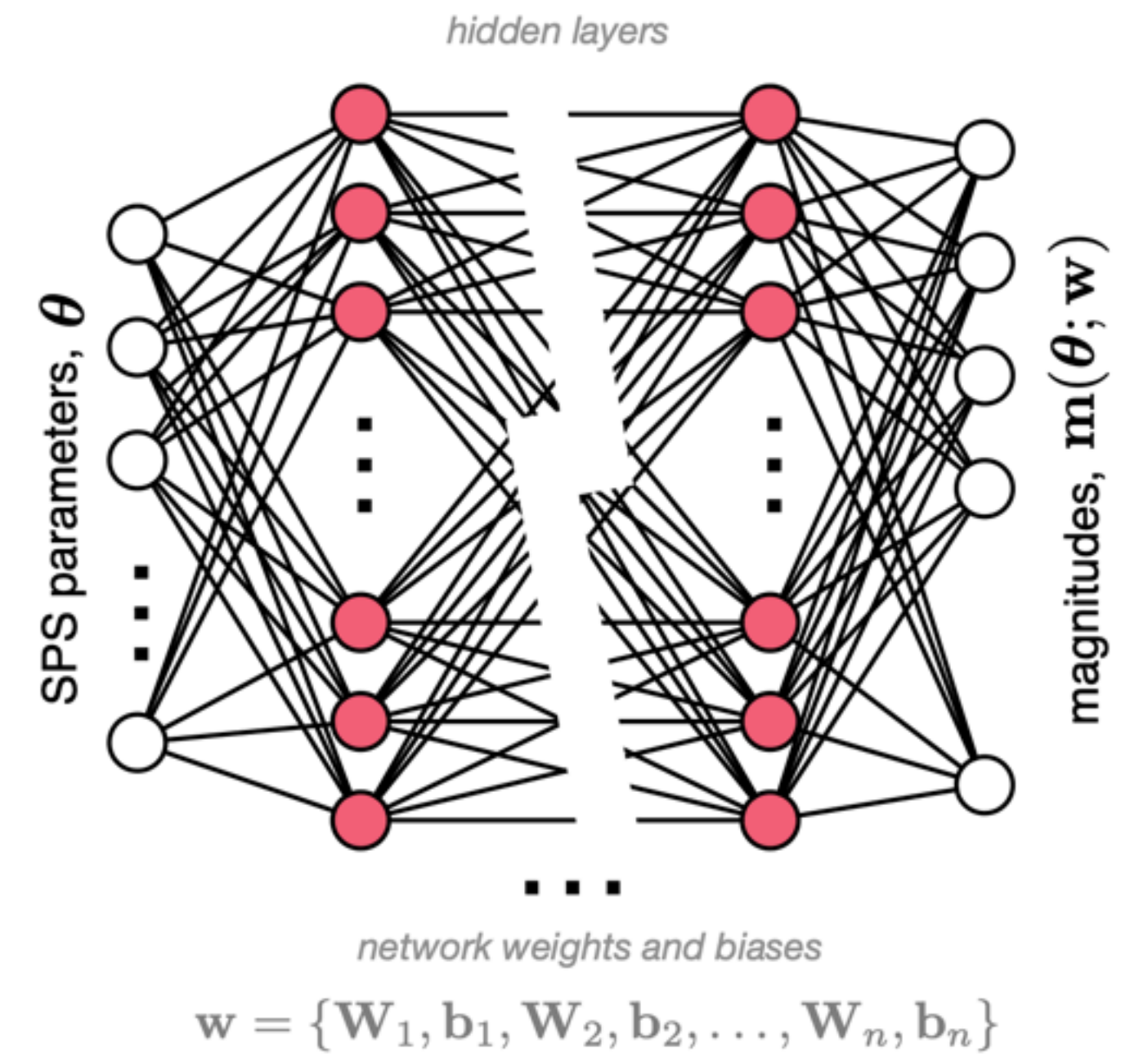


add up light from all the stars (at their ages and metallicities)

Speeding things up with neural emulators



Emulating spectra

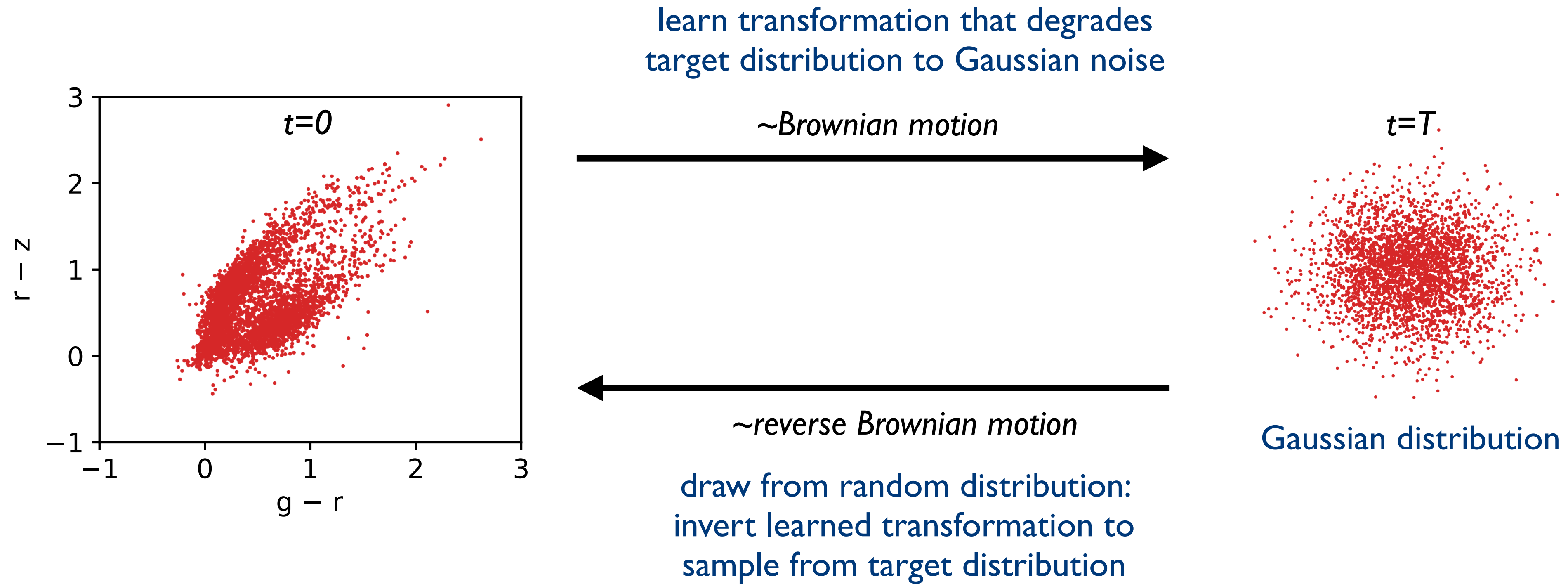


Emulating photometry

16-parameter SPS model | sub-percent accuracy | factors x 10⁴ speed-up | differentiable

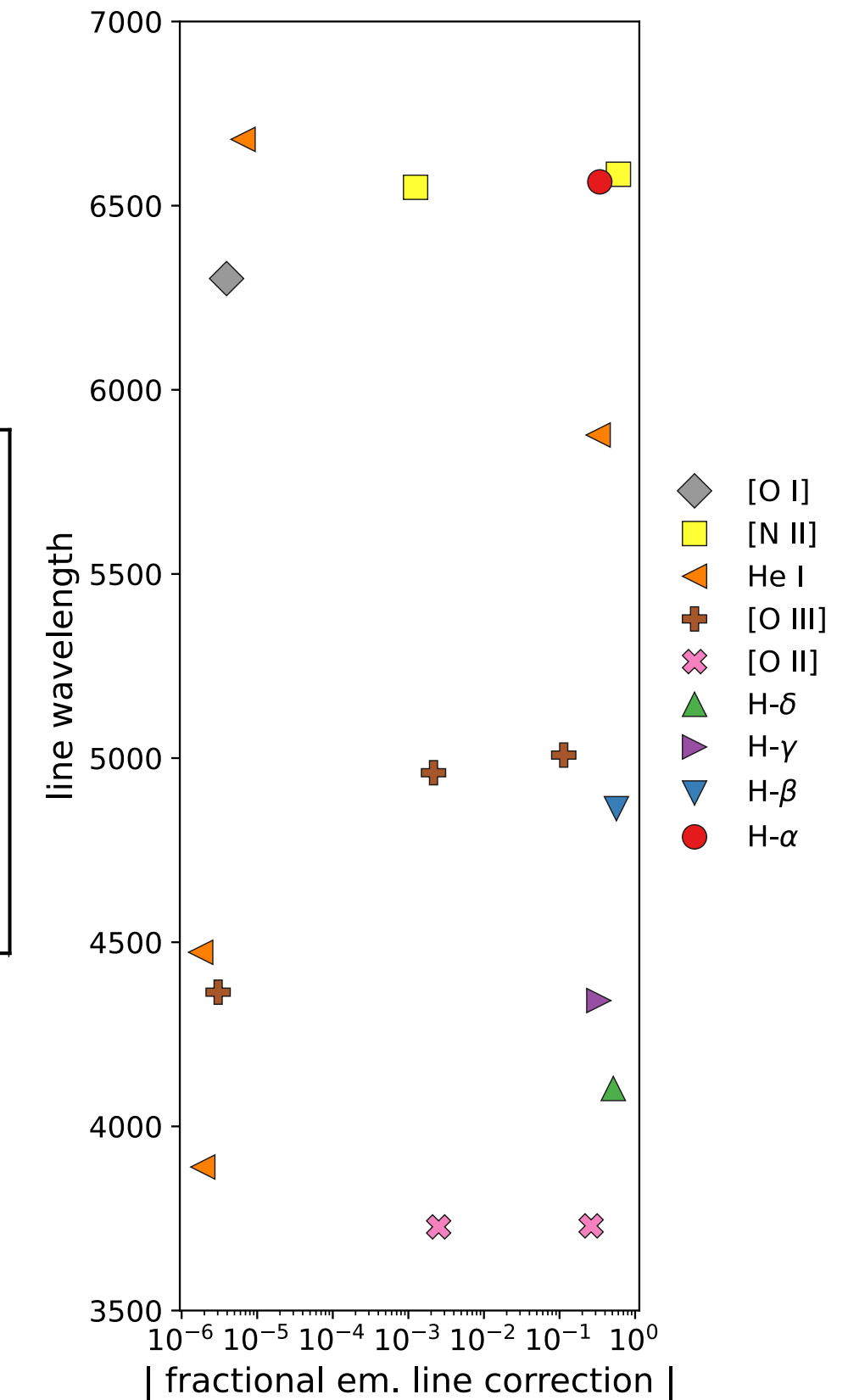
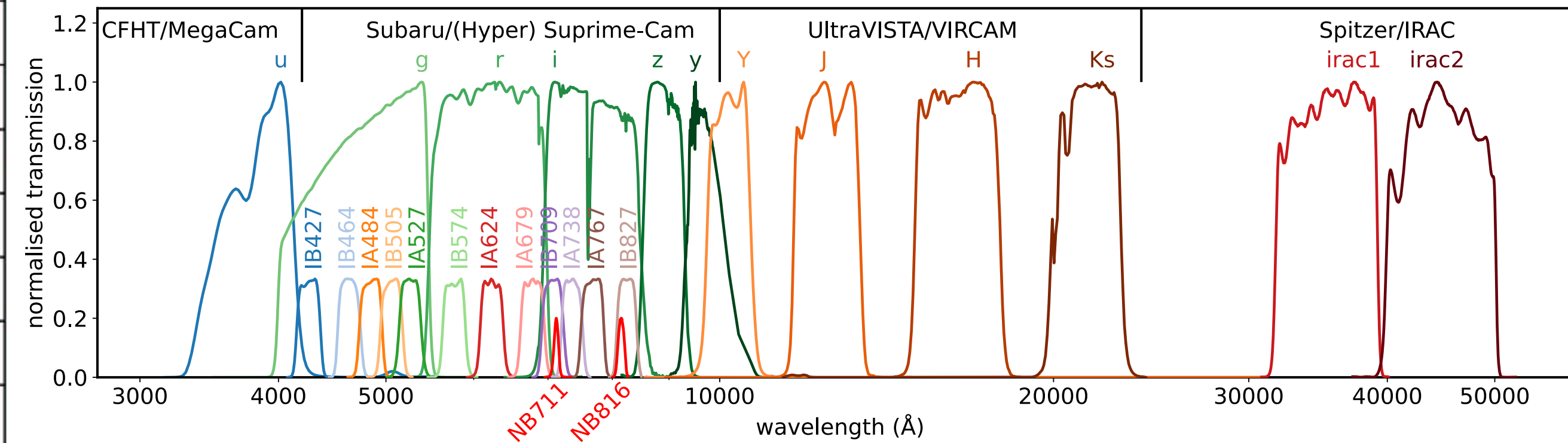
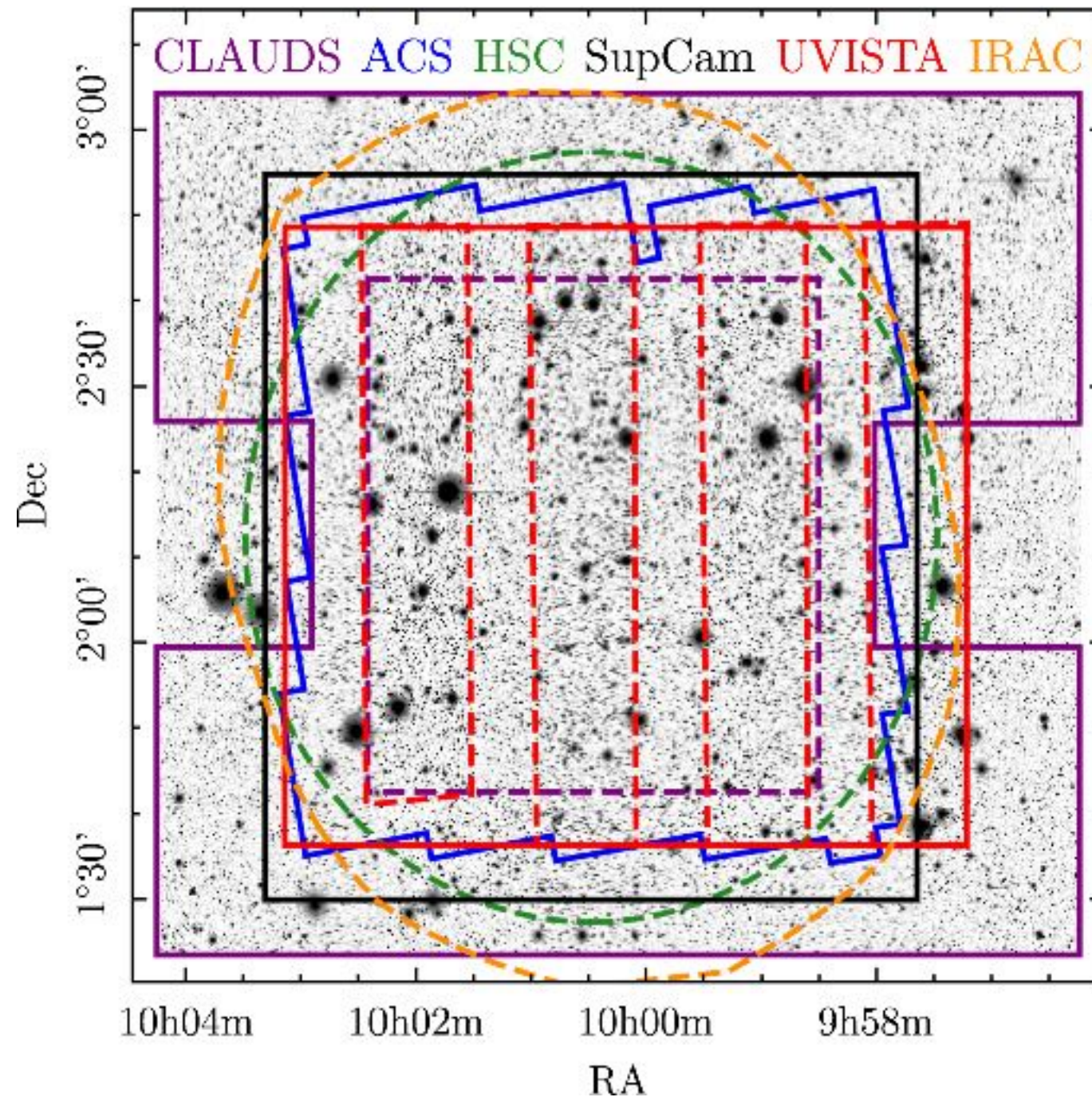
ALSING, PEIRIS, LEJA, HAHN, TOJEIRO, MORTLOCK, LEISTEDT, JOHNSON, CONROY (APJS, 2020)

Flexible neural models for distribution of galaxy properties



score-based diffusion model

pop-cosmos: galaxy population model calibrated with COSMOS2020

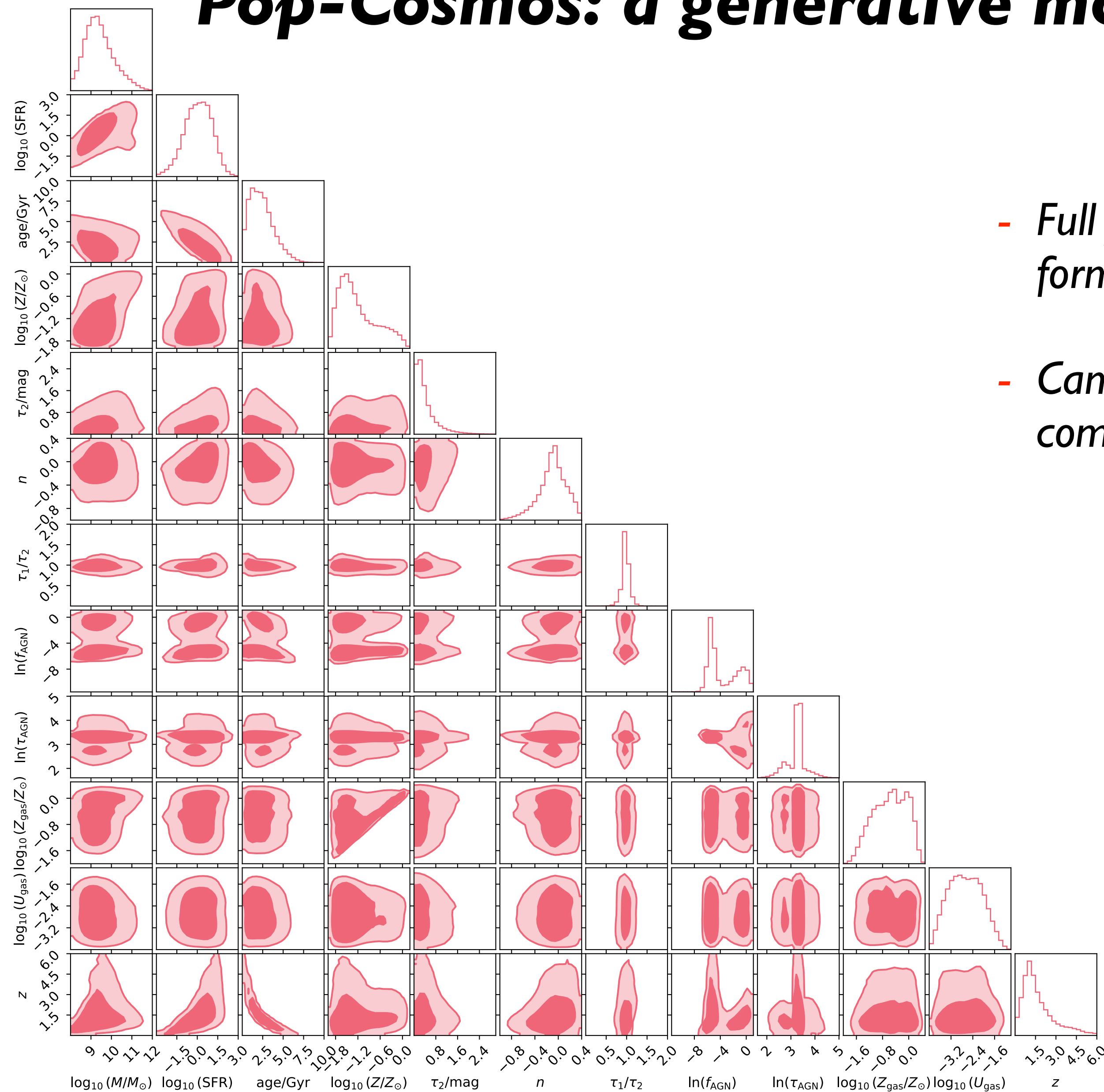


~400,000 galaxies | 26 bands near-UV to mid-IR | deep $z < 4$ | simple selection

Zero-point calibration | emission line corrections | Student-t uncertainty model

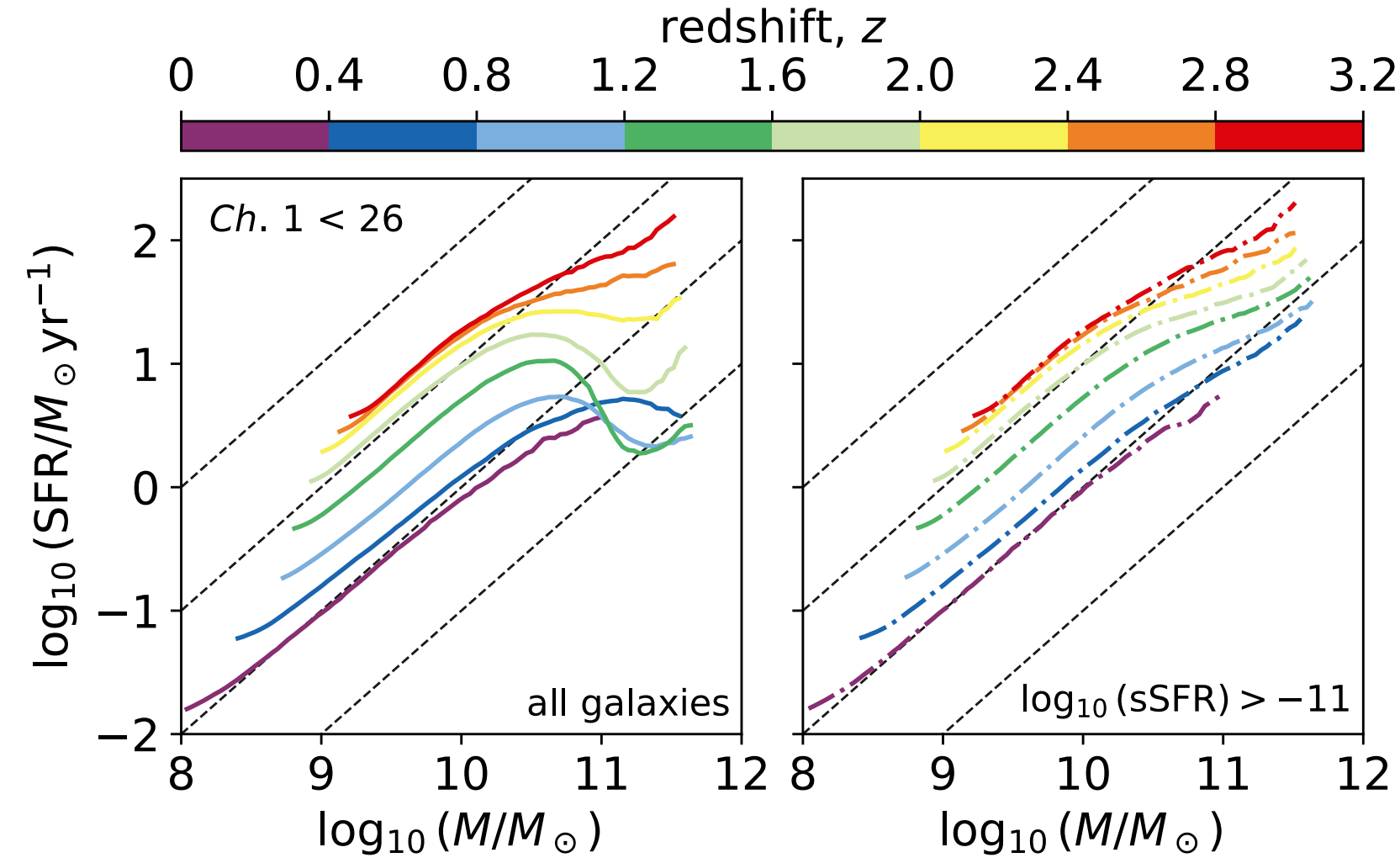
WEAVER ET AL (2021), ALSING ET AL (2022, APJS), LEISTEDT ET AL. (2022, APJS), ALSING ET AL (2024, APJS), THORP ET AL (2025)

Pop-Cosmos: a generative model for galaxy surveys

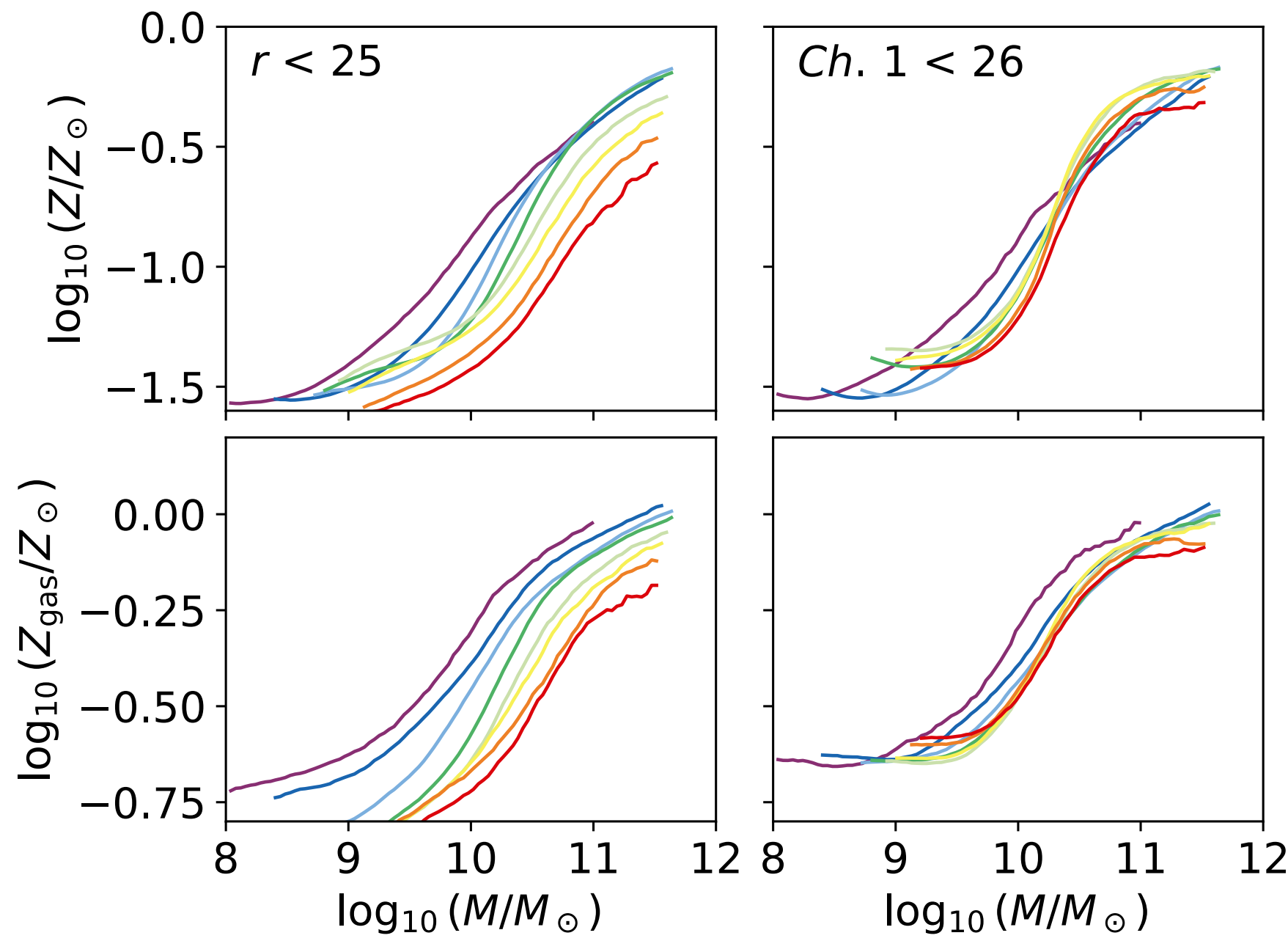


- Full joint density of galaxy properties learned in the form of a physics model (SPS+redshifts)
- Can predict properties of any catalogue of comparable / shallower depth than COSMOS2020

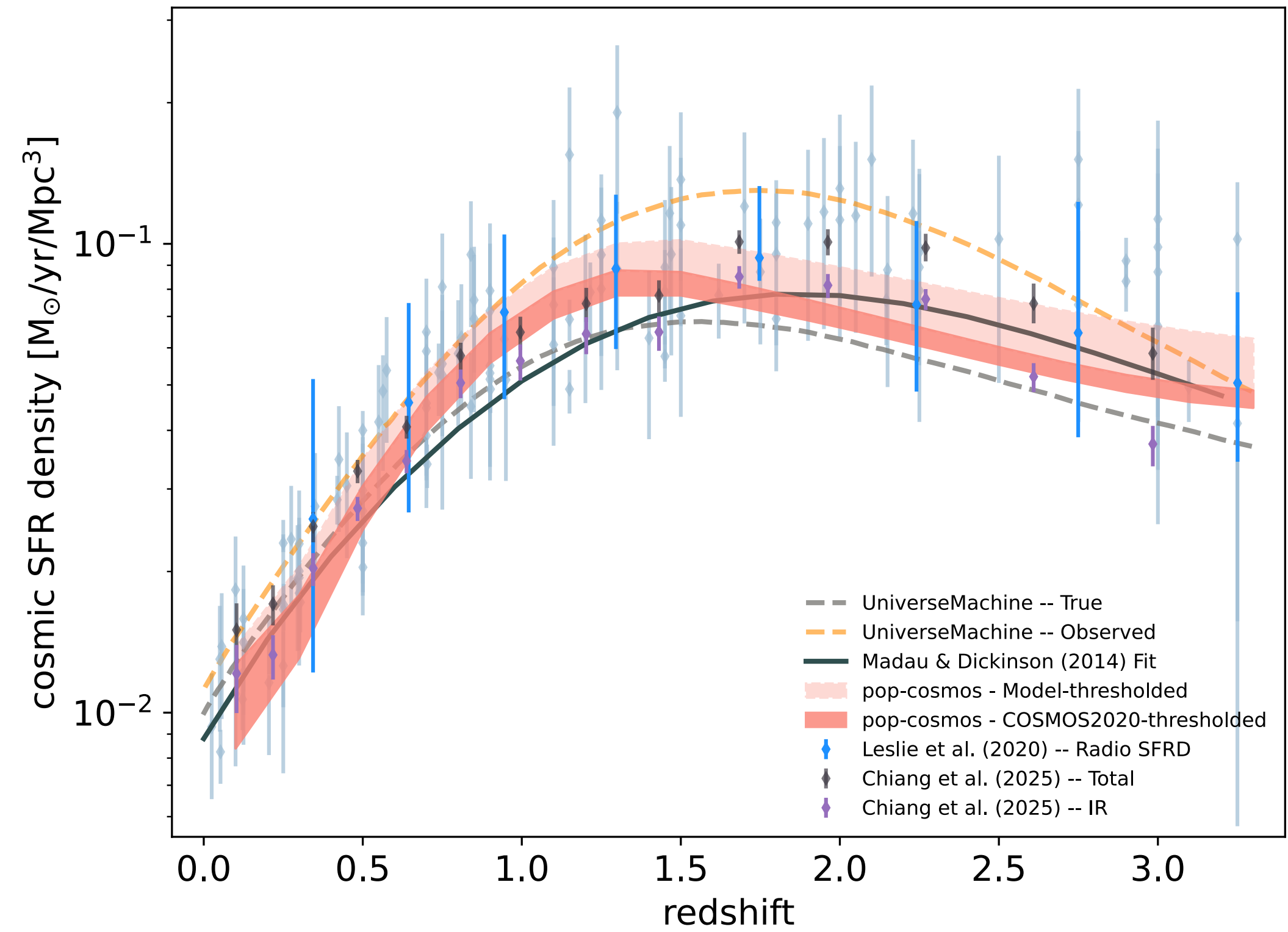
Information on full galaxy population over cosmic time



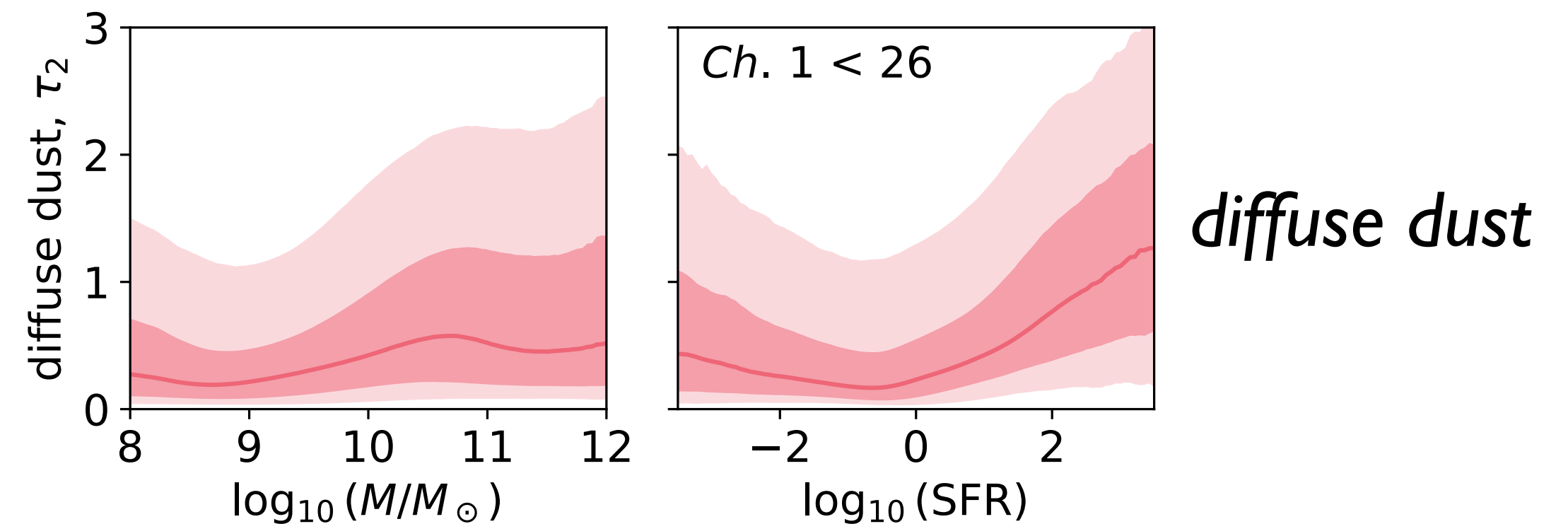
star-forming sequence



mass-metallicity relations

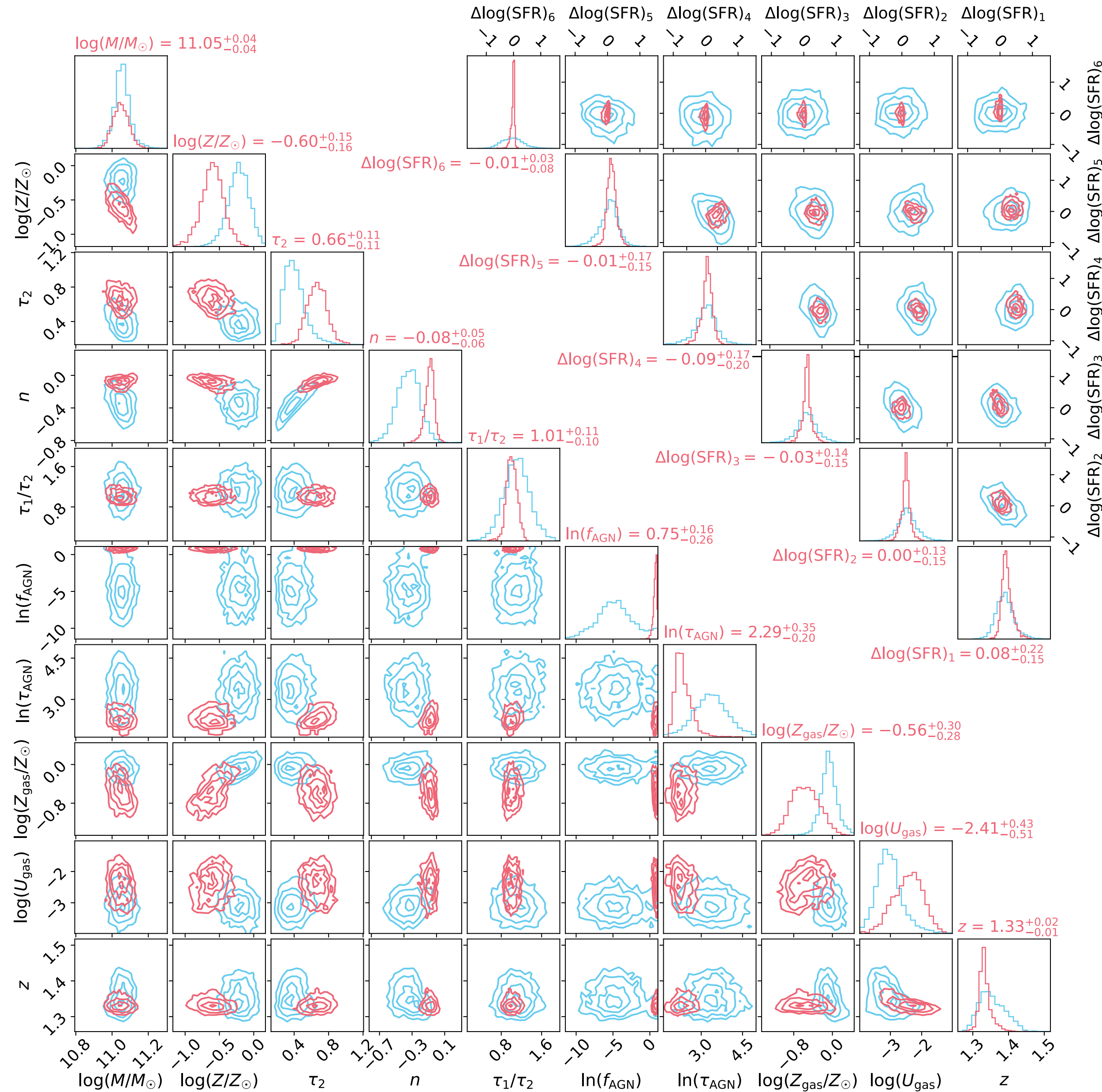


cosmic star formation rate density



diffuse dust

Full Bayesian SED fitting of large photometric catalogues



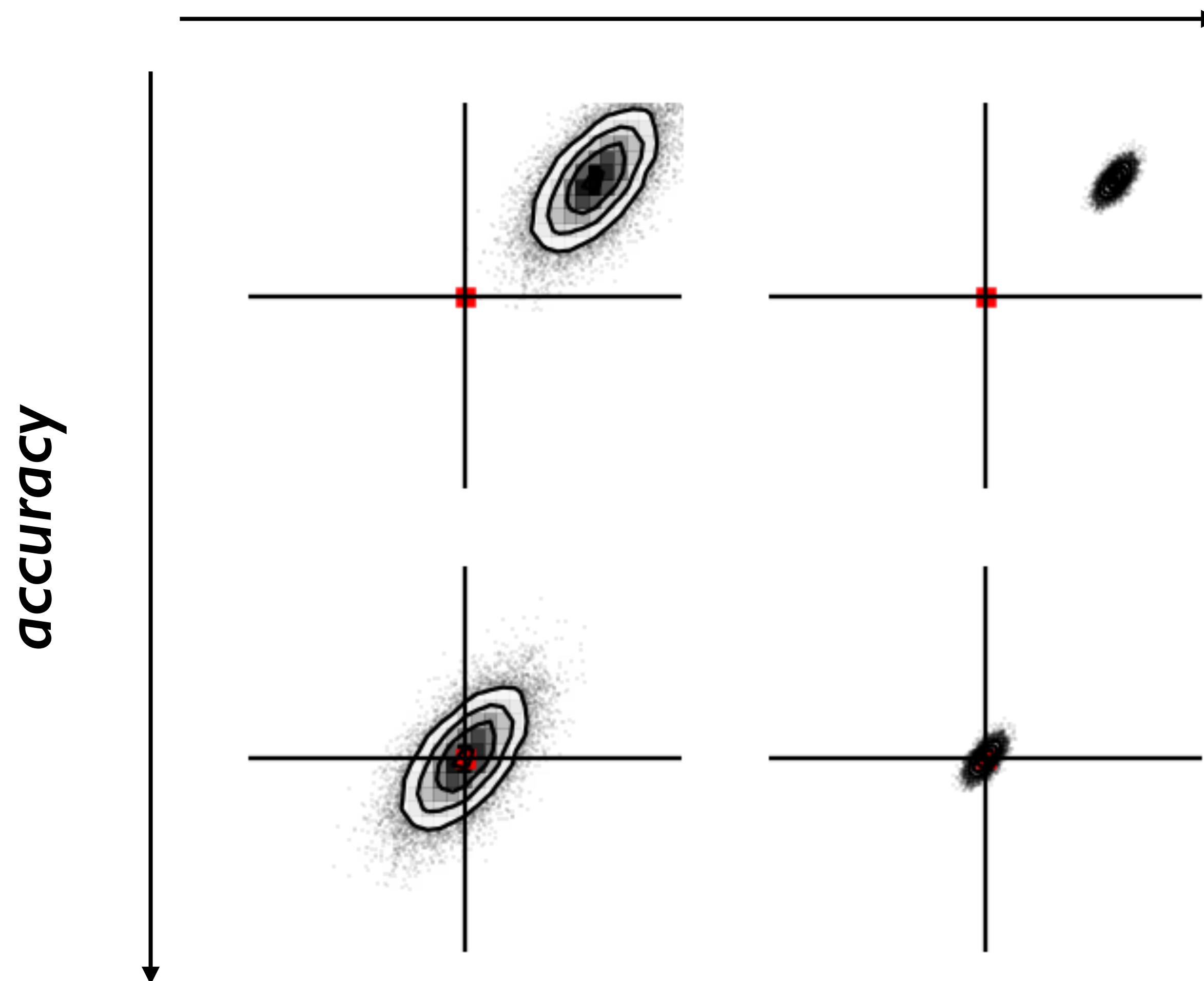
15 GPU-sec / galaxy under pop-cosmos prior

<https://zenodo.org/records/15850536>

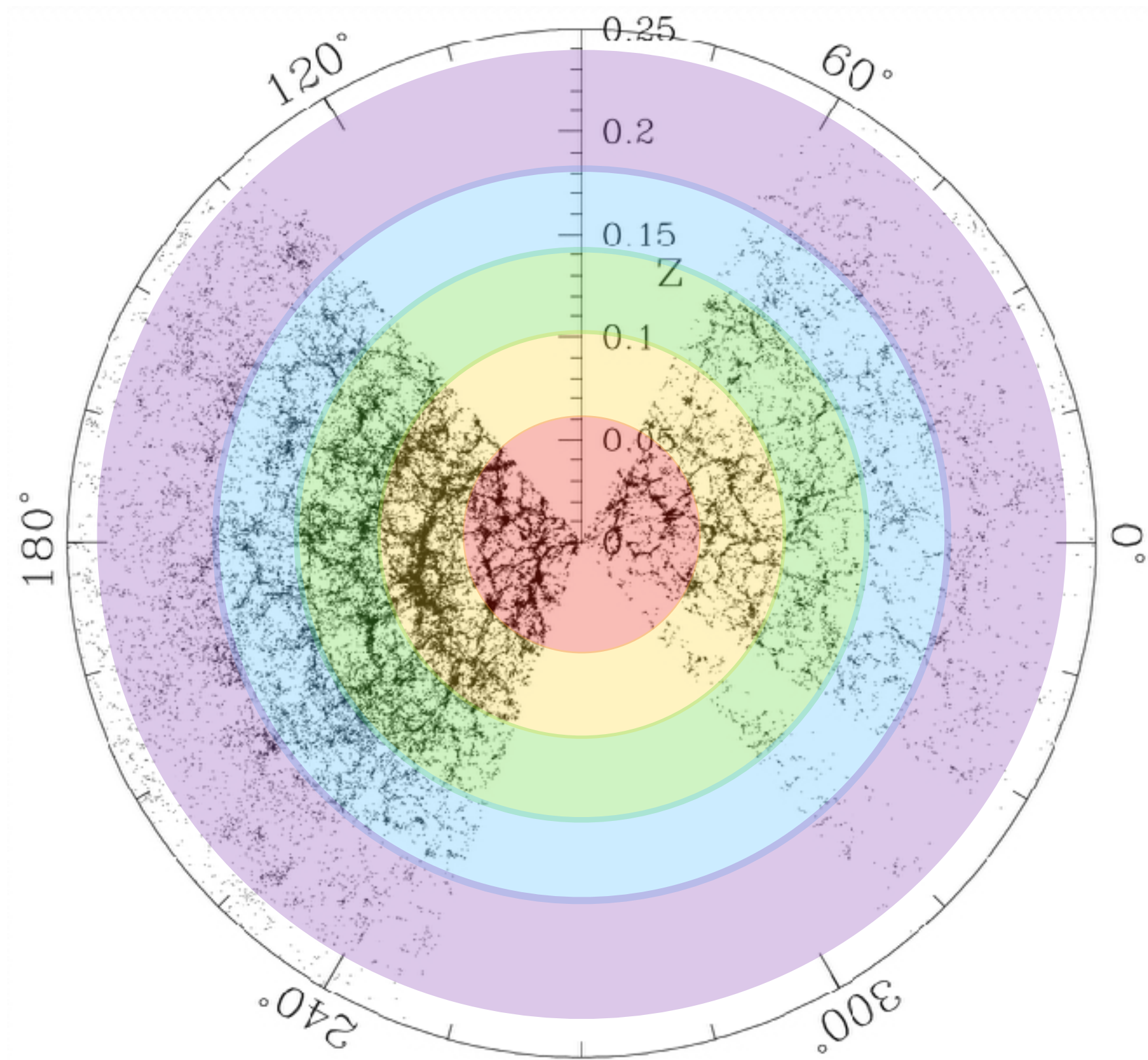
- Demonstration analysis of $\sim 300,000$ COSMOS2020 galaxies
- Comparison: FSPS under Prospector
25 CPU-hrs / galaxy

TL;DR for theorists

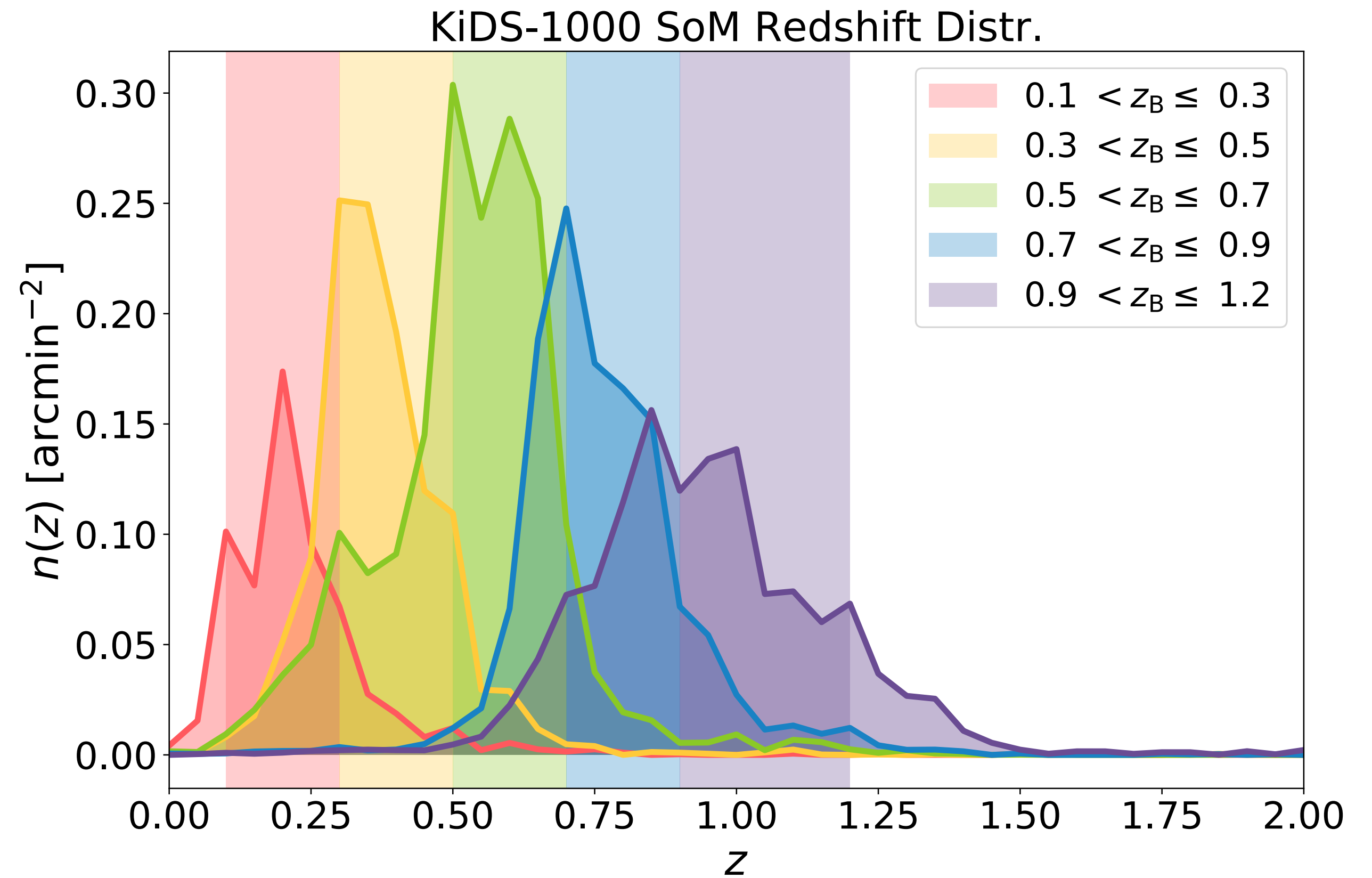
precision



Photometric catalogues require redshift estimation



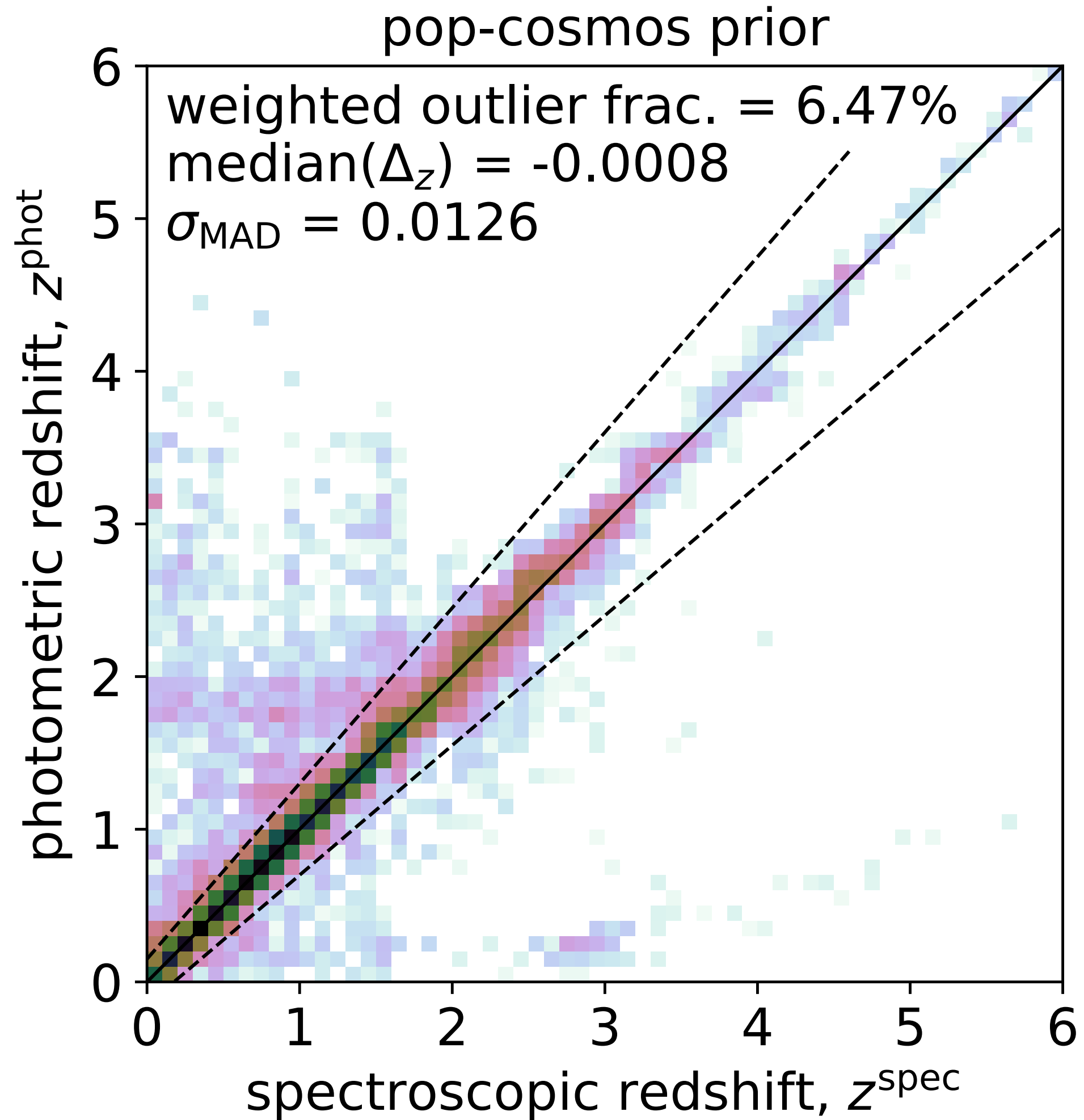
Blanton et al. (2003)



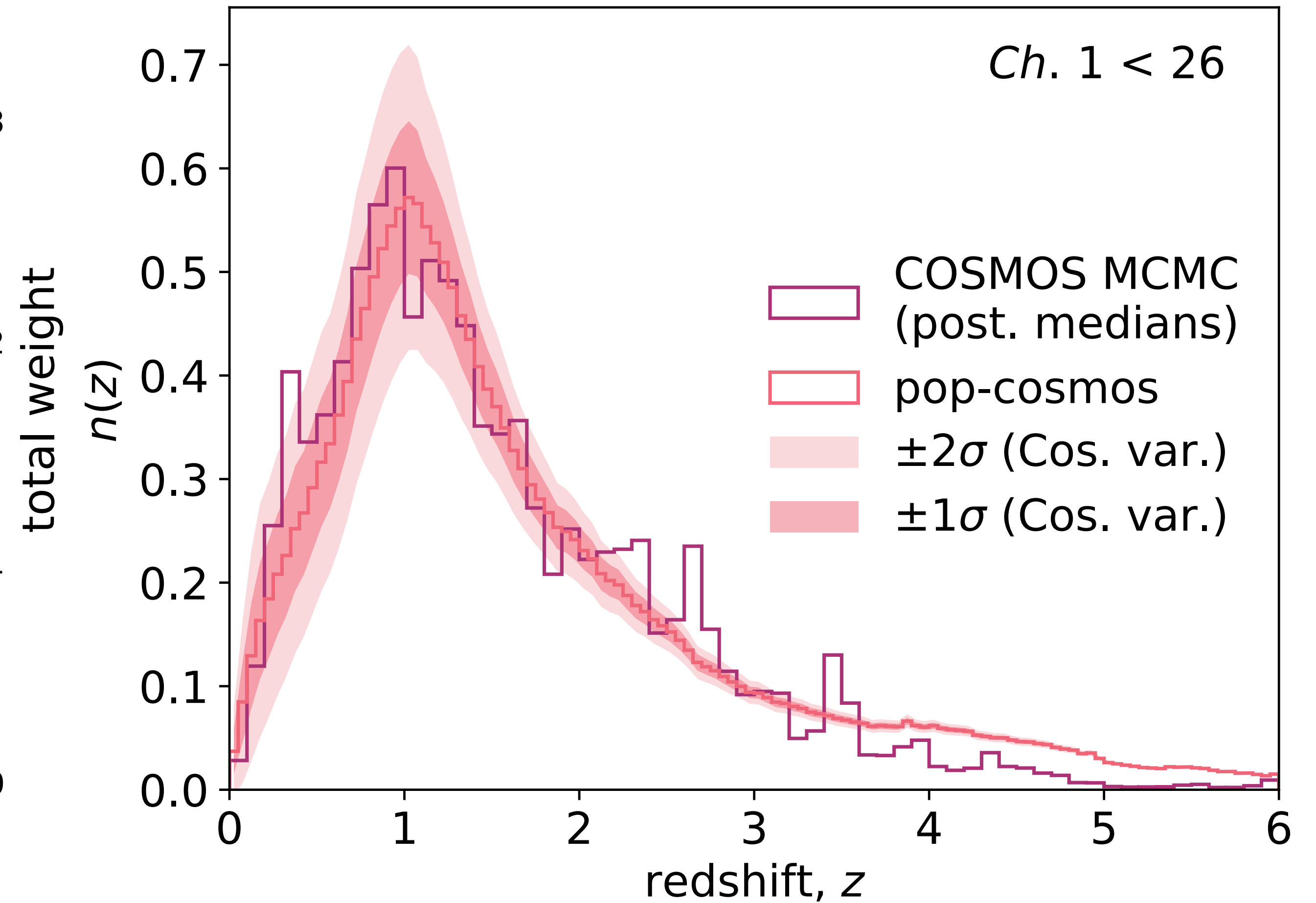
Loureiro et al. (2023)

Tomography: sort galaxies into bins and estimate underlying redshift distribution

pop-cosmos as a redshift inference engine



Deep spectroscopic validation with Khostovan (2026) spec-z compilation

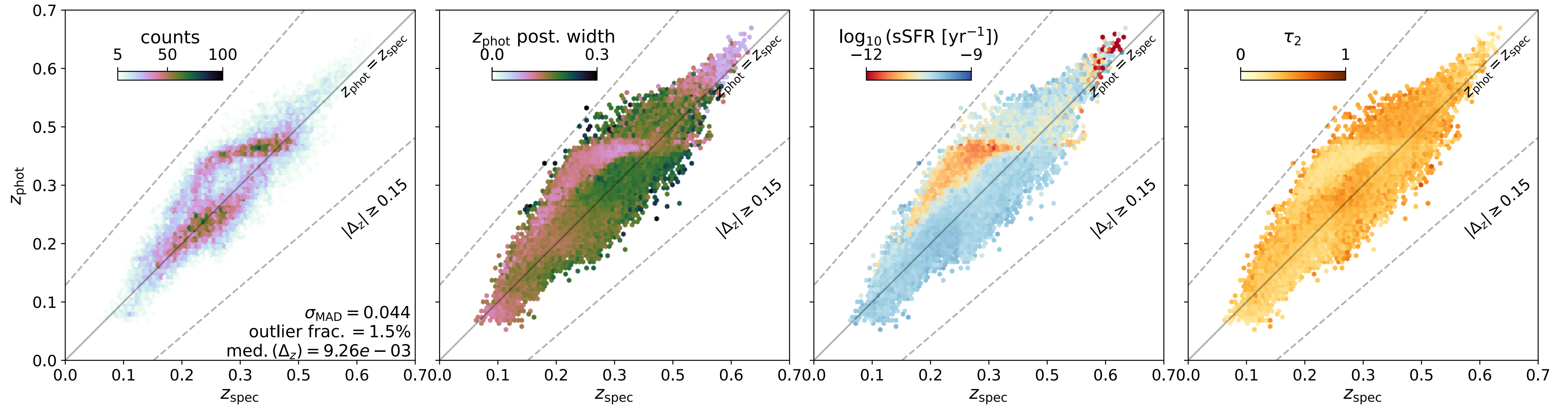


Pop-cosmos prediction for COSMOS2020 galaxy redshift distribution

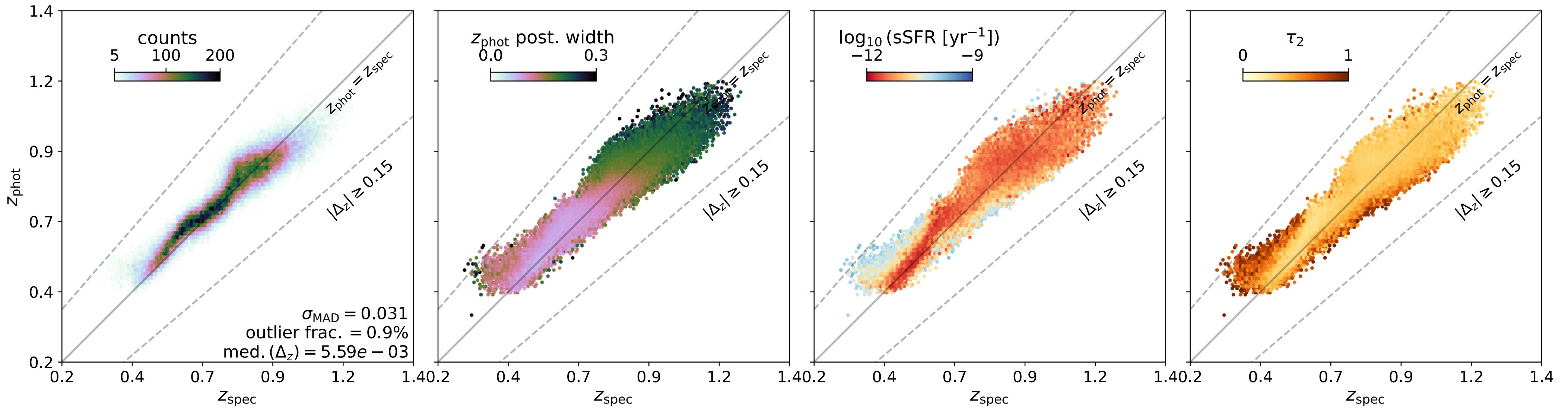
Generalisation to other surveys: DESI

KiDS-1000 × DESI BGS cross-matched galaxies; Inference under DESI (DECaLS + WISE) photometry

BGS



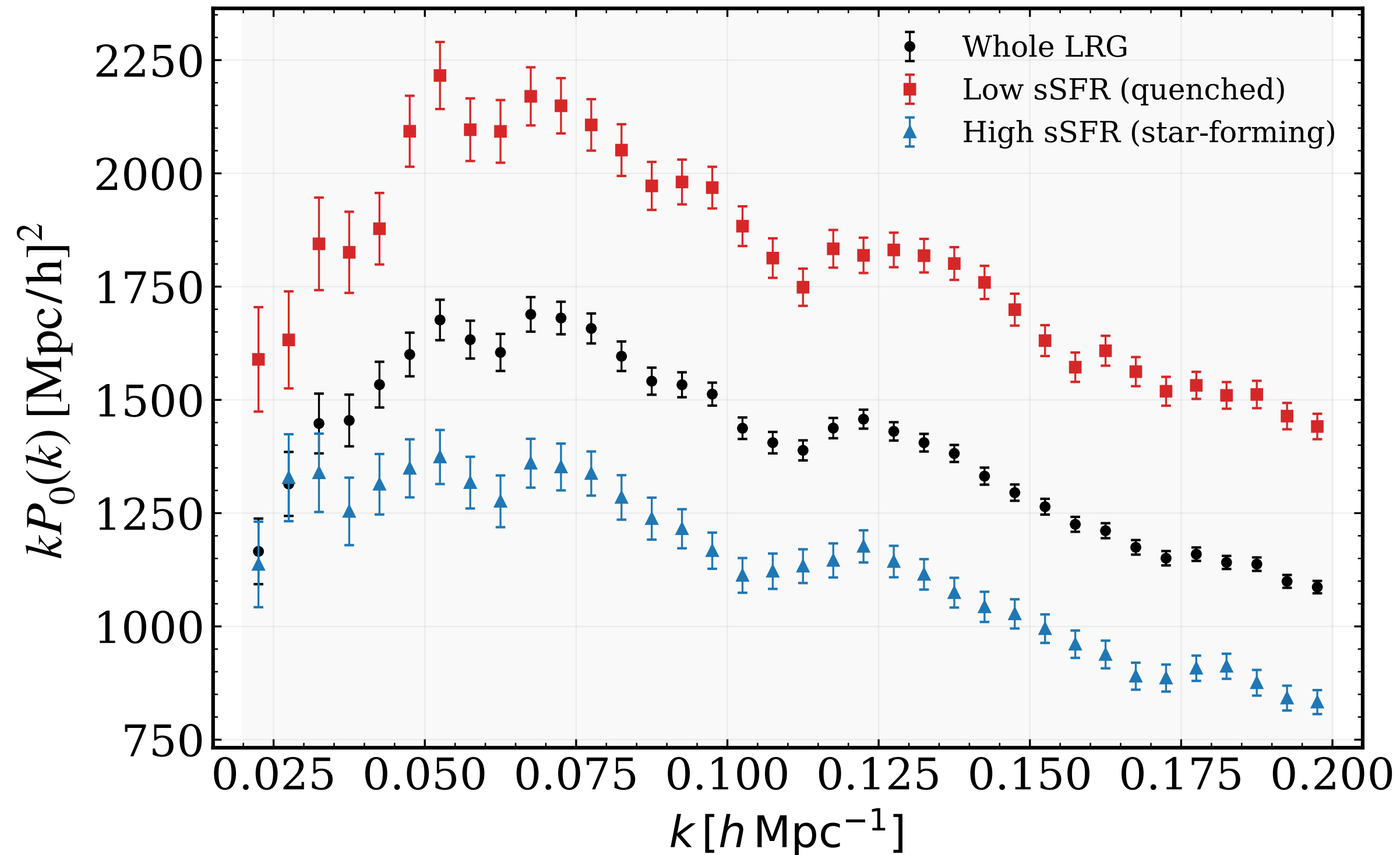
LRG



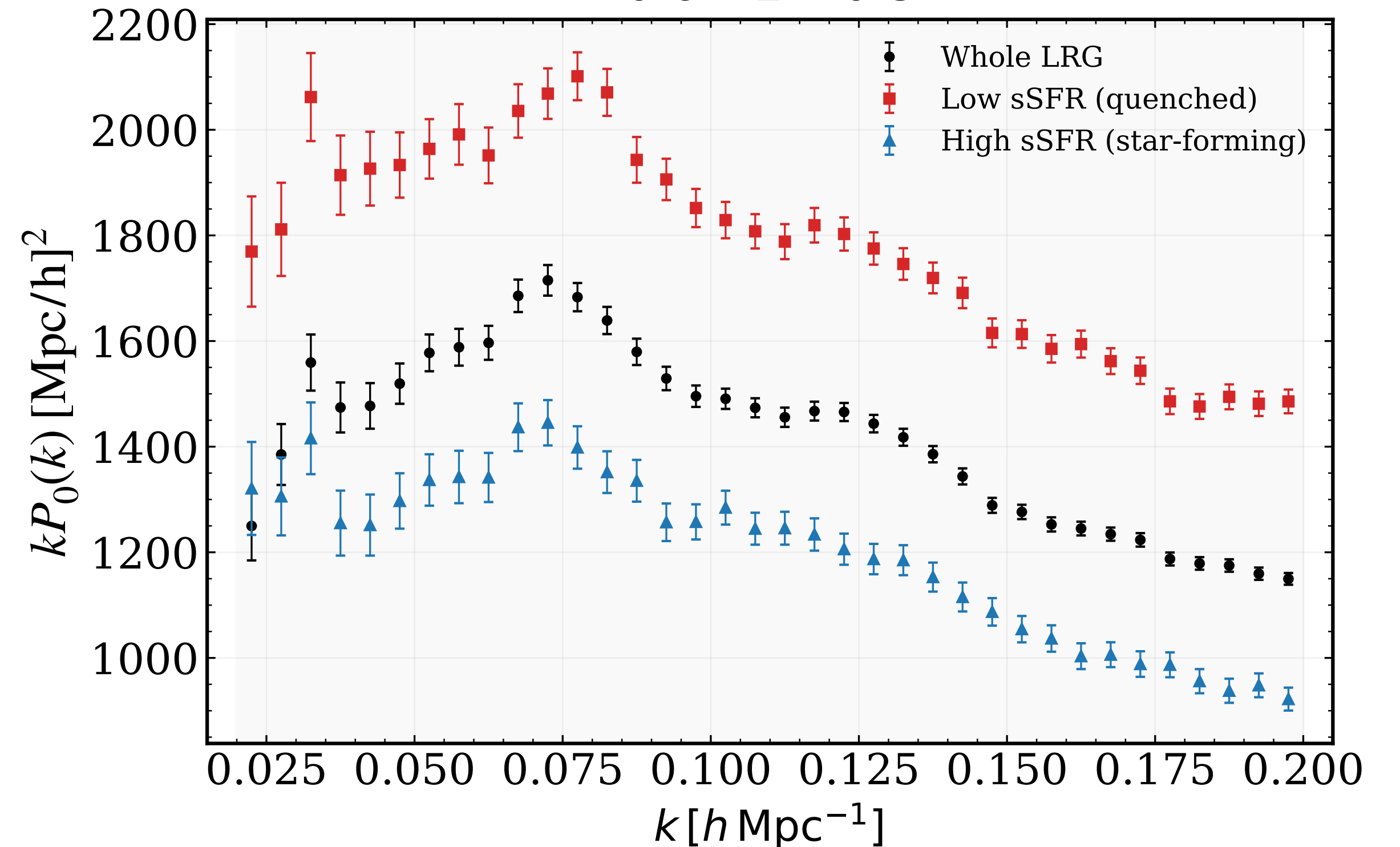
DESI spec-z, validating inference with DECaLS+WISE photometry under pop-cosmos prior

Multi-tracer cosmology (preliminary!)

$0.4 < z < 0.6$



$0.6 < z < 0.8$



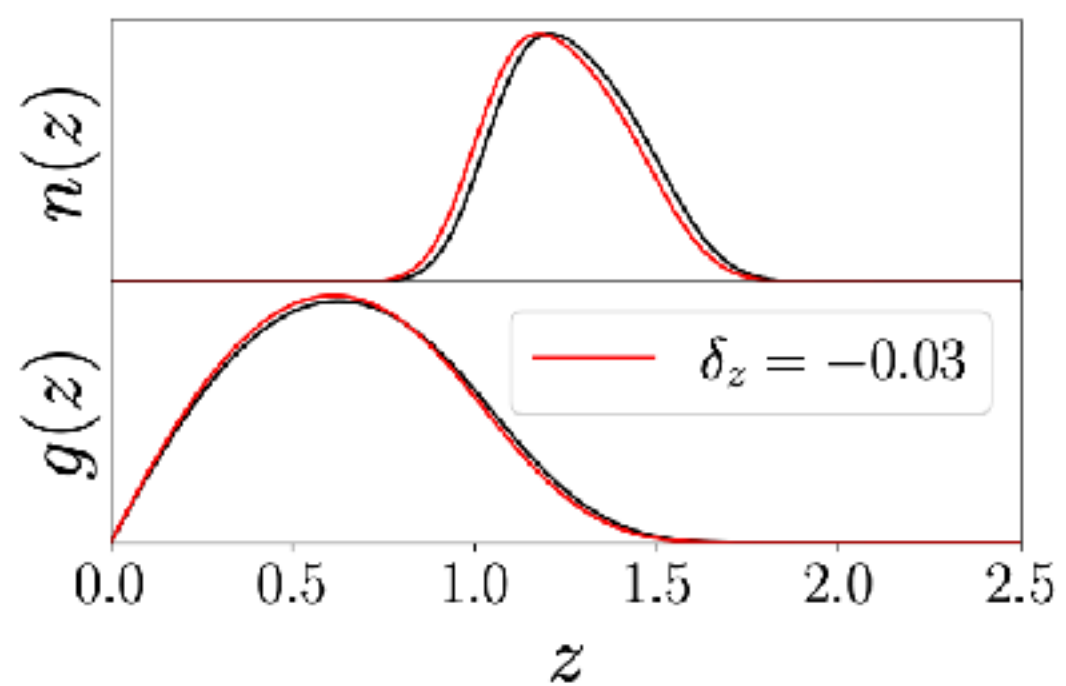
- DESI DRI Luminous Red Galaxies (LRG) \sim constant stellar mass $\sim 10^{11.3} M_{\odot}$
- Selection on physical properties (quiescent v star-forming); clustering measurements
- $\sim 240k$ galaxies (BIN1) and $\sim 370k$ galaxies (BIN2) in each quenched/SF sample

PRELIMINARY WORK LED BY DORSA SADAT HOSSEINI (PERIMETER INSTITUTE) AND ANIK HALDER (CAMBRIDGE)

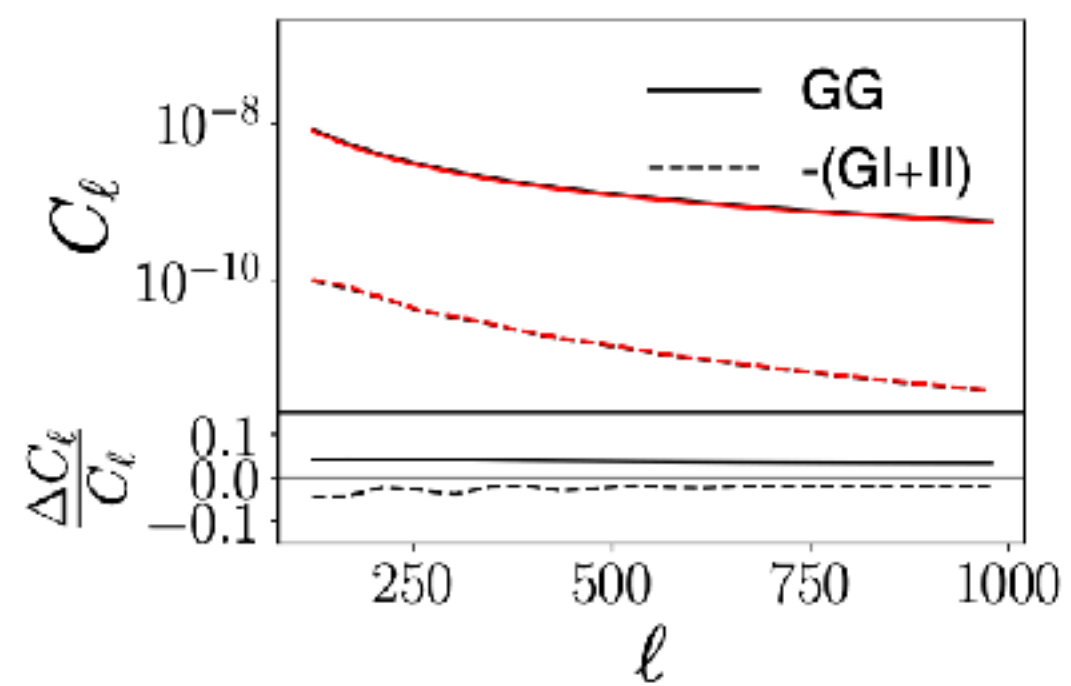


Accurate estimation of redshift distributions crucial

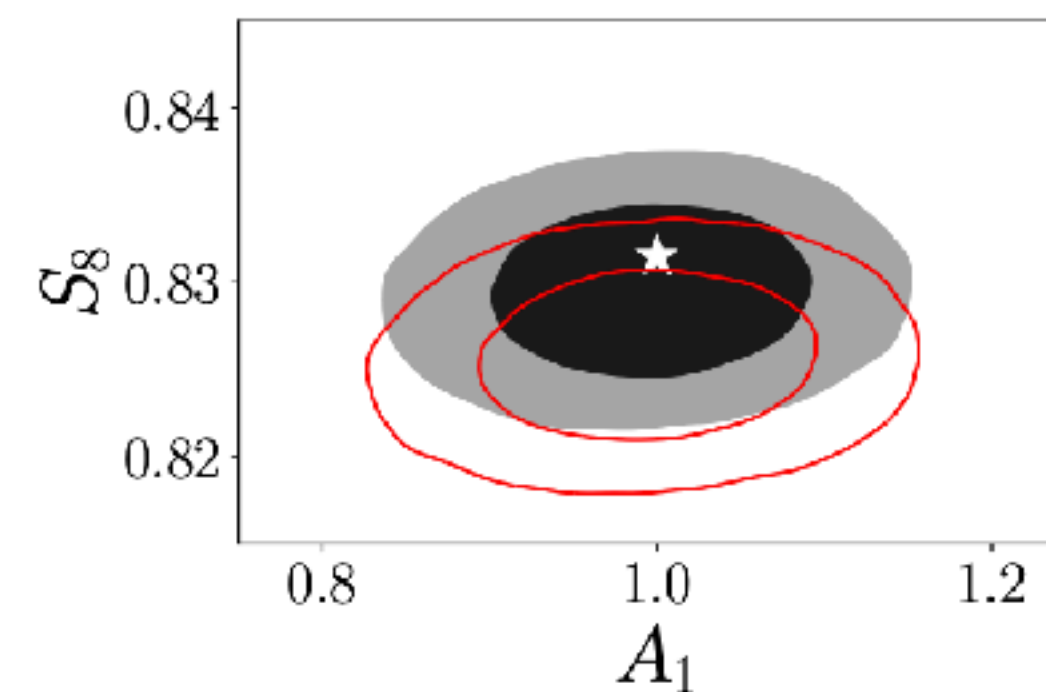
shifts



(a)

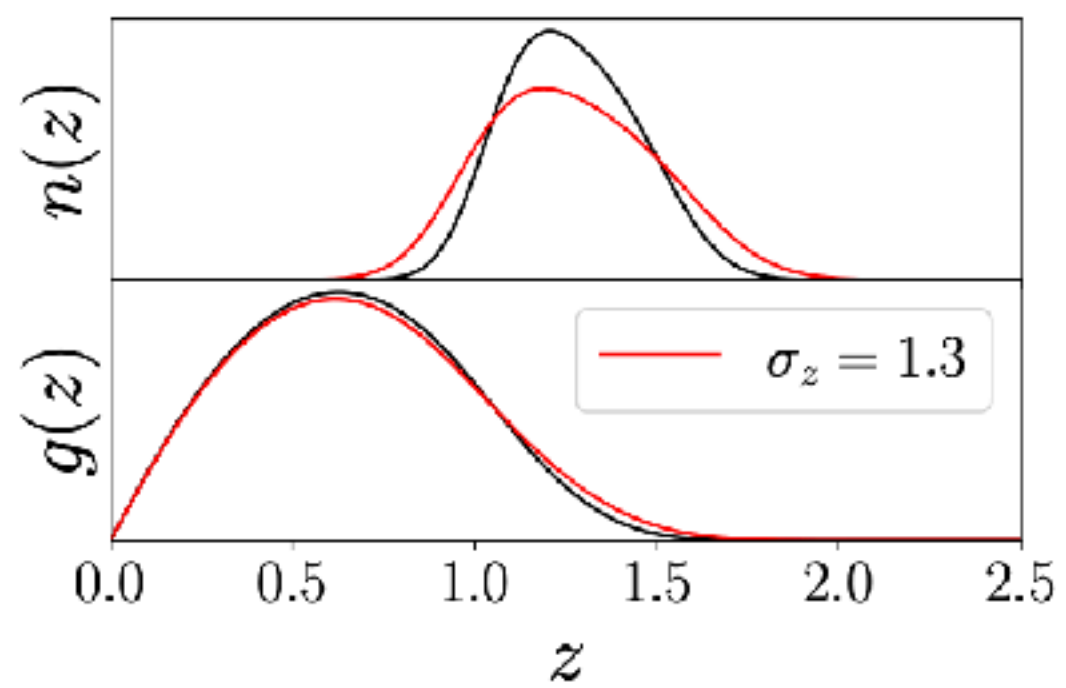


(b)

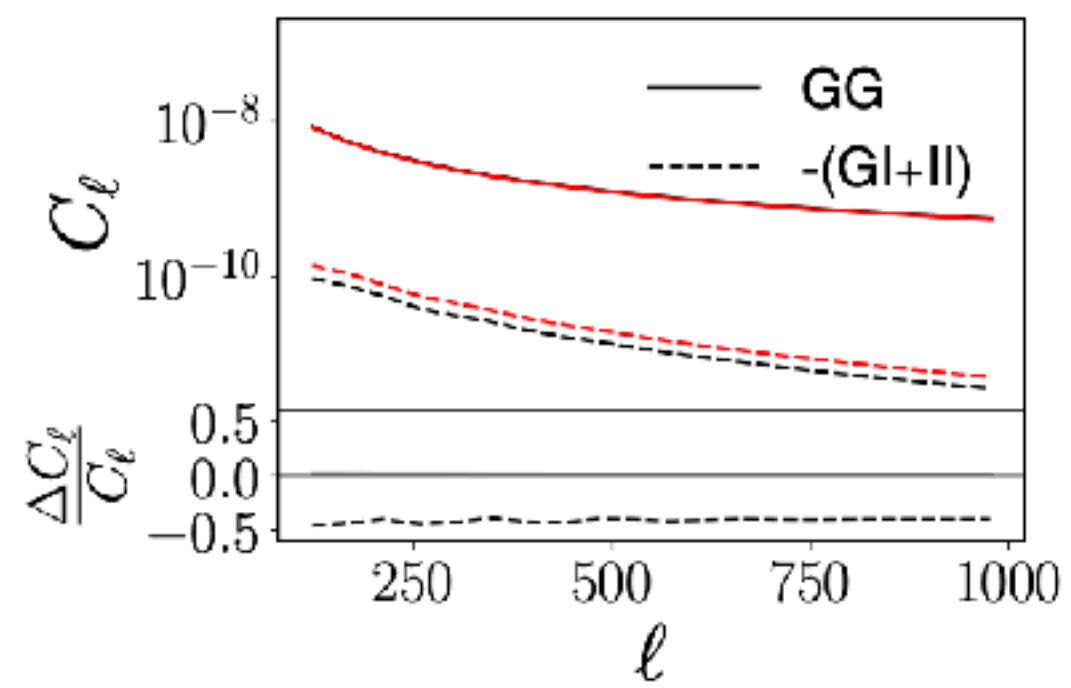


(c)

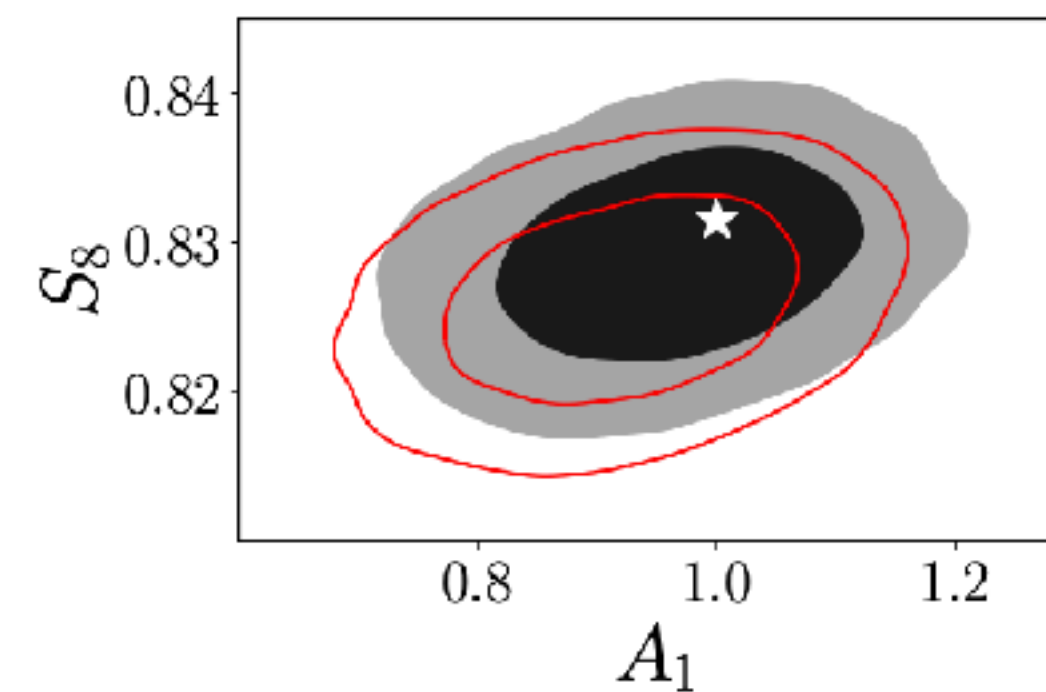
shape



(d)



(e)



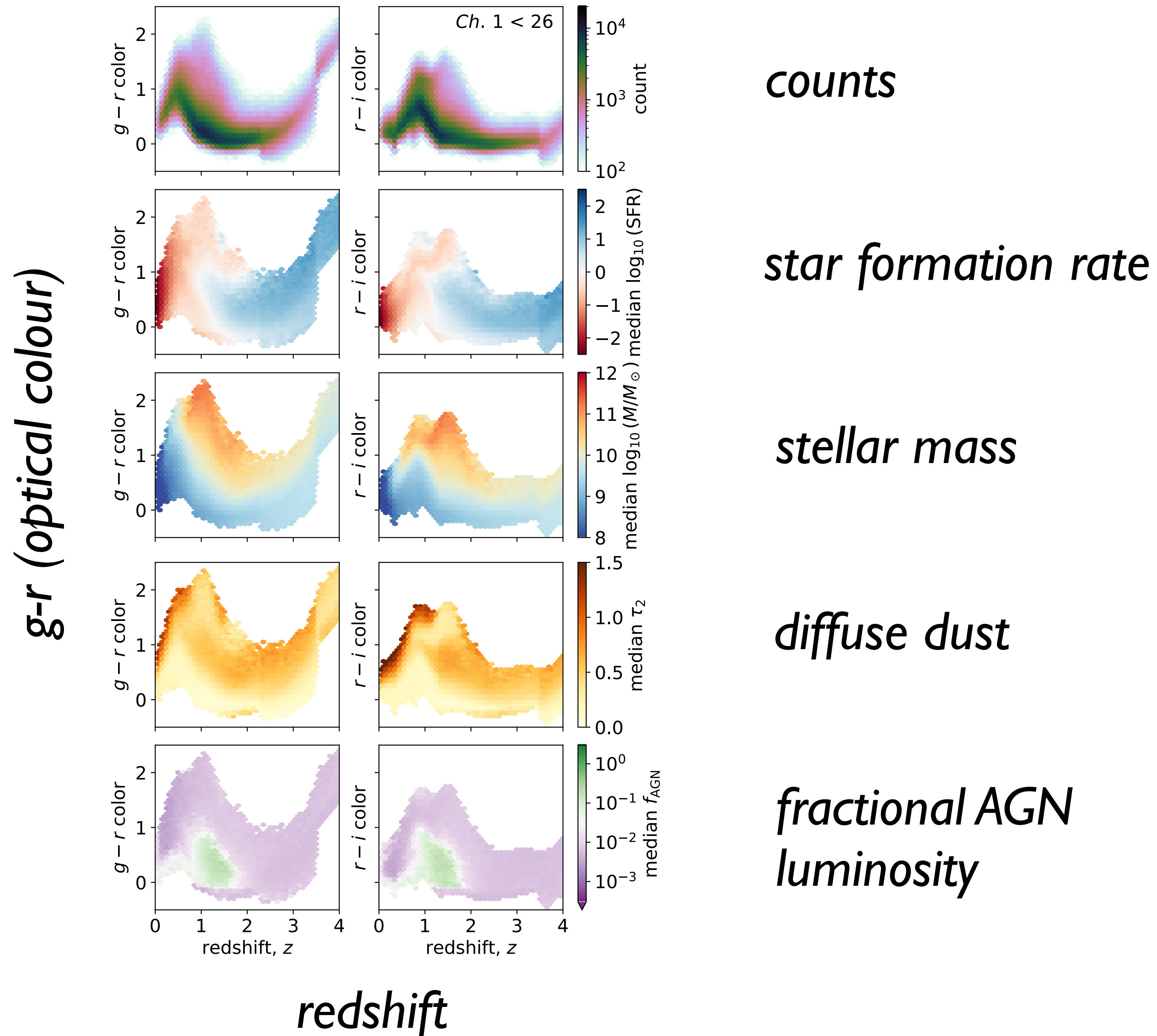
(f)

redshift distribution

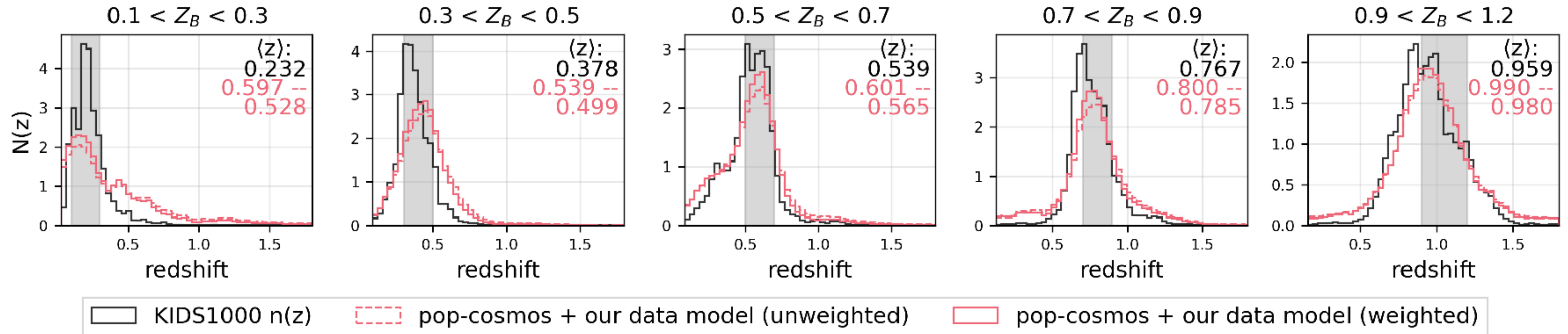
power spectra

parameters

What matters: the colour-redshift relation



pop-cosmos estimates for KiDS-1000 weak lensing tomography

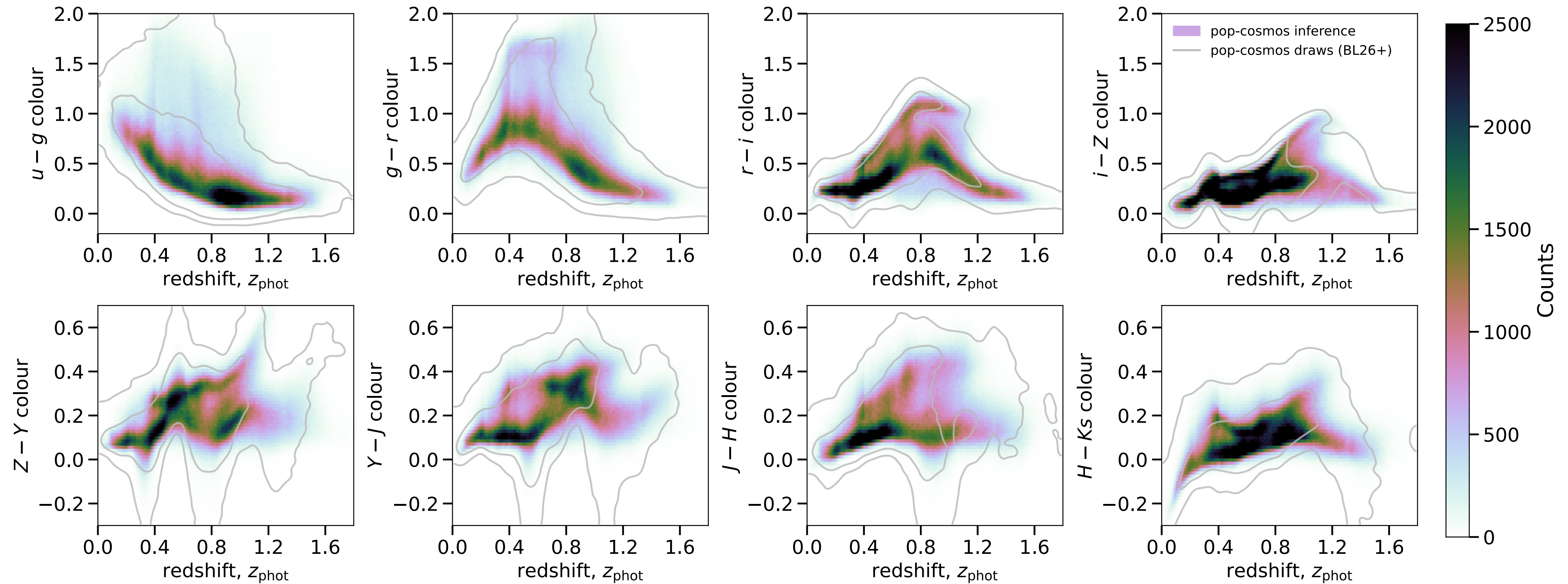


pop-cosmos & noise & detection & selection

- Data model, selection + detection calibrated on SKILLS image simulations, kindly provided by **Shun-Sheng Li**

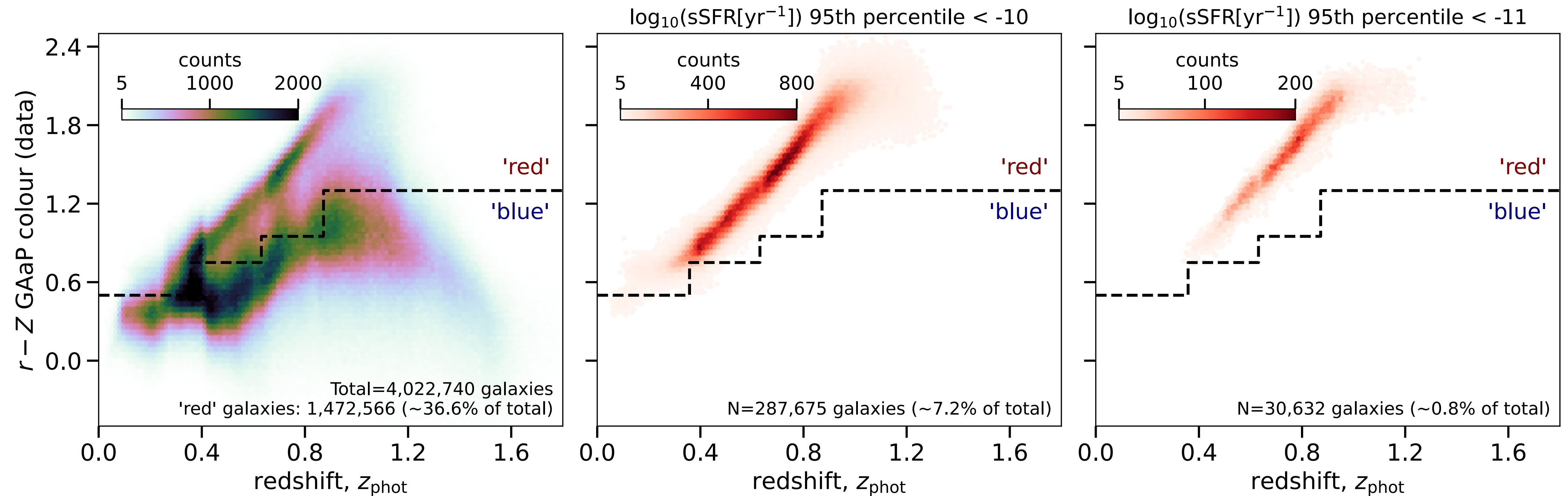
- Thanks to **Angus Wright** and KiDS team for sharing proprietary data and analysis settings!

Inference at scale: KiDS-1000



- Full Bayesian MCMC SED-fitting under *pop-cosmos* prior for 4 million KiDS-1000 galaxies using **affine** (6.5 GPU-sec / galaxy)

systematics-cleaned catalogues for cosmology



- Quiescent galaxies dominate intrinsic alignment (IA) signal: **star-forming selection** to mitigate IA
- Colour-selection (e.g. McCullough+2024) does not equal star-forming selection; cuts ~40% of sample
- Selection on physical properties cuts only 10%, cleaning systematics and preserving statistical power

Welcome to:



LEONARDO
CINECA



EuroHPC
Joint Undertaking

* Red Hat Enterprise Linux 8.7 (Ootpa)

*

* Booster module:

* Atos Bull Sequana X2135 "Da Vinci" Blade

* 3456 compute nodes with:

* - 32 cores Ice Lake at 2.60 GHz

* - 4 x NVIDIA Ampere A100 GPUs, 64GB

* - 512 GB RAM

*

* DataCentric General Purpose module (DCGP):

* Atos BullSequana X2140 Blade

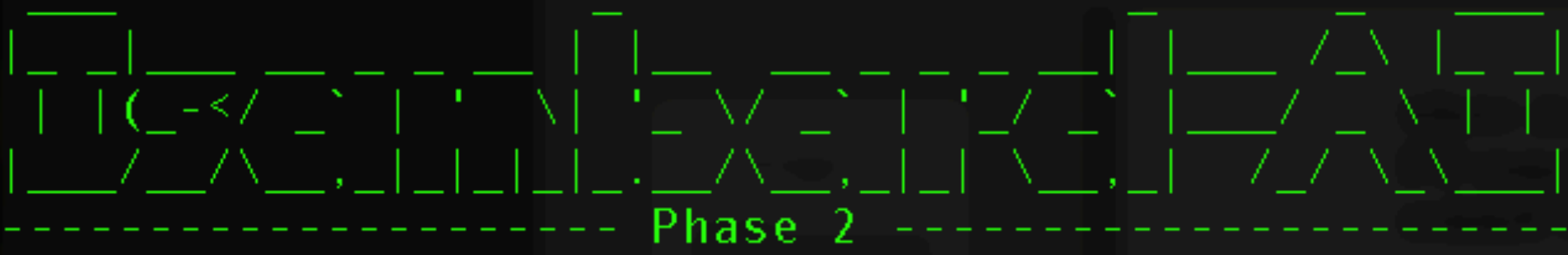
* 1536 compute nodes with:

* - 2x56 cores Intel Sapphire Rapids at 2.00 GHz

* - 512 GB RAM

rw-r--r--	1	interactive	1.7G	May	19	12	18	momc_summaries_0_1000000.h5
-rw-r--r--	1	interactive	1.7G	Jun	9	21	55	momc_summaries_1000000-2000000.h5
-rw-r--r--	1	interactive	1.7G	Jun	9	22	07	momc_summaries_1000000-11000000.r5
-rw-r--r--	1	interactive	1.7G	Jun	9	21	58	momc_summaries_1100000-12000000.r5
-rw-r--r--	1	interactive	1.7G	Jun	9	22	00	momc_summaries_1200000-13000000.r5
-rw-r--r--	1	interactive	1.7G	Jun	9	22	02	momc_summaries_1300000-14000000.r5
-rw-r--r--	1	interactive	1.7G	Jun	9	21	57	momc_summaries_1400000-15000000.r5
-rw-r--r--	1	interactive	1.7G	Jun	9	22	00	momc_summaries_1500000-16000000.r5
-rw-r--r--	1	interactive	1.7G	Jun	9	22	20	momc_summaries_1600000-17000000.r5
rw-r--r--	1	interactive	1.7G	Jun	9	22	06	momc_summaries_1700000-18000000.r5
-rw-r--r--	1	interactive	1.7G	Jun	9	22	08	momc_summaries_1800000-19000000.r5
-rw-r--r--	1	interactive	1.7G	Jun	9	22	12	momc_summaries_1900000-20000000.r5
-rw-r--r--	1	interactive	1.7G	Jun	9	21	59	momc_summaries_2000000-30000000.h5
rw-r--r--	1	interactive	1.7G	Jun	9	22	14	momc_summaries_2000000-21000000.r5
-rw-r--r--	1	interactive	1.7G	Jun	9	22	09	momc_summaries_2100000-22000000.r5
-rw-r--r--	1	interactive	1.7G	Jun	9	21	59	momc_summaries_2200000-23000000.r5
-rw-r--r--	1	interactive	1.7G	Jun	9	22	17	momc_summaries_2300000-24000000.r5
-rw-r--r--	1	interactive	1.7G	Jun	9	22	17	momc_summaries_2400000-25000000.r5
-rw-r--r--	1	interactive	1.7G	Jun	10	10	19	momc_summaries_2500000-26000000.r5
-rw-r--r--	1	interactive	1.7G	Jun	10	10	21	momc_summaries_2600000-27000000.r5
-rw-r--r--	1	interactive	1.7G	Jun	11	14	58	momc_summaries_2700000-28000000.r5
-rw-r--r--	1	interactive	1.7G	Jun	11	15	03	momc_summaries_2800000-29000000.r5
rw-r--r--	1	interactive	1.7G	Jun	11	15	00	momc_summaries_2900000-30000000.r5
-rw-r--r--	1	interactive	1.7G	Jun	9	21	59	momc_summaries_3000000-40000000.h5
-rw-r--r--	1	interactive	1.7G	Jun	15	00	12	momc_summaries_3100000-32000000.r5
-rw-r--r--	1	interactive	1.7G	Jun	15	00	34	momc_summaries_3200000-33000000.r5
rw-r--r--	1	interactive	1.7G	Jun	15	00	34	momc_summaries_3300000-34000000.r5
-rw-r--r--	1	interactive	1.7G	Jun	15	00	13	momc_summaries_3400000-35000000.r5
-rw-r--r--	1	interactive	1.7G	Jun	16	15	30	momc_summaries_3500000-36000000.r5
-rw-r--r--	1	interactive	1.7G	Jun	17	08	38	momc_summaries_3600000-37000000.r5
-rw-r--r--	1	interactive	1.7G	Jun	9	22	02	momc_summaries_4000000-50000000.h5
-rw-r--r--	1	interactive	1.7G	Jun	9	22	00	momc_summaries_5000000-60000000.h5
-rw-r--r--	1	interactive	1.7G	Jun	9	22	02	momc_summaries_6000000-70000000.h5
-rw-r--r--	1	interactive	1.7G	Jun	9	22	06	momc_summaries_7000000-80000000.h5
-rw-r--r--	1	interactive	1.7G	Jun	9	21	56	momc_summaries_8000000-90000000.h5
rw-r--r--	1	interactive	1.7G	Jun	9	21	57	momc_summaries_9000000-100000000.h5

**Bayesian inference for
>41 million KiDS-Legacy galaxies
84,000 GPU-hrs**



User Documentation: <https://docs.isambard.ac.uk>
 BriCS Helpdesk: <https://support.isambard.ac.uk>
 Service Status: <https://status.isambard.ac.uk>

Last login: Mon Jun 15 10:20:31 2026 from [REDACTED]

[(base) [REDACTED]]

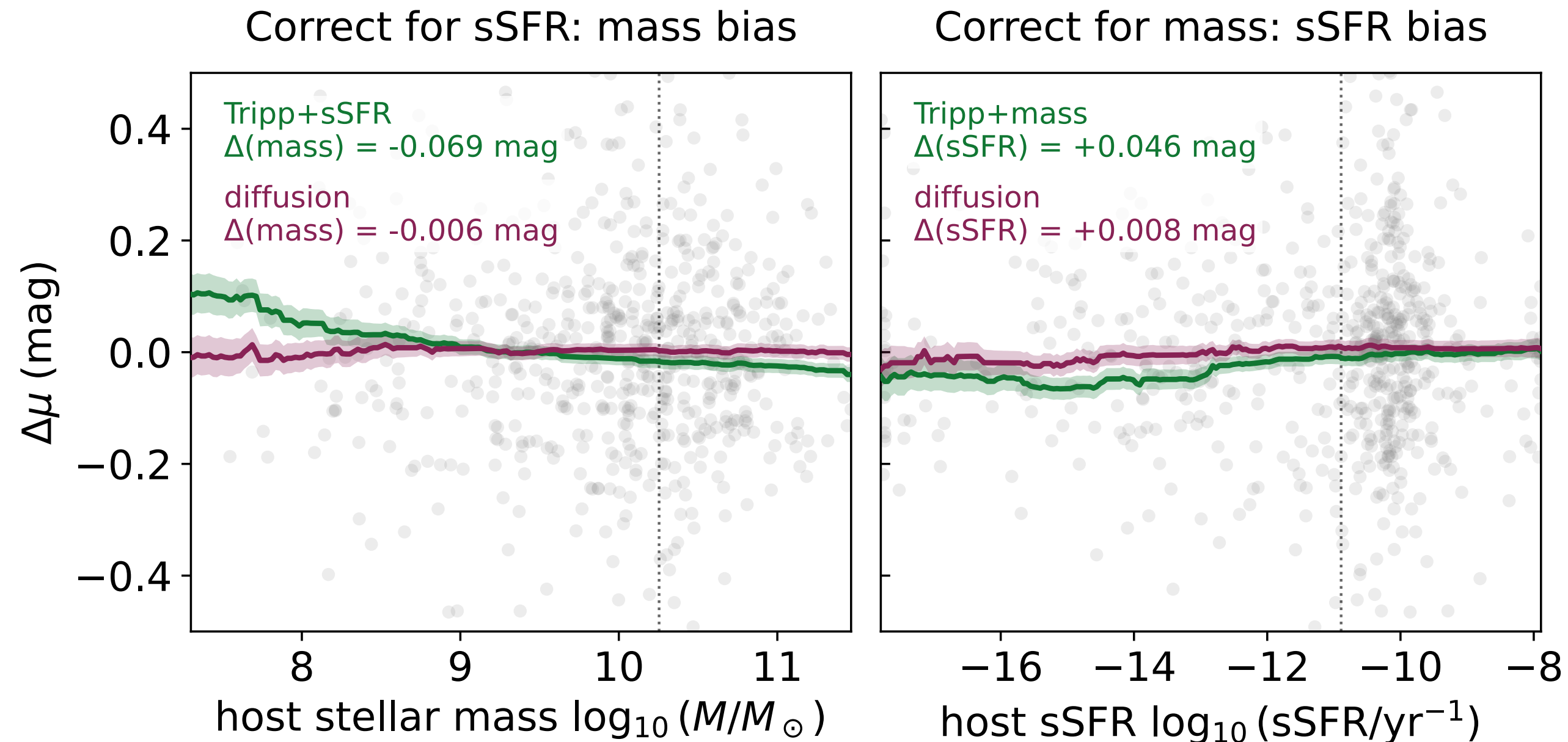
Fri Jun 19 09:33:02 2026

JOBID	PARTITION	NAME	USER	STATE	TIME	TIME_LIMI	NODES	NODELIST(REASON)
5299278_[198-199]	workq	BGS_5-6M	[REDACTED]	PENDING	0:00	1-00:00:00	1	(Priority)
5299185_[0-109]	workq	BGS_5-6M	[REDACTED]	PENDING	0:00	1-00:00:00	1	(Priority)
5292701_108	workq	BGS_4-5M	[REDACTED]	RUNNING	14:19:46	1-00:00:00	1	nid011099
5292701_109	workq	BGS_4-5M	[REDACTED]	RUNNING	14:19:46	1-00:00:00	1	nid011244
5292701_98	workq	BGS_4-5M	[REDACTED]	RUNNING	14:20:38	1-00:00:00	1	nid010208
5292701_99	workq	BGS_4-5M	[REDACTED]	RUNNING	14:20:38	1-00:00:00	1	nid010368
5292701_100	workq	BGS_4-5M	[REDACTED]	RUNNING	14:20:38	1-00:00:00	1	nid010624
5292701_101	workq	BGS_4-5M	[REDACTED]	RUNNING	14:20:38	1-00:00:00	1	nid010624
5292701_102	workq	BGS_4-5M	[REDACTED]	RUNNING	14:20:38	1-00:00:00	1	nid010717
5292701_103	workq	BGS_4-5M	[REDACTED]	RUNNING	14:20:38	1-00:00:00	1	nid010482
5292701_104	workq	BGS_4-5M	[REDACTED]	RUNNING	14:20:38	1-00:00:00	1	nid010505
5292701_105	workq	BGS_4-5M	[REDACTED]	RUNNING	14:20:38	1-00:00:00	1	nid011029
5292701_106	workq	BGS_4-5M	[REDACTED]	RUNNING	14:20:38	1-00:00:00	1	nid011042
5292701_107	workq	BGS_4-5M	[REDACTED]	RUNNING	14:20:38	1-00:00:00	1	nid011042
5292701_94	workq	BGS_4-5M	[REDACTED]	RUNNING	14:21:35	1-00:00:00	1	nid010647
5292701_95	workq	BGS_4-5M	[REDACTED]	RUNNING	14:21:35	1-00:00:00	1	nid010789
5292701_96	workq	BGS_4-5M	[REDACTED]	RUNNING	14:21:35	1-00:00:00	1	nid010257
5292701_97	workq	BGS_4-5M	[REDACTED]	RUNNING	14:21:35	1-00:00:00	1	nid010272
5292701_62	workq	BGS_4-5M	[REDACTED]	RUNNING	14:22:26	1-00:00:00	1	nid010200
5292701_63	workq	BGS_4-5M	[REDACTED]	RUNNING	14:22:26	1-00:00:00	1	nid010200
5292701_64	workq	BGS_4-5M	[REDACTED]	RUNNING	14:22:26	1-00:00:00	1	nid010315
5292701_65	workq	BGS_4-5M	[REDACTED]	RUNNING	14:22:26	1-00:00:00	1	nid010315
5292701_66	workq	BGS_4-5M	[REDACTED]	RUNNING	14:22:26	1-00:00:00	1	nid010921
5292701_67	workq	BGS_4-5M	[REDACTED]	RUNNING	14:22:26	1-00:00:00	1	nid010479
5292701_68	workq	BGS_4-5M	[REDACTED]	RUNNING	14:22:26	1-00:00:00	1	nid010479



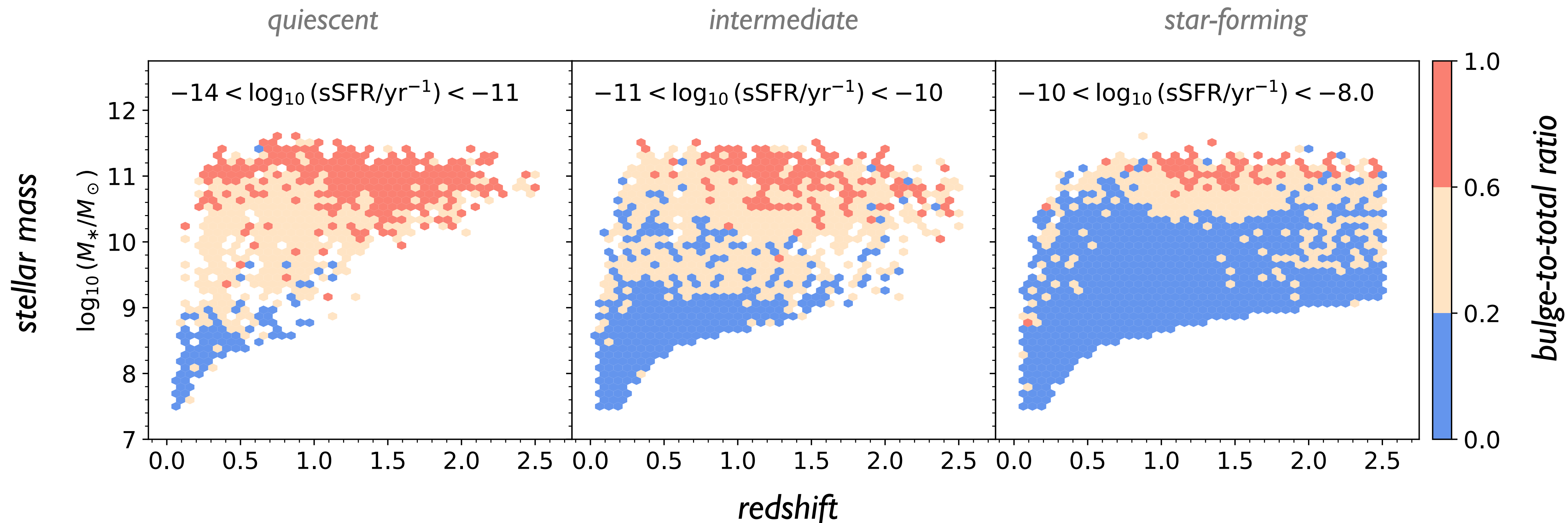
Bayesian inference for
 billion-galaxy surveys
 620,000 GPU-hrs

Supernova-host connection



- ZTF DR2 spectroscopically-confirmed SNe Ia (633 hosts after cuts)
- Host properties from fitting multiwavelength photometry (Blast: SDSS/DES/Pan-STARRS/2MASS/WISE/GALEX) under pop-cosmos population-level prior
- Single-property-step Tripp standardisation vs generalised diffusion model over joint set of host properties (mass, sSFR, dust, metallicity).

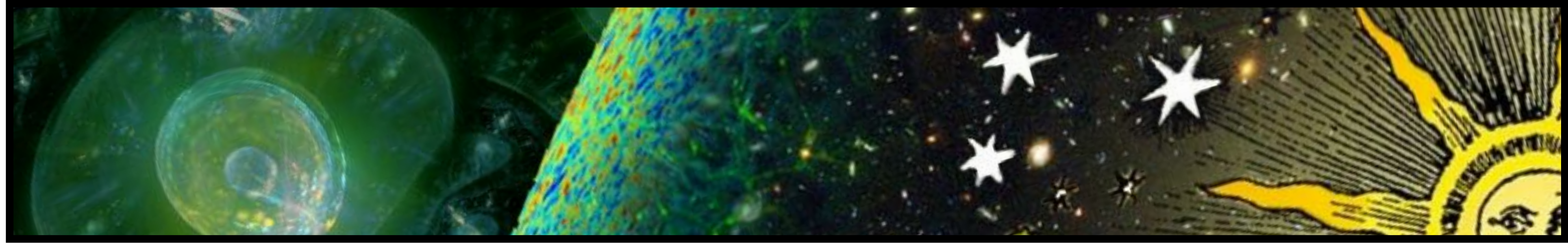
Towards a generative model for galaxy morphology



- colour-coded by bulge-to-total ratio computed from JWST photometry for three sSFR bins
- in process of extending generative model to galaxy morphology+size dimension with image features
- structural trends cannot be pulled out of even the best ground-based photometry (HSC)

In the pipeline

- ***Weak lensing analysis of LSST precursor KiDS Legacy catalogue using star-forming selection***
- ***Even faster inference with Amortized Neural Posterior Estimation (ANPE)***
- ***Galaxy-halo connection via conditional neural abundance matching***



Papers



ALSING+ (2020)



ALSING+ (2024)



THORP+ (2024A)



THORP+ (2024B)



THORP+ (2025)



DEGER+ (2025)



HALDER+ (2026)



LEISTEDT+ (2026)

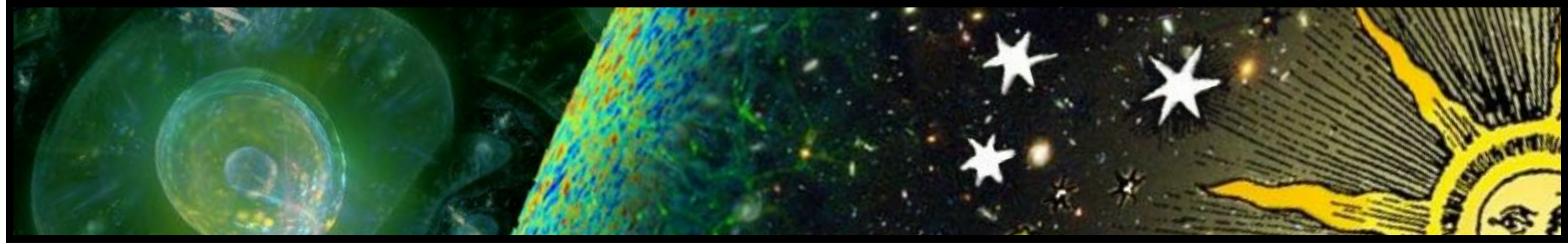


VAN DEN BUSSCHE+ (2026)



erc

European Research Council
Established by the European Commission



Software and data products



Speculator



pop-cosmos



affine



flowfusion

cosmo-pop.github.io



