pop-cosmos: Insights from generative modeling of a deep, infrared-selected galaxy population

Stephen Thorp [ID],[1,2] Hiranya V. Peiris [ID],[2,1] Gurjeet Jagwani [ID],[2,3] Sinan Deger [ID],[2] Justin Alsing [ID],[1]
Boris Leistedt [ID],[4] Daniel J. Mortlock [ID],[4,5] Anik Halder [ID],[2] and Joel Leja [ID][6,7,8]

[1] *The Oskar Klein Centre, Department of Physics, Stockholm University, AlbaNova University Centre, SE 106 91 Stockholm, Sweden*
[2] *Institute of Astronomy and Kavli Institute for Cosmology, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK*
[3] *Research Computing Services, University of Cambridge, Roger Needham Building, 7 JJ Thomson Ave, Cambridge CB3 0RB, UK*
[4] *Astrophysics Group, Imperial College London, Blackett Laboratory, Prince Consort Road, London, SW7 2AZ, UK*
[5] *Department of Mathematics, Imperial College London, London SW7 2AZ, UK*
[6] *Department of Astronomy & Astrophysics, The Pennsylvania State University, University Park, PA 16802, USA*
[7] *Institute for Computational & Data Sciences, The Pennsylvania State University, University Park, PA 16802, USA*
[8] *Institute for Gravitation & the Cosmos, The Pennsylvania State University, University Park, PA 16802, USA*

ABSTRACT

We present an extension of the pop-cosmos model for the evolving galaxy population up to redshift $z \sim$ 6. The model is trained on distributions of observed colors and magnitudes, from 26-band photometry of $\sim 420,000$ galaxies in the COSMOS2020 catalog with *Spitzer* IRAC *Ch. 1* < 26. The generative model includes a flexible distribution over 16 stellar population synthesis (SPS) parameters, and a depth-dependent photometric uncertainty model, both represented using score-based diffusion models. We use the trained model to predict scaling relationships for the galaxy population, such as the stellar mass function, star-forming main sequence, and gas-phase and stellar metallicity vs. mass relations, demonstrating reasonable-to-excellent agreement with previously published results. We explore the connection between mid-infrared emission from active galactic nuclei (AGN) and star-formation rate, finding high AGN activity for galaxies above the star-forming main sequence at $1 \lesssim z \lesssim 2$. Using the trained population model as a prior distribution, we perform inference of the redshifts and SPS parameters for 429,669 COSMOS2020 galaxies, including 39,588 with publicly available spectroscopic redshifts. The resulting redshift estimates exhibit minimal bias (median$[\Delta_z] = -8 \times 10^{-4}$), scatter ($\sigma_{\mathrm{MAD}} = 0.0132$), and outlier fraction (6.19%) for the full $0 < z < 6$ spectroscopic compilation. These results establish that pop-cosmos can achieve the accuracy and realism needed to forward-model modern wide–deep surveys for Stage IV cosmology. We publicly release pop-cosmos software, mock galaxy catalogs, and COSMOS2020 redshift and SPS parameter posteriors.

1. INTRODUCTION

The next generation of imaging surveys, such as those which will be undertaken by the Vera C. Rubin Observatory (Ž. Ivezić et al. 2019), *Euclid* (Euclid Collaboration et al. 2025a), and the *Nancy Grace Roman Space Telescope* (D. Spergel et al. 2015), will produce photometric catalogs containing many billions of galaxies. The scale of these data-sets means that traditional data analysis techniques, typically based on parametric models and individual object likelihoods, will not be viable. To properly account for the complexity of these data-sets, including selection, will require the use of non-parametric

models, forward simulation of the data, and fast emulation of observable galaxy properties.

The need for accurate photometric redshifts for Stage IV cosmology surveys (see R. Mandelbaum 2018; J. A. Newman & D. Gruen 2022) has motivated the development of sophisticated forward-modeling and simulation based inference techniques for photometric surveys. To infer a population-level redshift distribution, or obtain accurate redshifts for individual objects, it is essential to understand how galaxy spectral energy distributions (SEDs) evolve over cosmic time. Moreover, modeling the full galaxy population provides a way of controlling systematics such as intrinsic alignments (see, e.g., R. Mandelbaum et al. 2006, 2011; M. A. Troxel & M. Ishak 2015; B. Joachimi et al. 2015), supernova–host correlations (see, e.g., M. J. Childress et al. 2014;

Corresponding author: Stephen Thorp
stephen.thorp@ast.cam.ac.uk

P. Wiseman et al. 2022, 2023; D. O. Jones et al. 2024; K. Karchev et al. 2025), and the effect of baryonic feedback (see, e.g., I. G. McCarthy et al. 2017; J. Schaye et al. 2023). In this way, redshift inference and cosmology are inextricably linked to the evolution of the galaxy population (for an overview, see J. Alsing et al. 2023).

Numerous parallel developments over the past decade have been moving toward addressing these challenges for Stage IV cosmology. These have included: fully differentiable implementations of physical models (e.g., A. P. Hearin et al. 2021, 2023; A. Alarcon et al. 2023); emulators for stellar population synthesis (SPS) models (e.g., Y. Han & Z. Han 2012, 2014; J. Alsing et al. 2020; K. J. Kwon et al. 2023; E. P. Mathews et al. 2023; L. Tortorelli et al. 2025); machine-learning accelerators for hydrodynamical or radiative transfer models (e.g. C. C. Lovell et al. 2019; S. Gilda et al. 2021; Y. Qiu & X. Kang 2022); parametric population models combined with rapid image simulation based on morphology indicators (J. Herbel et al. 2017; B. Moser et al. 2024; S. Fischbacher et al. 2024a,b, 2025b; L. Tortorelli et al. 2025); flow-based models for the color–redshift relation (e.g J. F. Crenshaw et al. 2024); SPS-based forward-modeling for spectro-photometric data (C. Hahn et al. 2023a, 2024); flow-based population models (J. Li et al. 2024); galaxy- and population-level Bayesian model comparison (Y. Han & Z. Han 2019; Y. Han et al. 2023); explicit physically-motivated priors on galaxy properties in template fitting (e.g. M. Tanaka 2015); hierarchical models for clustering/cross-correlation redshift distributions (e.g., C. Sánchez & G. M. Bernstein 2019; M. M. Rau et al. 2020, 2023); stratified learning techniques (M. Autenrieth et al. 2024); and neural posterior estimators (C. Hahn & P. Melchior 2022; B. Wang et al. 2023a).

Our work on galaxy population modeling was also driven initially by the problem of photometric redshift distributions, starting with a Bayesian hierarchical model for inferring galaxy redshift distributions (B. Leistedt et al. 2016). One significant limitation was the use of a finite set of predefined template galaxy SEDs, motivating more flexible data-driven approaches to SED modeling (B. Leistedt & D. W. Hogg 2017; B. Leistedt et al. 2019). The limited constraining power of photometric data alone also motivated the inclusion of physics constraints via SPS (for a review of SPS see, e.g., B. M. Tinsley 1980; C. Conroy 2013; K. G. Iyer et al. 2025), for which we developed a fast neural network emulator, Speculator (J. Alsing et al. 2020). Speculator was combined with a parametric model of the galaxy population (including SPS parameters) to infer galaxy redshift distributions with sufficient fidelity to meet Stage III survey requirements (J. Alsing et al. 2023) and to self-calibrate the data model, e.g., zero points (B. Leistedt et al. 2023). However, for Stage IV surveys the use of parametric distributions for the galaxy population and SPS parameters is too restrictive. To over-

come this limitation, in J. Alsing et al. (2024, hereafter A24) we used a flexible diffusion model which was calibrated on the COSMOS2020 catalog (J. R. Weaver et al. 2022) from the Cosmic Evolution Survey (COSMOS; N. Scoville et al. 2007a) to produce the pop-cosmos galaxy population model. We also developed a complementary framework for rapid SPS parameter inference for individual objects using the pop-cosmos population model as a prior (S. Thorp et al. 2024a, hereafter T24).

In this paper, we present an extended version of the pop-cosmos generative model, calibrated on a *Spitzer* IRAC-selected galaxy sample from the COSMOS2020 catalog (J. R. Weaver et al. 2022), which is a magnitude deeper than the sample used in A24. We also present a first public release of software for sampling from and evaluating the pop-cosmos population model, along with COSMOS-like mock galaxy catalogs drawn from it, and an update to the posterior samples from T24 for individual COSMOS galaxies.

In Section 2, we describe the selection of our training data. In Section 3, we give an overview of the pop-cosmos modeling framework and the improvements made since A24, with the performance of the new model evaluated in Section 4. Astrophysical insights extracted from the model are presented in Section 5. We discuss the wider context of our results, and directions for future developments in Section 6. This paper will be followed by both a more detailed astrophysical application of the new model (S. Deger et al., in preparation) and two different applications to the Kilo Degree Survey (KiDS; K. Kuijken et al. 2019) catalogs (B. Leistedt et al. and A. Halder et al., both in preparation).

Where relevant, we assume a flat ΛCDM cosmology with $H_0 = 67.66$ km s^{-1} Mpc^{-1} and $\Omega_{\mathrm{m}} = 0.3097$ (Planck Collaboration et al. 2020). Magnitudes and relative fluxes are all based on the AB system with a reference flux density of $f_{\mathrm{AB}} = 3631$ Jy (J. B. Oke & J. E. Gunn 1983; M. Fukugita et al. 1996).

2. DATA

We now describe the multi-wavelength photometric data we use for calibrating the generative model presented in this work, and the spectroscopic data we use for validation.

2.1. *Photometric Data*

As in A24, we use 26-band photometry from the COSMOS2020 catalog (J. R. Weaver et al. 2022); specifically v2.1 of the Farmer catalog that uses a profile-fitting method for the photometric extraction (J. R. Weaver et al. 2023a). From COSMOS2020, we use u-band data from the Canada–France–Hawai'i–Telescope (C. Laigle et al. 2016; M. Sawicki et al. 2019), $grizy$ data from Subaru Hyper Suprime-Cam (HSC; H. Aihara et al. 2019), $YJHK_S$ data from UltraVISTA (H. J. McCracken et al. 2012), Channel 1 & 2 of *Spitzer* IRAC (C. L. Steinhardt et al. 2014; M. L. N. Ashby et al. 2013, 2015, 2018; Eu-

clid Collaboration et al. 2022), and optical narrow- and medium-band data from Suprime-Cam (Y. Taniguchi et al. 2007, 2015). We select our sample within the 1.27 deg² "combined" mask for COSMOS2020, which includes HSC, UltraVISTA, and IRAC coverage, and avoids bright stars (J. R. Weaver et al. 2022).

We follow J. R. Weaver et al. (2023b) in limiting our sample to < 26 mag in *Spitzer* IRAC Channel 1 (*Ch. 1*). Based on deeper data available from the Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey (CANDELS; N. A. Grogin et al. 2011; A. M. Koekemoer et al. 2011) in a subset of the COSMOS field, specifically the catalog from H. Nayyeri et al. (2017), J. R. Weaver et al. (2022, 2023b) estimate that the COSMOS2020 detection efficiency[9] at *Ch. 1* = 26 is 75%. The loss of efficiency towards faint *Ch. 1* magnitudes stems from the fact that detection is based on a co-added (as in A. S. Szalay et al. 1999; A. Drlica-Wagner et al. 2018) $izYJHK_S$ image, which does not lead to a significant detection for all sources with $25 < Ch. 1 < 26$ (see also I. Davidzon et al. 2017)[10].

To enable the best use of low signal-to-noise (S/N) data, and to avoid applying a lower flux limit of zero (as was done in A24), we convert the COSMOS2020 catalog fluxes into asinh magnitudes (R. H. Lupton et al. 1999)[11]. These retain the key feature of logarithmic magnitudes (N. Pogson 1856) — a reduced dynamic range between the brightest and faintest sources — while remaining well-behaved mathematically close to a survey's detection limit. All training data take the form of asinh magnitudes and colors. The asinh magnitude corresponding to flux[12] f is defined as

$$ m = -\frac{5}{2} \log_{10}(e) \left[\mathrm{asinh}\left(\frac{f}{2f_b} \right) + \ln\left(\frac{f_b}{f_{\mathrm{AB}}} \right) \right], \quad (1) $$

where $f_{\mathrm{AB}} = 3{,}631$ Jy is the standard reference flux in the AB system (J. B. Oke & J. E. Gunn 1983; M. Fukugita et al. 1996), and f_b is the flux below which asinh magnitudes transition from being logarithmic to linear, referred to as the softening parameter (R. H. Lupton et al. 1999). For reference, the corresponding definition of a logarithmic AB magnitude would be

$$ m = -\frac{5}{2} \log_{10}\left(\frac{f}{f_{\mathrm{AB}}} \right). \quad (2) $$

The value of f_b in 1 can be different for each band and survey; it is typically chosen to be a factor of a few times the background flux uncertainty.

We set the softening parameter f_b by estimating an effective depth from the reported COSMOS flux uncertainties and logarithmic AB magnitudes; the resultant values are tabulated in Table 5 in Appendix A. These are generally fainter than the 3σ depths reported by J. R. Weaver et al. (2022) based on a $3''$ aperture. Our choices of f_b correspond to a transition between linear and logarithmic behavior at a signal-to-noise ratio (S/N) of ~ 5. We do not propagate the COSMOS flux uncertainties into asinh magnitude uncertainties at any stage of our pipeline. When training our uncertainty model (Section 3.4), and performing galaxy-level inference (Section 3.6), we work with the COSMOS flux uncertainties directly.

A total of 423,262 galaxies from COSMOS2020 pass the *Ch. 1* < 26 cut, three times as many as the $r < 25$ training set in A24. This is the sample on which we calibrate our generative model in this work. As in A24, this cut also includes the elimination of objects considered likely stars by J. R. Weaver et al. (2022)[13]. Unlike A24, we apply no further cuts on low flux or S/N objects. We do not omit galaxies that are likely AGN hosts, flagged as having a coincident *Chandra* X-ray detection in the F. Civano et al. (2016) catalog; within the *Ch. 1* < 26 sample there are 1,951 such galaxies ($\sim 0.5\%$ of the catalog). We find 6,407 COSMOS2020 sources with *Ch. 1* $\geq$ 26 but $r < 25$, implying that there is a small MIR-faint but optically bright population that will be missed by our sample selection criterion[14].

2.2. *Spectroscopic Data*

For our update to the T24 SPS parameter and redshift inference for individual galaxies, we validate against the new COSMOS spectroscopic archive from A. A. Khostovan et al. (2025). Of the 97,929 unique objects in A. A. Khostovan et al. (2025), 60,203 have an identified counterpart in the COSMOS2020 Farmer catalog from J. R. Weaver et al. (2022), with 48,291 of these being in our *Ch. 1* < 26 sample. The cumulative numbers of *Ch. 1* < 26 sources assigned a given confidence level (CL) by A. A. Khostovan et al. (2025) are: CL $\geq$ 97%, 16,902 sources; CL $\geq$ 95%, 22,848 sources; CL $\geq$ 80%, 33,368 sources; CL $\geq$ 50%, 39,588 sources. The CL $\geq$ 50% sources correspond to 39,579 unique COSMOS2020 Farmer galaxies; further validation of the cross-match is described in Appendix B.

---

[9] J. R. Weaver et al. (2022, 2023b) find that their photometric extraction routines recover 75% of *Ch. 1* $\approx$ 26 sources in the CANDELS survey area within the COSMOS field.

[10] Detection of sources in the mid-infrared (MIR) alone is inefficient for *Ch. 1* > 22 (see Euclid Collaboration et al. 2022) due to the difficulty of deblending the low-resolution IRAC images without a prior informed by the near-infrared (NIR) data.

[11] See also E. P. Mathews et al. (2023) for a recent motivating application in an SED fitting context.

[12] This is formally a flux density; flux is used here for brevity.

[13] From the COSMOS2020 public catalog, we remove objects with lp_type=1 and lp_type=-99, but not objects with lp_type=2 (X-ray AGN identified in *Chandra*; F. Civano et al. 2016).

[14] These do not all pass the selection criteria used by A24 and T24. Due to the additional flux and S/N cuts applied there, the number of sources with *Ch. 1* $\geq$ 26 is 824 out of 140,745.

3. METHODS

The full `pop-cosmos` pipeline is outlined in Section 3.1, with Sections 3.2–3.5 containing more specific details about some of the components. Parameters and notation are summarised in Tables 1 & 2.

3.1. *Overview of the Generative Model*

The `pop-cosmos` generative model contains the following elements:

1. A population distribution $P(\boldsymbol{\vartheta}|\boldsymbol{\psi})$ over 16 SPS parameters including redshift $\boldsymbol{\vartheta} = (\boldsymbol{\varphi}, z)$. This is a score-based diffusion model (Y. Song et al. 2021) defined by a set of hyperparameters $\boldsymbol{\psi}$. The architecture and initialization of this model are given in Section 3.3.

2. A primary emulator (`Speculator`; J. Alsing et al. 2020) for the Flexible Stellar Population Synthesis library (FSPS; C. Conroy et al. 2009, 2010; C. Conroy & J. E. Gunn 2010). This deterministically predicts model fluxes $\boldsymbol{f}_{\mathrm{SPS}}(\boldsymbol{\vartheta})$ in the 26 COSMOS passbands given a set of input parameters drawn from $P(\boldsymbol{\vartheta}|\boldsymbol{\psi})$. This is described in Section 3.2.

3. A secondary emulator (J. Alsing et al. 2020; B. Leistedt et al. 2023) for the flux contributions of 44 key nebular emission lines $\boldsymbol{F}_{\mathrm{EM}}(\boldsymbol{\vartheta})$.

4. A set of population-wide calibration parameters $(\boldsymbol{\alpha}_{\mathrm{ZP}}, \boldsymbol{\beta}_{\mathrm{EM}}, \boldsymbol{\gamma}_{\mathrm{EM}})$ that model respectively: a zero-point offset in each passband; corrections to the strength of the 44 key nebular emission lines; and variances in the strength of the emission lines.

5. A non-parametric uncertainty model $P(\boldsymbol{\sigma}_P|\boldsymbol{f}, \boldsymbol{\chi})$ that predicts flux uncertainty $\boldsymbol{\sigma}_P$ conditional on the calibrated model fluxes, $\boldsymbol{f} = \boldsymbol{\alpha}_{\mathrm{ZP}} \circ [\boldsymbol{f}_{\mathrm{SPS}}(\boldsymbol{\vartheta}) + \boldsymbol{\beta}_{\mathrm{EM}} \cdot \boldsymbol{F}_{\mathrm{EM}}(\boldsymbol{\vartheta})]$. The uncertainty model is a score-based diffusion model (Y. Song et al. 2021), with hyperparameters $\boldsymbol{\chi}$. This model is defined in Section 3.4.

6. A Student's t error model $P(\boldsymbol{d}|\boldsymbol{f}, \boldsymbol{\sigma})$ that predicts observable fluxes given model fluxes $\boldsymbol{f}$ and a total uncertainty $\boldsymbol{\sigma} = \sqrt{\boldsymbol{\sigma}_P^2 + \boldsymbol{\sigma}_{\mathrm{EM}}^2}$, where $\boldsymbol{\sigma}_{\mathrm{EM}} = [\boldsymbol{\gamma}_{\mathrm{EM}} \circ (1 + \boldsymbol{\beta}_{\mathrm{EM}})] \cdot \boldsymbol{F}_{\mathrm{EM}}(\boldsymbol{\vartheta})$, and $\circ$ denotes an element-wise product.

In our main training loop (Section 3.5), we simultaneously optimize the population level parameters $(\boldsymbol{\psi}, \boldsymbol{\alpha}_{\mathrm{ZP}}, \boldsymbol{\beta}_{\mathrm{EM}}, \boldsymbol{\gamma}_{\mathrm{EM}})$ by comparing the mock photometry to the COSMOS2020 photometry.

3.2. *Emulated Stellar Population Synthesis*

Our 16-parameter SPS model largely follows the configuration used in A24, which is in turn based on

Table 1. Summary of main notation.

Symbol	Description
	Population-level quantities
$\boldsymbol{\psi}$	Population model hyperparameters
$\boldsymbol{\chi}$	Uncertainty model hyperparameters
$\boldsymbol{\alpha}_{\mathrm{ZP}}$	Fractional zero-point corrections
$\boldsymbol{\beta}_{\mathrm{EM}}$	Fractional emission line corrections
$\boldsymbol{\gamma}_{\mathrm{EM}}$	Fractional emission line std. devs.
	Galaxy-level quantities
$\boldsymbol{d}$	Noisy model fluxes
$\boldsymbol{f}$	Calibrated model fluxes
$\boldsymbol{m}$	Calibrated model asinh magnitudes
$\boldsymbol{f}_{\mathrm{SPS}}(\boldsymbol{\vartheta})$	SPS model fluxes
$\boldsymbol{F}_{\mathrm{EM}}(\boldsymbol{\vartheta})$	Emission line model fluxes
$\boldsymbol{\sigma}$	Model flux uncertainties
$\boldsymbol{\sigma}_P$	Model photometric uncertainty
$\boldsymbol{\sigma}_{\mathrm{EM}}$	Emission line uncertainty
$\boldsymbol{\vartheta}$	Galaxy level parameters $(\boldsymbol{\varphi}, z)$
$\boldsymbol{\varphi}$	Non-redshift SPS parameters
	Photometry-related
f	Flux in the AB system
f_{AB}	Flux of the AB standard source
f_b	Flux softening parameter
m	asinh magnitude
	Diffusion-related
$\boldsymbol{w}(t)$	Brownian motion process
$\boldsymbol{x}(t)$	State vector at time t
t	Time index in the diffusion process
T	Maximum time, $\boldsymbol{x}(T) \sim$ base density
$p_t(\boldsymbol{x})$	Marginal density of $\boldsymbol{x}(t)$ at time t
$\nabla_{\boldsymbol{x}} p_t(\boldsymbol{x})$	Score at time t
$\boldsymbol{s}_{\boldsymbol{\Theta}}(\boldsymbol{x}, t)$	Score network with hyperparameters $\boldsymbol{\Theta}$
$\beta(t)$	Scaling factor in VP SDE
$\beta_{\min}, \beta_{\max}$	Minimum/maximum scaling factor

Prospector-α model (J. Leja et al. 2017; B. D. Johnson et al. 2021a) and `v3.2` of FSPS (C. Conroy et al. 2009, 2010; C. Conroy & J. E. Gunn 2010). The free parameters that describe a galaxy are given in Table 2. The core ingredients of the FSPS model are: the G. Chabrier (2003) initial mass function (IMF); MILES stellar templates (P. Sánchez-Blázquez et al. 2006; J. Falcón-Barroso et al. 2011); isochrones from the MESA Isochrones and Stellar Tracks library (MIST;

Table 2. Summary of SPS parameters.

Symbol / Unit	Lower Bound	Upper Bound	Description
			Base parameters
$\log_{10}(M_{\rm form}/M_\odot)$	-	-	Logarithm of stellar mass formed
$\log_{10}(Z/Z_\odot)$	-1.98	0.19	Logarithm of stellar metallicity
$\Delta\log_{10}({\rm SFR})_{\{1:6\}}$	-5.0	5.0	Ratios of SFR between adjacent bins of SFH
$\tau_2/{\rm mag}$	0.0	4.0	Diffuse dust optical depth
n	-1.0	0.4	Power-law index for diffuse dust attenuation law
τ_1/τ_2	0.0	2.0	Birth cloud dust optical depth relative to τ_2
$\ln(f_{\rm AGN})$	$-5\ln(10)$	$\ln(3)$	Logarithm of AGN bolometric luminosity fraction
$\ln(\tau_{\rm AGN})$	$\ln(5)$	$\ln(150)$	Logarithm of AGN torus optical depth
$\log_{10}(Z_{\rm gas}/Z_\odot)$	-2.0	0.5	Logarithm of gas-phase metallicity
$\log_{10}(U_{\rm gas})$	-4.0	-1.0	Logarithm of gas ionization, $U_{\rm gas} = \frac{\text{photon density}}{\text{H density}}$
z	0.0	6.0	Redshift
			Derived quantities
$t_{\rm age}/{\rm Gyr}$	-	-	Mass-weighted age
$\log_{10}(M/M_\odot)$	-	-	Logarithm of stellar mass remaining
$\log_{10}({\rm SFR}/M_\odot\,{\rm yr}^{-1})$	-	-	Logarithm of SFR
$\log_{10}({\rm sSFR}/{\rm yr}^{-1})$	-	-	Logarithm of specific SFR per unit mass remaining

A. Dotter 2016; J. Choi et al. 2016); the B. T. Draine & A. Li (2007) polycyclic aromatic hydrocarbon (PAH) model for dust emission[15]; the M. Nenkova et al. (2008a,b) models for the dusty torus around active galactic nuclei (see J. Leja et al. 2018); and the P. Madau (1995) model for extinction by the intergalactic medium. We use a 7-bin piecewise constant star-formation history (SFH; see J. Leja et al. 2019a). MIR emission due to dust shells around asymptotic giant branch (AGB) stars is included (A. Villaume et al. 2015).

Our base nebular emission model uses the standard FSPS treatment (N. Byler et al. 2017), computed using the `Cloudy` photoionization code (G. J. Ferland et al. 2013) for the MIST stellar evolutionary tracks. We use the (recommended) dust-free configuration from N. Byler et al. (2017), based on a stellar ionizing spectrum within a spherical shell of pure gas. Dust attenuation is included separately using a modified D. Calzetti et al. (2000) dust law (S. Noll et al. 2009) with the

M. Kriek & C. Conroy (2013) prescription for the UV bump strength. An additional birth-cloud attenuation component is included for stars younger than 10 Myr (S. Charlot & S. M. Fall 2000), with attenuation proportional to λ^{-1}.

We use `Prospector` (B. D. Johnson et al. 2021a), `python-fsps` (B. D. Johnson et al. 2021b), and `sedpy` (B. D. Johnson 2021) to generate FSPS photometry in 26 passbands for 6.4 million galaxies using the configuration described above, with the 16 SPS parameters ($\boldsymbol{\vartheta}$) sampled randomly for each galaxy within the ranges given in Table 2 (where the ranges for the non-redshift parameters, $\boldsymbol{\varphi}$, are the same as in A24). To ensure coverage of the deeper IRAC-selected sample, we allow $z \in [0,6]$, compared to $z \in [0, 4.5]$ for the r-band-selected sample used in A24. Using the generated FSPS photometry, we train an emulator using `Speculator` (J. Alsing et al. 2020) to predict magnitude as a function of the input parameters $\boldsymbol{\vartheta} = (\boldsymbol{\varphi}, z)$. The emulator is a four-layer dense neural network with 128 units per layer, and the bespoke activation function introduced by J. Alsing et al. (2020). Training uses the `Adam` optimizer (D. P. Kingma & J. Ba 2015) to minimize the mean squared error between predicted and true magnitude. Unlike in J. Alsing et al. (2020, 2024), where this loss was computed for absolute magnitudes per unit mass, here we use apparent asinh magnitudes (R. H. Lupton et al. 1999) to avoid giving excessive weight to parts of parameter space that would be too faint ($\gtrsim 30$ mag) to observe.

---

[15] We use this model with fixed parameters, $U_{\min} = 1$, $\gamma_e = 1\%$, and $Q_{\rm PAH} = 2\%$. The choice of $U_{\min} = 1$ and $\gamma_e = 1\%$ corresponds to a minimum starlight intensity of 1, which $1 - \gamma_e = 99\%$ of dust mass is exposed to; the remaining $\gamma_e = 1\%$ of dust mass will be exposed to a power law distribution of intensities $> U_{\min}$ (as in D. A. Dale & G. Helou 2002). The choice of $Q_{\rm PAH} = 2\%$ corresponds to a dust model with 2% of total dust mass in PAHs. As our reddest data (IRAC *Ch. 2*) covers $\sim$ 4–5 µm in the observer frame, we expect to be relatively insensitive to these choices, as these features would be strongest further into the IR (see J. Leja et al. 2017).

In addition to the photometry emulator, we also use a secondary `Speculator`-based emulator to compute the strengths of the 44 nebular emission lines whose strengths we modify relative to the N. Byler et al. 2017 `Cloudy` model grids. We set up this emulator following the procedure detailed in B. Leistedt et al. (2023, §3.4.2; hereafter L23), modeling the emission lines as delta functions, and the COSMOS photometric bandpasses as generalized Gaussian mixture models.

3.3. *Population Model Initialization*

In A24, the population model was a score-based diffusion model (Y. Song et al. 2021). The free parameters in this model are the weights and biases, ψ, of a score network, $s_\psi(x, t)$, which approximates a score function $\nabla_x p_t(x)$. The role of the score function is to define a continuous-time transform of a vector $x(t)$ between a base density $p_T(x)$ at time $t = T$, and a target density $p_0(x)$ at time $t = 0$. In `pop-cosmos`, the target density at $t = 0$ is the population distribution over 16 SPS parameters, $P(\vartheta | \psi)$.

We use a diffusion model based on a variance-exploding stochastic differential equation (VE SDE; Y. Song et al. 2021), with a unit Gaussian base density as in A24. The VE SDE can be thought of as a continuous-time version of score matching with Langevin dynamics (Y. Song & S. Ermon 2019). Our score model $s_\psi(x, t)$ is a four-layer dense neural network with 128 units per layer, and sigmoid linear unit (SiLU) as an activation function. We have found that the initialization and predictive performance are comparable when using a variance-preserving SDE (VP SDE) as the basis of the diffusion model, but that the VE SDE is faster to train and sample from.

We initialize the model based on a set of maximum a posteriori (MAP) estimates of ϑ, obtained under the Prospector-α prior (J. Leja et al. 2017) for the COSMOS2020 `Farmer` catalog (J. R. Weaver et al. 2022). From this distribution of MAP estimates, ψ is optimized based on the denoising score-matching loss function (A. Hyvärinen 2005; P. Vincent 2011; Y. Song & S. Ermon 2019; Y. Song et al. 2021).

3.4. *Non-parametric Uncertainty Model*

In A24, the uncertainty model was a single-component mixture density network (MDN; C. M. Bishop 2006) of the form $P(\sigma_P | f, \tilde{\chi}) = \mathcal{N}(\sigma_P | \mu(f), \Sigma(f))$, where μ and Σ were small dense neural networks with weights and biases $\tilde{\chi}$. The use of a single-component model was motivated by the need for something that can be represented by a bijective transform of a simple base density. Such a representation is necessary for gradient-based optimization of `pop-cosmos`. However, such a model is unable to capture multimodal distributions of σ_P given f. Such multimodality is most pronounced for photometric bands that do not have uniform depth across the COSMOS field: u-band (including data from M. Sawicki

et al. 2019 with a new u filter, and data from C. Laigle et al. 2016 with an older filter); $YJHK_S$ bands (due to the four ultra-deep stripes in UltraVISTA, detailed in H. J. McCracken et al. 2012); and the *Spitzer* IRAC channels (which include deeper regions tracking the UltraVISTA ultra-deep stripes from M. L. N. Ashby et al. 2018).

In this work, we use a score-based diffusion model (Y. Song et al. 2021) to replace the MDN used in A24. Specifically, we set up an uncertainty model of the form $P(\ln \sigma_P | m, \chi)$ that predicts the logarithm of flux uncertainty conditional on calibrated model magnitudes m. The free parameters χ define a conditional score network $s_\chi(x, t; m)$ that approximates the time-dependent score of the diffusion process for a given set of input magnitudes m. Our score network in this case is a five-layer densely connected neural network with 256 units per layer and tanh activation functions. The diffusion process follows a VP SDE,

$$d x = -\frac{1}{2}\beta(t) x \, dt + \sqrt{\beta(t)} \, dw, \qquad (3)$$

where w is a Brownian motion process and $\beta(t) = \beta_{\min} + \frac{t}{T}(\beta_{\max} - \beta_{\min})$. We use $\beta_{\min=0.1}$ and $\beta_{\max} = 20.0$ (J. Ho et al. 2020; Y. Song et al. 2021), and sample $x(t = T)$ from a Gaussian base density with unit variance. The VP SDE can be viewed as a continuous generalization of a denoising diffusion probabilistic model (J. Sohl-Dickstein et al. 2015; J. Ho et al. 2020; J. Song et al. 2021; Y. Song et al. 2021).

We fit the uncertainty model to the reported magnitudes and flux errors from the COSMOS2020 `Farmer` catalog (J. R. Weaver et al. 2022), by minimizing a denoising score-matching loss function (A. Hyvärinen 2005; P. Vincent 2011; Y. Song & S. Ermon 2019; Y. Song et al. 2021) using the `Adam` optimizer (D. P. Kingma & J. Ba 2015). In the main population model training loop (Section 3.5), we fix χ.

3.5. *Population Model Training*

We use a training procedure similar to A24, minimizing a loss that measures similarity between the observed COSMOS2020 photometry and the model photometry produced by the generative model. Our loss function is based on a set of summary statistics of the color and magnitude distributions. From the forward simulated model magnitudes, and observed COSMOS magnitudes, we compute 11 broadband and 13 narrow-band colors from all adjacent pairs of magnitudes. This gives $26 + 11 + 13 = 50$ colors and magnitudes for each model or observed galaxy. We use these colors and magnitudes to construct a vector of summary statistics containing: the mean of each color and magnitude; the standard deviation of each color and magnitude; the full set of color–color, magnitude–magnitude, and color–magnitude correlation coefficients; and a set of 19 percentiles (5th,

10th, ..., 95th) of each color and magnitude distribution. Our loss function is then the mean squared error across the collected set summary statistics. We arrived at this set of summaries through numerical experiments, finding that these gave the best predictive performance. In future work, it would be worthwhile to formulate a compressed set of summaries, e.g., using a principal component analysis (S. Thorp et al. 2025b), or by making use of non-linear correlations identified using mutual information (e.g. N. Chartab et al. 2023).

We minimize our loss function using the `Adam` optimizer (D. P. Kingma & J. Ba 2015). We optimize the parameters $(\psi, \boldsymbol{\alpha}_{ZP}, \boldsymbol{\beta}_{EM}, \boldsymbol{\gamma}_{EM})$ simultaneously for 120 iterations. We initialize the calibration parameters $(\boldsymbol{\alpha}_{ZP}, \boldsymbol{\beta}_{EM}, \boldsymbol{\gamma}_{EM})$ to the A24 values.

3.6. *Individual-Galaxy Inference*

As demonstrated in T24, the trained population model can be used as a prior over the 16 SPS parameters in downstream inference tasks; for example SED fitting for individual galaxies. Using an empirically calibrated population distribution for this task breaks degeneracies and leads to more accurate parameter constraints (see T24; or B. Wang et al. 2023b for a different approach to the construction of a more informative prior in SPS parameter space; see also C. K. Jespersen et al. 2025a for further motivation). Since the population distribution over SPS parameters in `pop-cosmos` is a score-based diffusion model (Section 3.3), it is possible to evaluate the prior probability of a set of SPS parameters as described in T24 (for further background, see W. Grathwohl et al. 2018; R. T. Q. Chen et al. 2018; Y. Song et al. 2021).

Using the trained model as a prior in this way, we perform posterior inference of the SPS parameters for the 423,262 COSMOS2020 galaxies with *Ch. 1* < 26, as well as 6,407 galaxies with *Ch. 1* ≥ 26 but *r* < 25 that were included in the A24 training sample. As a baseline, we also perform this inference under a broader, `Prospector`-α-like prior (J. Leja et al. 2017), as was done in T24. Under both priors, we compute model fluxes using the `Speculator` emulator (J. Alsing et al. 2020) described in Section 3.2 (including zero-point and emission line strength corrections set using the $\boldsymbol{\alpha}_{ZP}$ and $\boldsymbol{\beta}_{EM}$ values fitted in Section 3.5). Our likelihood is a Student's *t*-distribution (L23) in flux, with a scale parameter that includes a quadrature sum of the reported COSMOS2020 flux errors from J. R. Weaver et al. (2022), the emission line variances ($\boldsymbol{\gamma}_{EM}$) fitted in Section 3.5, and the fractional error floors estimated in L23. As we work directly with the reported fluxes at this stage, we do not need to propagate the flux uncertainty to an asinh magnitude uncertainty.

Under both priors, we generate ∼ 5,000 independent posterior samples for each galaxy using a bespoke batched and GPU-accelerated affine-invariant MCMC sampler, `affine`[16], implemented in PyTorch (based on the algorithm from J. Goodman & J. Weare 2010; D. Foreman-Mackey et al. 2013). We run the sampler with two ensembles of 256 walkers, using the parallel stretch move from D. Foreman-Mackey et al. (2013), for 1000 iterations. This yields 512,000 posterior samples per galaxy, which we thin by a factor of 100. The MCMC is initialized using maximum a posteriori (MAP) estimates obtained using the `Adam` optimizer. Under the `Prospector`-α-like prior, we batch computations over 2,000 galaxies simultaneously, with a single batch executing in ∼ 23 min on an NVIDIA A100 GPU card. Under the `pop-cosmos` prior we batch over 500 galaxies, with a single batch taking ∼ 82 mins to complete on an A100 GPU. This gives an average throughput of 0.7 GPU-sec per galaxy under the `Prospector` prior, or < 10 GPU-sec per galaxy under the `pop-cosmos` prior. The total processing time for the 429,669 galaxies was around 1,000 GPU-hrs.

3.7. *Mock Catalog Generation*

For all subsequent analyses in the paper, we use two mock galaxy catalogs (one *r*-band limited at 25 mag; one *Ch. 1*-limited at 26 mag) that we have made publicly available on Zenodo. For each of these catalogs, we generated galaxies from the forward model until 2 million galaxies were produced are above the relevant magnitude limit (based on their noisy model magnitudes). For the *Ch. 1* < 26 catalog, this process takes around 2 GPU-hrs on a single NVIDIA A100 card.

Each galaxy in the mock catalogs has 16 SPS parameters and 26-band COSMOS-like photometry associated with it. When we show photometric predictions from these catalogs, we show logarithmic magnitudes[17]. In the public release we include fluxes, logarithmic magnitudes, and asinh magnitudes. We include both noiseless and noisy model fluxes for each galaxy, as well as the model photometric uncertainties that were applied to obtain the latter from the former. The magnitudes in our plots and in the public catalogs are noisy magnitudes, obtained by applying either Equation 1 or 2 to the noisy model fluxes. We do not propagate the model flux uncertainties to magnitude uncertainties, or report noiseless model magnitudes (although these can be easily obtained from the noiseless model fluxes).

From the 16 base SPS parameters (listed in the upper part of Table 2), we compute several derived quantities that are used in our downstream analyses and included in the public catalogs; these are listed in the lower part of Table 2. We compute the stellar mass remaining using

---

[16] https://github.com/justinalsing/affine/tree/torch

[17] Although we carry out our fits using asinh magnitudes, we show predictions in logarithmic magnitudes for consistency with A24 and other literature. Above the effective depths listed in Table 5, the two systems are almost equivalent (see also R. H. Lupton et al. 1999).

a neural network emulator that approximates the fraction of surviving stellar mass (calculated by FSPS) as a function of the 16 base parameters[18]. The (s)SFR and mass weighted age are computed directly using the star-formation rate ratios, $\Delta \log_{10}(\text{SFR})_{1:6}$, with the (s)SFR being averaged over the last 100 Myr of a galaxy's evolution (corresponding to the two most recent bins of the SFH; J. Leja et al. 2019a). A full description is given in Appendix C.

4. VALIDATION

Given the algorithmic and numerical complexity of the `pop-cosmos` model described in Section 3, it is critical to carefully validate all the different aspects of the model's performance. In Section 4.1, we show the colors and magnitudes predicted by the `pop-cosmos` model, compared to those in the J. R. Weaver et al. (2022) COSMOS2020 `Farmer` catalog. Section 4.2 demonstrates the performance of the uncertainty model relative to the COSMOS2020 flux errors reported by J. R. Weaver et al. (2022). Sections 4.3–4.5 validate the inferences made for individual galaxies under the new `pop-cosmos` prior, compared to the COSMOS spectroscopic archive (A. A. Khostovan et al. 2025), the results under a `Prospector`-like prior, and the `LePhare` results from J. R. Weaver et al. (2022). Finally, in Section 4.6, we examine the stellar mass vs. redshift relation predicted by the model to estimate the redshift-evolving mass completeness limit of our predictions. Supplementary results are included in Appendices D–F.

4.1. Colors and Magnitudes for Ch. 1 < 26

In Figure 1 and Figure 2, we show the broadband color and magnitude marginals for the $Ch.\,1 < 26$ mock galaxy sample, compared to COSMOS2020 (J. R. Weaver et al. 2022). There are some small offsets in the magnitude distributions. The most prominent of these is for $Ch.\,1$, where the turnover in number counts at $Ch.\,1 > 25$ is not well modeled. As discussed in Section 2.1 here (and I. Davidzon et al. 2017; J. R. Weaver et al. 2022, 2023b), there is likely a degree of unmodeled incompleteness at these faint magnitudes.

The marginal distributions of colors in Figure 2 show generally very good agreement. The most substantial offsets are seen in $g-r$ and $K_S - Ch.\,1$. The $K_S - Ch.\,1$ color is known to be particularly sensitive to the photometric extraction (J. R. Weaver et al. 2022), and exhibits some of the largest offsets between the `Farmer` J. R. Weaver et al. (2023a) profile fitting photometry and the `Classic` `SExtractor`-based (E. Bertin & S. Arnouts 1996) aperture photometry pipeline (comparable to COSMOS2015; C. Laigle et al. 2016), the latter

of which tends to be systematically redder. The u- and g-bands, and $g-r$ color at the faint end are also known to be sensitive to the photometric extraction in COSMOS2020 (J. R. Weaver et al. 2022). We have found that even with a broader `Prospector`-α-like prior, it is difficult to produce $K_S - Ch.\,1 < -1$ using our current SPS configuration (described in Section 3.2).

We show color vs. color plots in Figure 3, for the set of color pairs used by A24. These color choices are motivated by J. R. Weaver et al. (2022), who identified them as the most significantly informative COSMOS colors for galaxy evolution. We perform a detailed 9-dimensional validation of the broadband colors for $Ch.\,1 < 26$ in Appendix D (following the scheme from S. Thorp et al. 2025b). This validation includes a comparison based on quantile–quantile (Q–Q) and probability–probability (P–P) tests (M. B. Wilk & R. Gnanadesikan 1968), as well as statistical summaries in the form of Wasserstein distances (L. Kantorovich & G. S. Rubinstein 1958; L. N. Vaserstein 1969) and the Kolmogorov–Smirnov (K–S) statistic (A. Kolmogorov 1933; N. Smirnov 1948). These tests confirm that the quantitative agreement between the model predictions and data is comparable to agreement between the two COSMOS2020 photometric reductions, `Classic` and `Farmer`. We show plots of colors, magnitudes, and flux uncertainties for the $r < 25$ selection in Appendix F.

4.2. Uncertainty Model

We now inspect the trained uncertainty model, introduced in Section 3.4, by using it to generate flux uncertainties conditional on the `pop-cosmos` model magnitudes. Figure 4 shows the joint distribution of flux uncertainties and magnitudes drawn from the trained `pop-cosmos` model, compared to the reported magnitudes and flux errors from the COSMOS2020 `Farmer` catalog (J. R. Weaver et al. 2022). The agreement here is good by construction, as the uncertainty model is trained directly on the `Farmer` catalog, with any visible offsets arising due to the offsets in the magnitude distributions for this catalog (see Figure 1). The bimodal distributions of flux errors resulting from the mixed depths in the $uYJHK_S$ bands are successfully captured by the trained uncertainty model.

4.3. Individual Galaxy Inference: Redshifts

We obtain full posterior distributions for the properties of each galaxy, including redshift. For each galaxy, we take its posterior median redshift as a point estimate of z^{phot} when comparing to the z^{spec} collated by A. A. Khostovan et al. (2025). For the 39,588 objects with spectroscopic cross-matches in A. A. Khostovan et al. (2025) we follow T24 and L23 by characterising the mismatch using $\Delta_z = (z^{\text{phot}} - z^{\text{spec}})/(1 + z^{\text{spec}})$. We define an outlier as a galaxy with $|\Delta_z| > 0.15$, following the widely used convention in the literature. We find a total of 2,844 outliers (7.18% of all galaxies).

---

[18] For our parametrization, the fraction of mass remaining is set by the star-forming history and stellar metallicity; i.e. the salient parameters are z, $\Delta \log_{10}(\text{SFR})_{\{1:6\}}$, and $\log_{10}(Z/Z_\odot)$.
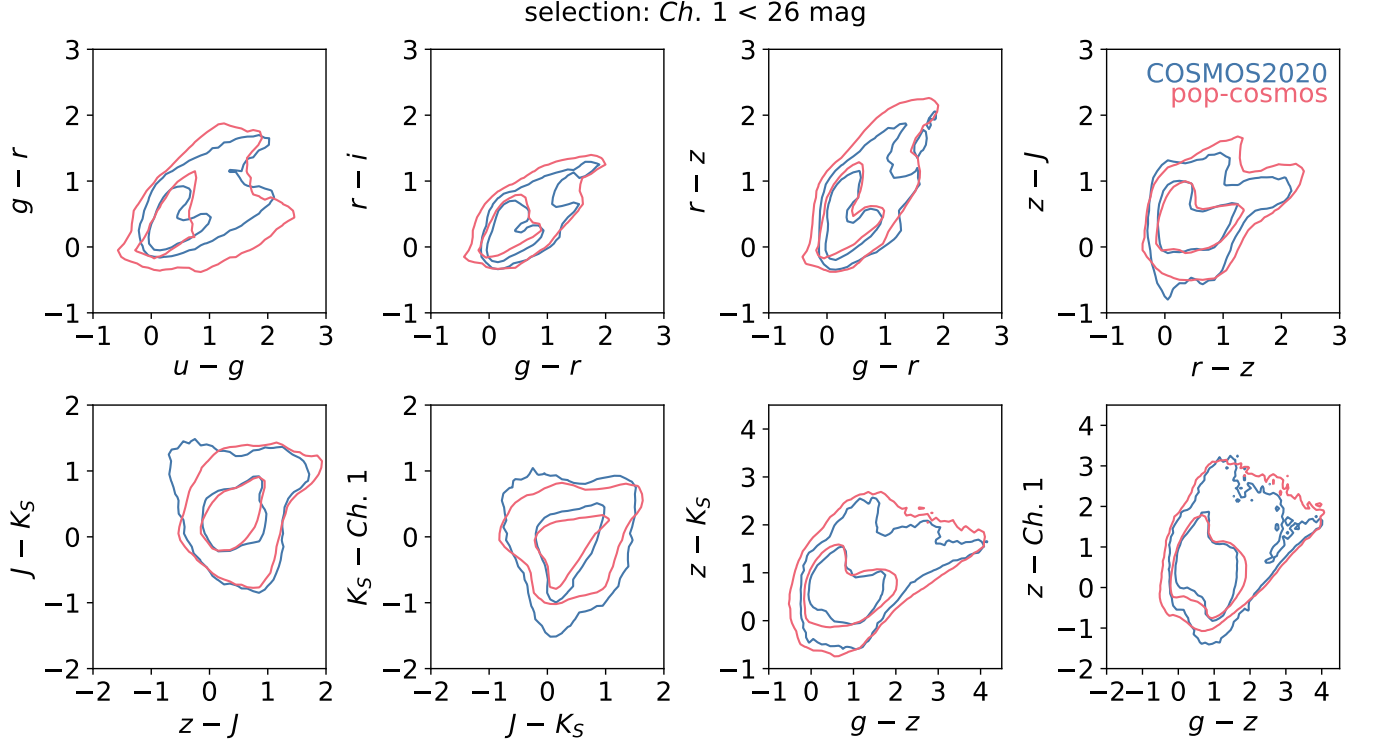
selection: *Ch.* 1 < 26 mag
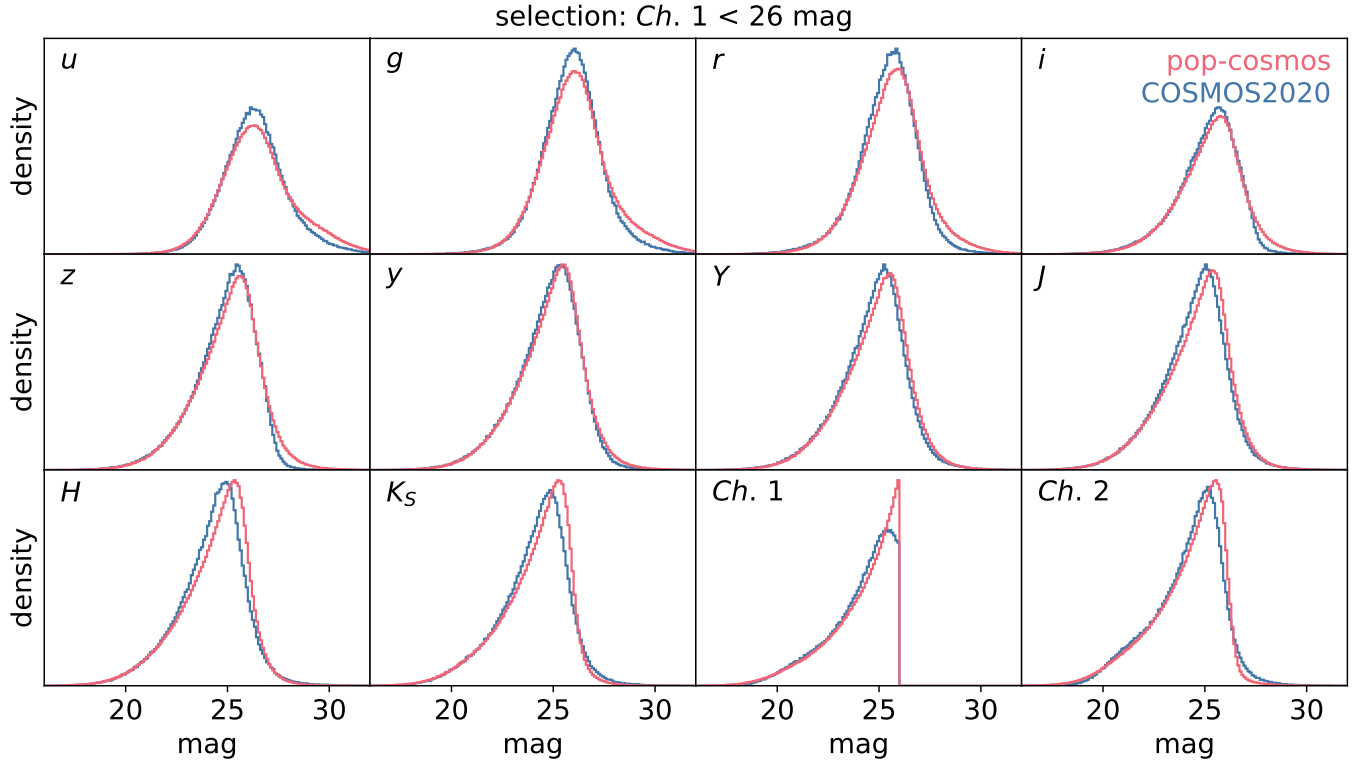
Figure 1. Broadband magnitudes (logarithmic) predicted by the new `pop-cosmos` generative model, compared to COSMOS2020 (J. R. Weaver et al. 2022) for *Ch. 1* < 26.

selection: *Ch.* 1 < 26 mag
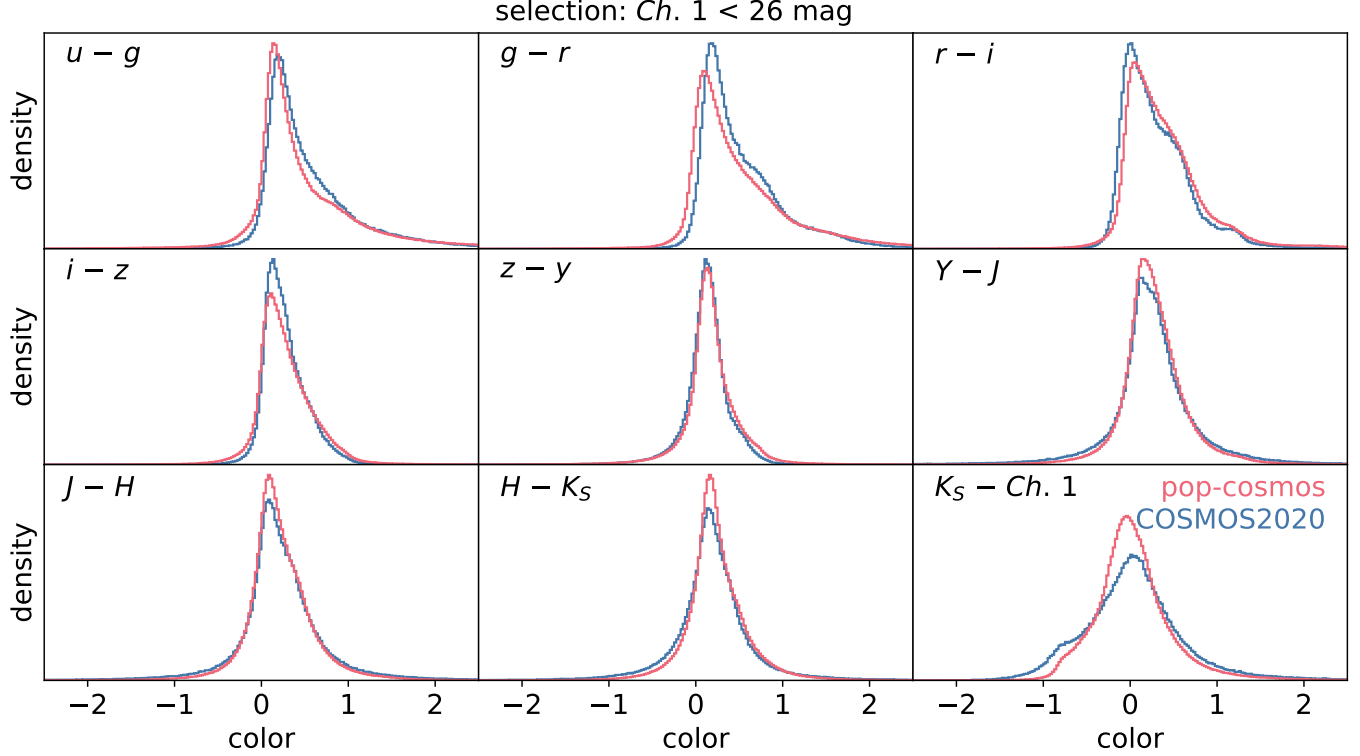
Figure 2. Broadband colors (logarithmic, adjacent bands) predicted by the new `pop-cosmos` generative model, compared to COSMOS2020 (J. R. Weaver et al. 2022) for *Ch. 1* < 26.

selection: *Ch.* 1 < 26 mag

Figure 3. Broadband colors (logarithmic, A24 band combinations) predicted by the new `pop-cosmos` generative model, compared to COSMOS2020 (J. R. Weaver et al. 2022) for *Ch. 1* < 26. Contours enclose the 68 and 95% highest density regions.

selection: *Ch.* 1 < 26 mag
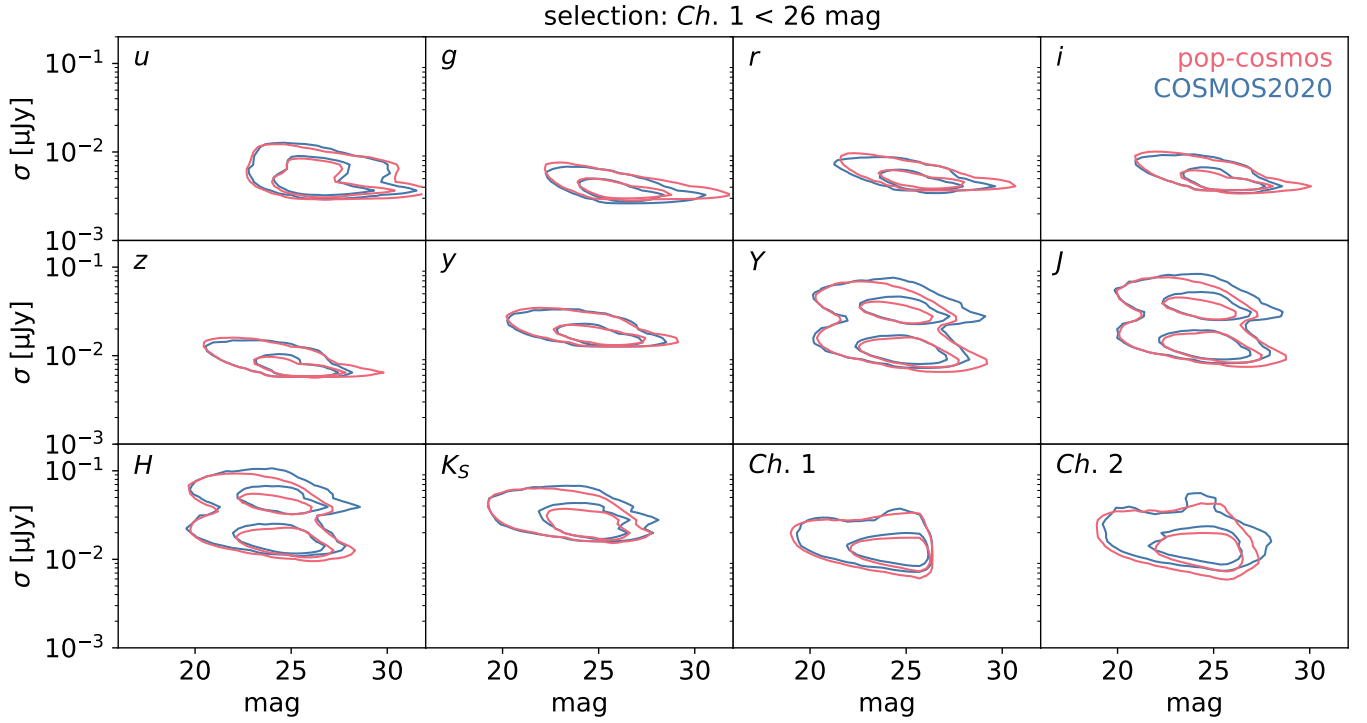
Figure 4. Flux uncertainty vs. (logarithmic) magnitude for the COSMOS broad bands, compared to the flux errors reported in the COSMOS2020 catalog (J. R. Weaver et al. 2022) for *Ch. 1* < 26. Note that the `pop-cosmos` uncertainty predictions are made conditional on the `pop-cosmos` model magnitudes. Contours enclose the 68 and 95% highest probability density regions.
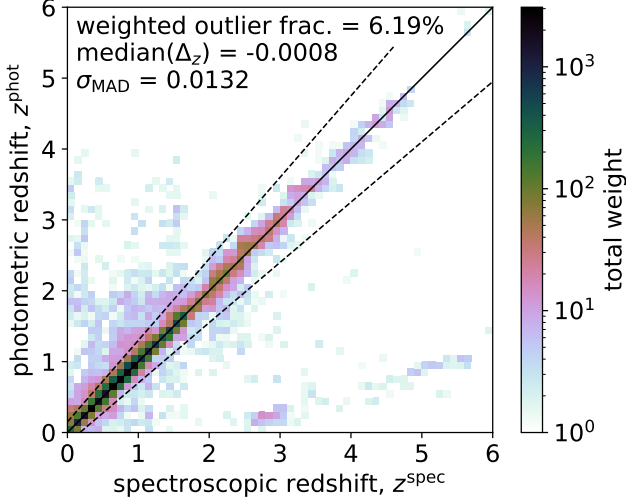
Figure 5. Comparison of z^{phot} vs. z^{spec} based on z^{phot} (posterior median z) inferred under the new `pop-cosmos` prior. Each galaxy is weighted by its spectroscopic confidence level from A. A. Khostovan et al. (2025). Solid and dashed lines show $z^{\mathrm{phot}} = z^{\mathrm{spec}}$ and $|\Delta_z| = 0.15$, respectively.

Table 3. Statistics on $\Delta_z = (z^{\mathrm{phot}} - z^{\mathrm{spec}})/(1 + z^{\mathrm{spec}})$. Outlier rate is weighted by spectroscopic confidence level (A. A. Khostovan et al. 2025).

Model	Median	σ_{MAD}	Outlier Rate (%)
`pop-cosmos`	-8×10^{-4}	0.0132	6.19%
`Prospector`	-3×10^{-4}	0.0148	6.57%
`LePhare`	-2×10^{-4}	0.0164	8.44%
`EAZY`	-4×10^{-4}	0.0132	6.84%

We find $\sigma_{\mathrm{MAD}} = 1.48 \times \mathrm{median}(|\Delta_z|) = 0.0132$, and $\mathrm{median}(\Delta_z) = -8 \times 10^{-4}$. Despite the new spectroscopic sample being a factor of $3.2\times$ larger than used in T24, and the redshift range being considerably larger, the outlier fraction here is comparable; the median and σ_{MAD} are both improved.

In Figure 5, we show a 2D histogram of z^{phot} vs. z^{spec}. In the histogram, we weight each galaxy based on the standardized z^{spec} confidence level assigned by A. A. Khostovan et al. (2025) (so a galaxy with CL = 50% will have weight 0.5, and so forth). The 39,588 galaxies have a total weight of 33,616.55 in this scheme. We show the weighted outlier fraction (6.19%) on Figure 5, defining this as the sum of weights of galaxies with $|\Delta_z| > 0.15$. We summarize the results for the `pop-cosmos` and `Prospector` priors in Table 3. We also show the results for `LePhare` (S. Arnouts et al. 1999; O. Ilbert et al. 2006) and `EAZY` (G. B. Brammer et al.
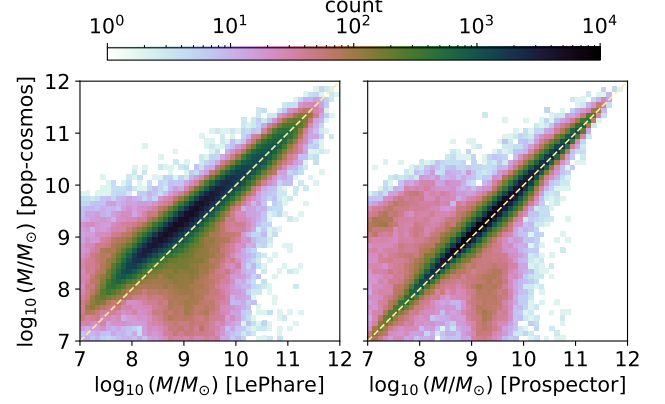
Figure 6. Comparison of stellar mass estimates for the 423,262 COSMOS2020 galaxies with $Ch.\,1 < 26$. The vertical axis shows the posterior median $\log_{10}(M/M_\odot)$ under the new `pop-cosmos` prior. The horizontal axis shows **(left)** the `LePhare` stellar mass estimates from J. R. Weaver et al. (2022), and **(right)** the posterior median under the `Prospector` prior.

2008) on the same spectroscopic sample, based on the photometric redshifts from J. R. Weaver et al. (2022).

4.4. *Individual Galaxy Inference: Stellar Masses*

In Figure 6, we compare the posterior median stellar masses estimated under the `pop-cosmos` prior to those estimated under the `Prospector` prior, and to the `LePhare` (S. Arnouts et al. 1999; O. Ilbert et al. 2006) stellar masses published by J. R. Weaver et al. (2022). We make this comparison for the full $Ch.\,1 < 26$ sample of 423,262 galaxies. We see that for $\log_{10}(M/M_\odot) \gtrsim 10$, the results under the new `pop-cosmos` prior agree closely with the `Prospector` prior and with `LePhare`. For lower masses, there is more scatter, and the `pop-cosmos` prior gives rise to slightly higher stellar mass estimates on average than `LePhare`. In the J. R. Weaver et al. (2022) `LePhare` fits, the template library used (from M. Polletta et al. 2007; O. Ilbert et al. 2009; M. Onodera et al. 2012, using SPS models from L. Silva et al. 1998; G. Bruzual & S. Charlot 2003) is based on a finite set of parametric SFH models[19]. Several works (e.g., J. Leja et al. 2019b; S. Lower et al. 2020) have shown that parametric SFH models tend to produce systematically lower stellar mass estimates than non-parametric models, likely accounting for the offset between our inferences and `LePhare`. That the difference is largest at

---

[19] The 19 elliptical and spiral galaxy templates based on M. Polletta et al. (2007) use a power-law SFR(z) using the L. Silva et al. (1998) `GRASIL` model. The additional star-forming galaxy templates from O. Ilbert et al. (2009) are based on 12 G. Bruzual & S. Charlot (2003) models that incorporate a smooth SFH with bursts. The 2 M. Onodera et al. (2012) elliptical galaxy templates use exponentially declining SFH models from G. Bruzual & S. Charlot (2003).

Table 4. Statistics on $\Delta \log_{10}(M/M_\odot)$ between the `pop-cosmos` and `Prospector` inferences. Galaxies are sorted into mass bins based on their `pop-cosmos` mass estimates.

Mass Bin	Count	Median	σ_{MAD}	Outlier Rate (%)	
				0.2 dex	0.5 dex
$[7, 12)$	421630	0.037	0.102	15.7	5.9
$[7, 8)$	16422	-0.077	0.293	49.8	36.6
$[8, 9)$	118585	0.049	0.113	17.7	7.3
$[9, 10)$	211883	0.032	0.096	13.2	4.0
$[10, 11)$	68494	0.039	0.088	12.2	2.2
$[11, 12)$	6246	0.039	0.074	12.1	2.4

Figure 7. Distribution of AGN bolometric luminosity fraction f_{AGN}, for galaxies with $Ch.\,1 < 26$. The filled red histogram shows the prior distribution computed from `pop-cosmos` model draws. The purple dashed histogram shows the posterior medians for 423,262 COSMOS2020 galaxies, with the cyan histogram showing posterior medians for the subset of 1,951 COSMOS2020 galaxies with *Chandra* X-ray detections.

low masses aligns with the results seen in J. Leja et al. (2019b).

Table 4 contains summaries of the difference in $\log_{10}(M/M_\odot)$ inferred under the `pop-cosmos` and `Prospector` priors. The quoted statistics are similar to those used in the spectroscopic redshift comparison: median$[\Delta \log_{10}(M/M_\odot)]$, $\sigma_{\mathrm{MAD}} = 1.48 \times$ median$[|\Delta \log_{10}(M/M_\odot)|]$, and the fraction of 0.2 dex and 0.5 dex outliers (the latter motivated by B. Mobasher et al. 2015). The logarithmic mass residuals are computed for `pop-cosmos` minus `Prospector`; i.e. a positive median$[\Delta \log_{10}(M/M_\odot)]$ would imply systematically higher masses under the `pop-cosmos` prior. Galaxies are sorted into the mass bins based on their inferred mass under the `pop-cosmos` prior. Across the full sample, the median offset in the inferred stellar mass is < 0.04 dex, with $\sigma_{\mathrm{MAD}} \sim 1$ dex. The fraction of galaxies where the two priors find $|\Delta \log_{10}(M/M_\odot)| > 0.5$ dex is very small at 5.9%. As seen in Figure 6, the outliers are predominantly at the low mass end, with the 0.5 dex outlier rate for $\log_{10}(M/M_\odot) \geq 10$ being $\sim 2\%$

4.5. *Individual Galaxy Inference: AGN*

To test the robustness of our inferences for AGN host galaxies, we look briefly at our results for the 1,951 COSMOS2020 $Ch.\,1 < 26$ galaxies with *Chandra* X-ray detections (based on the J. R. Weaver et al. 2022 cross-match against the F. Civano et al. 2016 catalog). In Figure 7, we compare the posterior median estimates of $\ln(f_{\mathrm{AGN}})$ for these galaxies to the posterior medians for the full COSMOS2020 catalog, and to the `pop-cosmos` prior over $\ln(f_{\mathrm{AGN}})$. Whilst the overlap between X-ray- and IR-bright AGN is not expected to be perfect (see, e.g., S. M. LaMassa et al. 2019; C. M. Carroll et al. 2021), both of these observational signatures are thought to correspond to radiatively efficient modes of accretion so a degree of correlation is expected. Examining the `pop-cosmos` model draws for all $Ch.\,1 < 26$ galaxies, we see a bimodal distribution in $\ln(f_{\mathrm{AGN}})$ similar to A24, with the two modes corresponding roughly to a

low-AGN ($f_{\mathrm{AGN}} \lesssim 3\%$) and high-AGN ($f_{\mathrm{AGN}} \gtrsim 3\%$) state. For the `pop-cosmos` model trained in this work, we find that $\sim 40\%$ of model galaxies have $f_{\mathrm{AGN}} > 3\%$. Looking at the f_{AGN} posteriors for the 1,951 *Chandra* X-ray detected galaxies, we find that $\sim 67\%$ of these galaxies have a posterior median AGN luminosity fraction greater than 3%, with the distribution of posterior medians in Figure 7 skewing strongly towards the high-AGN mode of the prior. Our posterior medians for the full COSMOS2020 $Ch.\,1 < 26$ catalog show a weaker skew towards the high-AGN mode.

4.6. *Mass Completeness*

It is important to quantify the regime of validity of the trained population model. We do this in two ways: (1) by estimating a mass completeness limit for the population described by the model; and (2) by considering the tails of the stellar mass distribution where a COSMOS-sized field may be subject to large sample variance (C. K. Jespersen et al. 2025b). Figure 8 shows the conditional distribution of stellar mass as a function of redshift for $0 \leq z \leq 4$, computed using the draws from the `pop-cosmos` generative model. The 99.5th percentile (gray shaded region) is shaded in gray. We follow A24 in estimating the mass completeness limit of the `pop-cosmos` mock galaxy catalog by identifying the mode or "turnover" of the stellar mass distribution as a function of redshift. This is illustrated with a dark shaded region, with a spline fit to this completeness limit shown as a black dashed line. In subsequent analyses, we focus on the mass complete part of the mock cat-
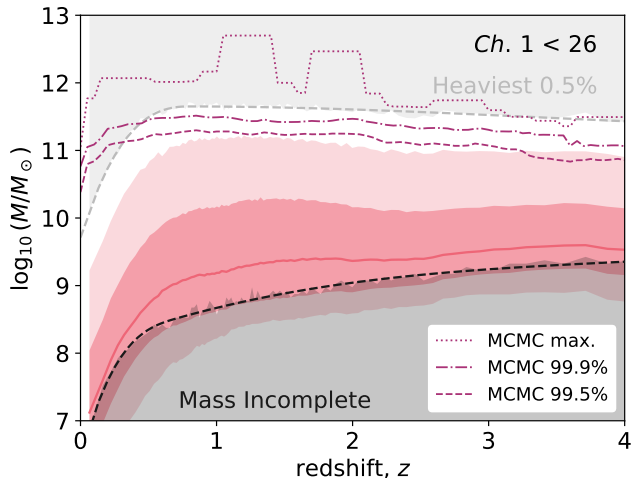
Figure 8. Stellar mass vs. redshift relation computed from the `pop-cosmos` model draws. Red shaded regions show the 68 and 95% conditional credible intervals, with the solid red line showing the median. The gray shaded region and dashed line correspond to the 99.5th percentile of stellar mass at a given redshift. The black shaded region and dashed line show the estimated mode or "turnover" of the stellar mass distribution as a function of redshift. Purple lines show upper limits on the masses of COSMOS2020 galaxies, estimated from the galaxy-level SED fits presented in Section 4.4.

Figure 9. Stellar mass vs. redshift computed from the `pop-cosmos` model draws for galaxies with $25 < Ch.\,1 < 26$ and $i > 27.6$ (3.5% of the $Ch.\,1 < 26$ mock catalog). The mass completeness limit from Figure 8 is shown as a yellow dashed line and hatched region.

alogs, omitting the model galaxies that fall below the estimated completeness threshold.

Our estimated mass completeness limit for `pop-cosmos` is consistent with the COSMOS2020 analysis by J. R. Weaver et al. (2023b), despite the difference in approach. J. R. Weaver et al. (2023b) estimate their stellar mass completeness limits using an approach similar to L. Pozzetti et al. (2010), with a limiting mass that is quadratic in $(1 + z)$. Our estimate tends to be slightly more conservative; around ~ 0.1–0.3 dex higher across the redshift range. For example, at redshifts $z = [1.0, 2.0, 4.0, 6.0]$ our estimated limiting mass is $\log_{10}(M/M_\odot) = [8.68, 9.05, 9.36, 9.70]$, compared to $\log_{10}(M/M_\odot) = [8.40, 8.78, 9.25, 9.56]$ using equation 3 from J. R. Weaver et al. (2023b).

Overlaid on Figure 8 with purple dashed and dash-dotted lines, we show respectively 99.5th and 99.9th percentiles of stellar masses estimated for individual COSMOS2020 galaxies from the SED fits in Section 4.4. This comparison shows that the most massive 0.5% of `pop-cosmos` model draws are likely not well-represented in the training data. Therefore, we excise these from subsequent analyses. Appendix E.1 provides a fitting function for masking these galaxies, shown as a gray dashed line in Figure 8. Appendix E.2 provides a fitting function for the mass incompleteness, shown as a black dashed line in 8.

To test the impact of the COSMOS2020 incompleteness at $Ch.\,1 \gtrsim 25$, we examine the stellar masses and redshifts of the model galaxies at the faint end of the $Ch.\,1$ magnitude distribution (i.e., in the region where `pop-cosmos` and COSMOS2020 begin to diverge in Figure 1). For the `pop-cosmos` model, 36.8% of galaxies with $Ch.\,1 < 26$ have $Ch.\,1 > 25$. For COSMOS2020, we find that 32.8% of detected galaxies lie within the $25 < Ch.\,1 < 26$ bin. We further require that model galaxies are optically faint, with $i > 27.6$ (based on the COSMOS2020 3σ depth; J. R. Weaver et al. 2022). We find that 3.5% of mock galaxies have $25 < Ch.\,1 < 26$ and $i > 27.6$, compared to just 0.9% of COSMOS2020 galaxies[20]. In Figure 9, we show the stellar masses and redshifts of `pop-cosmos` model galaxies with $25 < Ch.\,1 < 26$ and $i > 27.6$. We see that the majority of these galaxies are at $z \gtrsim 1$ and have low stellar masses, generally within ~ 0.5 dex of the completeness threshold defined in Figure 8. We calculate that $\sim 56\%$ of model galaxies with $25 < Ch.\,1 < 26$ and $i > 27.6$, or $\sim 69\%$ of galaxies with $25.5 < Ch.\,1 < 26$ and $i > 27.6$ lie below the mass completeness limit. We therefore infer that the $Ch.\,1 < 26$ galaxies missed by the COSMOS2020 detection pipeline are most likely to be low-mass galaxies in the mass-incomplete regime. Missing such galaxies from our training data is unlikely to

---

[20] The i-band is the deepest in the COSMOS detection image, so requiring an i band magnitude fainter the 3σ depth is an imperfect proxy for a galaxy's detectability. As discussed in depth by J. R. Weaver et al. (2023b), such a cut will not perfectly model the detection process, which can detect $i > 27.6$ galaxies that are bright in the shallower but redder K_S-band, or which have adequate S/N in the $izYJHK_S$ co-add (see Section 2.1).

significantly impact our conclusions, which are based on the mass-complete regime.

5. ASTROPHYSICAL PREDICTIONS

The data-driven calibration of the pop-cosmos model to the IRAC *Ch. 1*-selected COSMOS2020 catalog yields a population prior over the galaxy population probed by this selection. Samples drawn from this prior can be interpreted as synthetic galaxy populations, from which we can make predictions of population-level trends in galaxy-evolution. In effect, by marginalizing over different parameters in the learned SPS parameter distribution (or derived quantities thereof), we are able to make a range of astrophysical predictions. A selection of these are showcased in Section 5.1–5.7. A complete corner plot over the SPS parameter space is shown in in Appendix G.

These predictions from the trained generative model serve twin purposes. In the cases where well-explored trends in galaxy evolution exist in the literature, comparing the predictions from pop-cosmos with these relationships serves to further validate the model and increase trust in its robustness as a generative model for other galaxy surveys well-represented by the *Ch. 1 < 26* selection. However, as the model encapsulates the full joint distribution of the SPS parameters, which is highly challenging to access through other methods, its predictions can also provide insight into open questions in galaxy evolution.

5.1. *Redshift Distribution*

Inferring the redshift distribution of a population of galaxies, subject to selection, is one of the most important statistical challenges of doing cosmology with photometric survey data (for background, see B. Leistedt et al. 2016; A. I. Malz 2021; A. I. Malz & D. W. Hogg 2022; J. A. Newman & D. Gruen 2022). In Figure 10, we show the predicted redshift distribution for the limited COSMOS-like mock catalog with *Ch. 1 < 26*. We compute the redshift distribution of our mock galaxies with a redshift bin width of $\Delta z = 0.05$. For each bin, we compute the fractional cosmic variance based on B. P. Moster et al. (2011). In this calculation, we assume a COSMOS-sized field ($84' \times 84'$), and use the galaxy bias parameters for a stellar mass threshold of $\log_{10}(M/M_\odot) > 8.5$ in each bin (from table 4 of B. P. Moster et al. 2011).

Overlaid on Figure 10, we show a histogram of the posterior median redshifts for the actual COSMOS2020 galaxies, based on the inference presented in Section 4.3. For *r < 25*, we see a prominent spike for the $0.3 \le z < 0.4$ redshift bin, corresponding to large-scale structure (LSS) in the COSMOS field at $z \sim 0.35$ (see, e.g., N. Scoville et al. 2007b; I. K. Söchting et al. 2012; D. Chérouvrier et al. 2025). This feature is less prominent in the *Ch. 1 < 26* distribution. Several higher-z spikes are visible prominently in the *Ch. 1*-selected
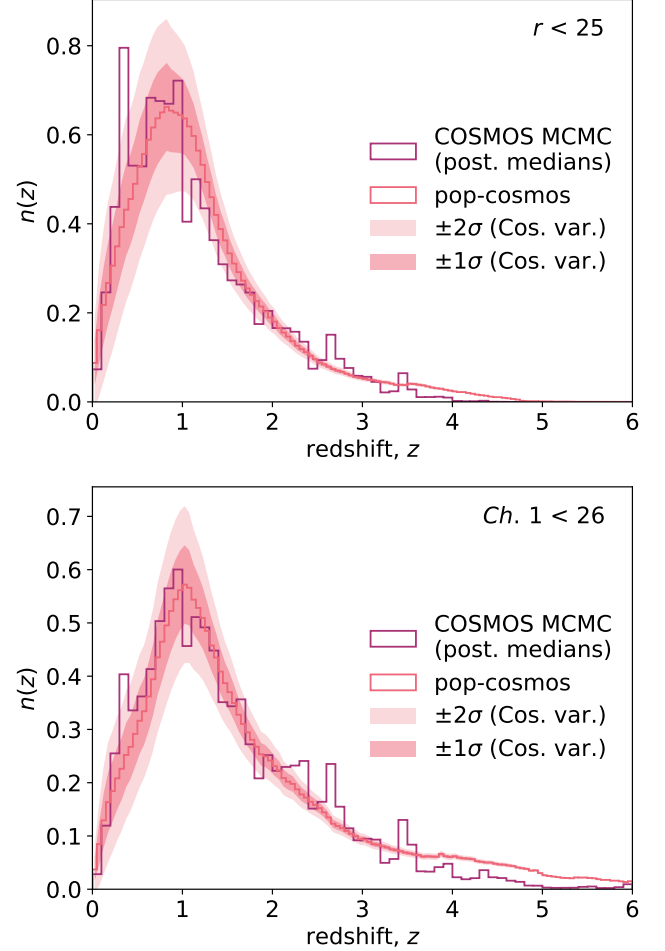
Figure 10. Normalized redshift distribution for **(top)** *r < 25* and **(bottom)** *Ch. 1 < 26*. Shaded regions show the fractional effect of cosmic variance that would be expected for a COSMOS-sized field. The purple histogram shows posterior median redshifts for **(top)** 146,252, and **(bottom)** 423,262 COSMOS galaxies that pass each selection cut (see Section 4.3).

redshift histogram, potentially correlating with $z \ge 2$ LSS and proto-clusters (for an extensive review see, e.g., G. Toni et al. 2025; I. McConachie et al. 2025). Confirmed structures at $z \sim 2.5$ (e.g., C. Diener et al. 2015; Y.-K. Chiang et al. 2015; O. Cucciati et al. 2018; D. Sikorski et al. 2025), $z \sim 2.7$ (K. Ito et al. 2023), and $z \sim 3.4$ (I. McConachie et al. 2022; B. Forrest et al. 2023) may be responsible for the two most prominent features.

Additionally visible when comparing the galaxy-level posterior medians to the pop-cosmos predictions is a less steep drop-off in sources in the model $n(z)$ at $z \gtrsim 4$ when compared to the galaxy-level inferences. This is seen prominently for the *Ch. 1 < 26* selection in the lower panel of Figure 10, and less strongly in the *r < 25* selection for $4 \lesssim z \lesssim 5$. We hypothesize that the galaxies "missing" from the COSMOS sample at $z \gtrsim 4$

are those lost due to the incompleteness of the data at $25 \lesssim Ch.\,1 \lesssim 26$ (see Sections 2.1 and 4.1, and I. David-zon et al. 2017; J. R. Weaver et al. 2022, 2023b). We find that $\sim 57\%$ of pop-cosmos model galaxies with $z > 5$ have $25 < Ch.\,1 < 26$, placing them in the range where the observed and model $Ch.\,1$ magnitude distributions being to diverge (see Figure 1).

5.2. Stellar Mass Function

The redshift evolving probability density function over stellar mass — the stellar mass function — has been a key observational target (for recent examples see, e.g., J. Leja et al. 2020; S. P. Driver et al. 2022; J. R. Weaver et al. 2023b; M. Shuntov et al. 2025a; Euclid Collabo-ration et al. 2025b), and is a fundamental test for cos-mological models (see, e.g., R. S. Somerville & R. Davé 2015). Figure 11 shows the stellar mass distribution of the draws from the pop-cosmos generative model for $Ch.\,1 < 26$. We shade out the most massive 0.5% of galaxies in each redshift bin on the basis of the re-sults in Figure 8. For $M \lesssim 10^{11} M_\odot$, we find that the distribution of stellar masses agrees well with the distribution implied by SED fits to the COSMOS2020 galaxies from Section 4.4. At $M \gtrsim 10^{11} M_\odot$, we see that pop-cosmos has a much heavier tail than found by the COSMOS SED fits from Section 4.4, although the fraction of the 2 million mock galaxies that lie in this region is very small, generally within the grayed-out 99.5th percentile. The completeness limit (based on the turnover of the pop-cosmos mass function) is shown in black at the left-hand edge of each panel. In the $0.2 < z < 0.4$ bin, the mass distribution implied by the SED fitting results shows a noticeable shoulder around $10^{10} \lesssim M/M_\odot \lesssim 10^{11}$. This may be driven by the higher prevalence of large scale structures seen in the COSMOS field around $z \sim 0.35$ (N. Scoville et al. 2007b; I. K. Söchting et al. 2012; D. Chérouvrier et al. 2025); see also Section 5.1 for further discussion.

Overlaid on Figure 11 is the double Schechter mass function constrained by J. Leja et al. (2020) using data from COSMOS2015 (C. Laigle et al. 2016) and 3D-HST (R. E. Skelton et al. 2014; K. E. Whitaker et al. 2014; I. G. Momcheva et al. 2016). The J. Leja et al. (2020) mass function is evaluated based on the posterior me-dian parameter values given in their figure 3. We chose to show the J. Leja et al. (2020) result as our compari-son here, as it is determined based on a deep sample se-lected from the COSMOS field, and as a self-consistent constraint on the star-forming main sequence is avail-able for comparison (see our Section 5.3) from J. Leja et al. (2022). Additionally, the J. Leja et al. (2020) mass function is inferred in a Bayesian framework, with full propagation of uncertainties, making it a particularly interesting comparison.

For $z > 0.4$, the low-mass slope of the pop-cosmos mass function shows excellent agreement with J. Leja et al. (2020) for masses above our estimated complete-
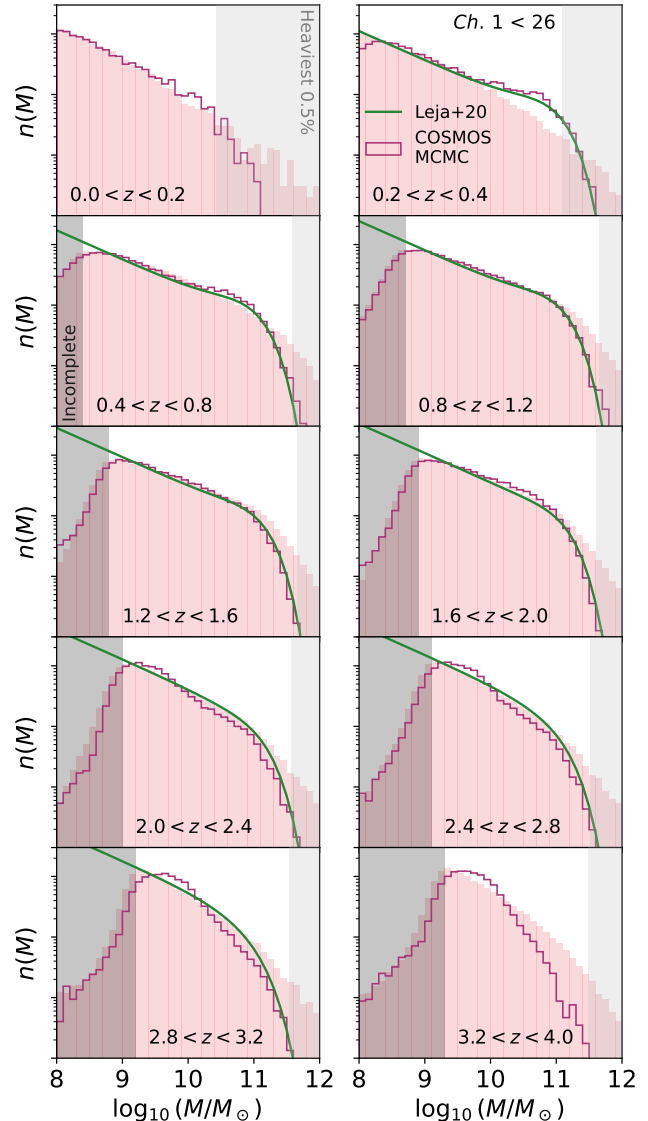
Figure 11. Stellar mass function (normalized, logarithmic axes) predicted by pop-cosmos (solid red histogram), com-pared to the stellar masses from Bayesian SED fitting (pur-ple) to COSMOS2020 galaxies selected on $Ch.\,1 < 26$ (see Section 4.4). For the pop-cosmos draws, we consider the most massive 0.5% of galaxies at a given redshift (rightmost gray shaded region) to be an extrapolation due to the rarity of such galaxies in a COSMOS-sized field. The histograms are normalized to integrate to 1 between the vertical shaded regions. In green we plot the double Schechter mass function from J. Leja et al. (2020), normalized within the same limits.

ness limit. The high-mass slope of the J. Leja et al. (2020) mass function agrees well with the mass distri-bution implied by our galaxy-level inference (the pur-ple unshaded histogram in Figure 11), declining more steeply past the "knee" of the mass function than the distribution implied by the pop-cosmos predictions. In-

terestingly, the results from J. Leja et al. (2020) are closer to the galaxy-level inference in the $0.2 < z < 0.4$ bin, with a more pronounced knee than is seen in the `pop-cosmos` predictions. At $z < 0.5$, the J. Leja et al. (2020) stellar mass function is only constrained by COSMOS2015, so any large-scale structure present in the field at low-z (e.g., the $z \sim 0.35$ feature seen in Figure 10 and by N. Scoville et al. 2007b; I. K. Söchting et al. 2012; D. Chérouvrier et al. 2025) could influence our galaxy-level inferences and the J. Leja et al. (2020) mass function. Inspecting the stellar mass function from our galaxy-level inference in narrower redshift bins, we find that the $0.3 \le z < 0.4$ slice has a noticeable excess of $10 \lesssim \log_{10}(M/M_\odot) \lesssim 11$ galaxies compared to the $0.2 \le z < 0.3$ and $0.4 \le z < 0.5$ bins, consistent with the presence of a cluster. Within this mass range, we find that the $0.3 \le z < 0.4$ bin also has an excess of galaxies with $\log_{10}(\mathrm{SFR}/M_\odot\,\mathrm{yr}^{-1})$ compared to the neighbouring bins, further aligning with this picture.

We forgo a more detailed to other mass function estimates in the literature (e.g., the COSMOS2020 results from J. R. Weaver et al. 2023b, or the Cosmic Dawn Survey results from Euclid Collaboration et al. 2025b), but will revisit the stellar mass function predicted by `pop-cosmos` in future work by S. Deger et al. (in prep.). Recent observational studies (J. R. Weaver et al. 2023b; M. Xiao et al. 2024; A. Weibel et al. 2024; Euclid Collaboration et al. 2025b; but see also M. Shuntov et al. 2025a) and simulations (e.g., J. S. Bennett et al. 2024) have favoured the presence of non-negligible numbers of very massive galaxies with $\log_{10}(M/M_\odot) \gtrsim 11.5$, and simulations (e.g., C. d. P. Lagos et al. 2018, 2019; R. Davé et al. 2019) often predict shallower high-mass slopes than observations (see, e.g., comparisons in N. J. Adams et al. 2021; J. R. Weaver et al. 2023b). However, observational constraints are sensitive to the strategy adopted for correcting Eddington bias (A. S. Eddington 1913, 1940), with this substantially influencing the inferences made about the high-mass end of the mass function (see, e.g., O. Ilbert et al. 2013; K. I. Caputi et al. 2015; A. Grazian et al. 2015; I. Davidzon et al. 2017; N. J. Adams et al. 2021; J. R. Weaver et al. 2023b; Euclid Collaboration et al. 2025b).

J. Leja et al. (2020) argues that their hierarchical Bayesian approach with principled uncertainty propagation from the photometry to the mass function ought to naturally avoid Eddington bias. A similar argument can be made for the population-level approach we have taken here with `pop-cosmos`. Irrespective of Eddington bias, the high-mass end of the stellar mass function is also generally more susceptible to the effects of cosmic variance when constrained based on small-area surveys (see, e.g., the discussion in C. K. Jespersen et al. 2025b). The calibration of the `pop-cosmos` stellar mass function at the massive end can likely be improved in future work by incorporating data from a wider area (e.g., the full
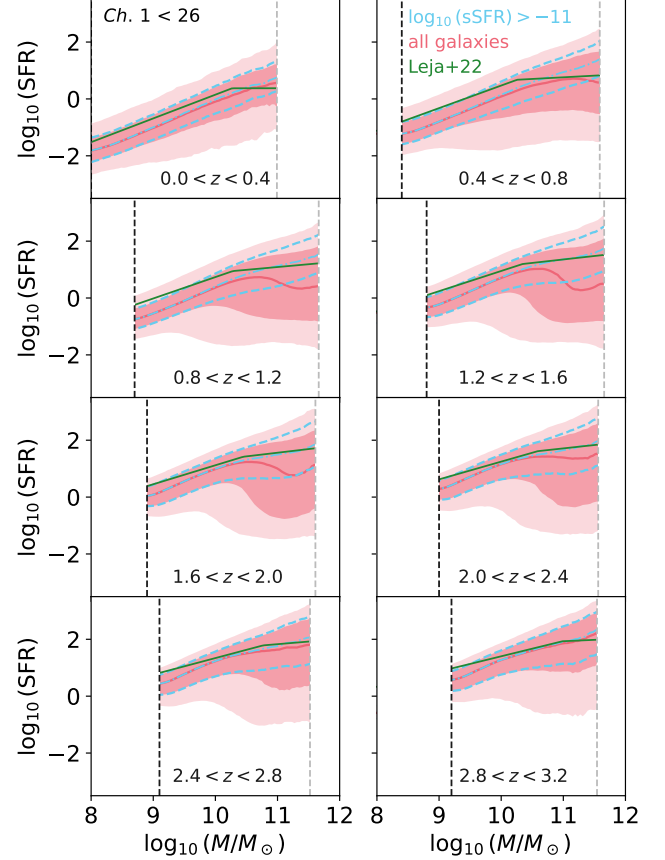
Figure 12. Star forming main sequence predicted by `pop-cosmos`. The solid red line shows the rolling median star formation rate $(M_\odot\,\mathrm{yr}^{-1})$ conditional on stellar mass. Red shaded regions show the rolling 68 and 95% credible intervals. Rolling quantities are computed in a stellar mass window of width $\delta\log_{10}(M/M_\odot) = 0.4$ dex. The mass limits from Figure 11 are applied and indicated with vertical dashed lines. The overlaid cyan lines show the median and 68% interval for the subset of star-forming model galaxies with $\log_{10}(\mathrm{sSFR}/\mathrm{yr}^{-1}) > -11$. The overlaid green line shows the double power-law from J. Leja et al. (2022).

670 deg^2 HSC survey area; H. Aihara et al. 2022; or the 1347 deg^2 KiDS catalog; A. H. Wright et al. 2024).

5.3. *Star-Forming Main Sequence*

The tight correlation between stellar mass and SFR amongst star forming galaxies (see E. Daddi et al. 2007; K. E. Whitaker et al. 2012, 2014; J. S. Speagle et al. 2014; P. Popesso et al. 2023) has important implications for their evolutionary pathways (e.g., Y.-J. Peng et al. 2010; S. N. Leitner 2012; L. E. Abramson et al. 2015), and has seen a growing qualitative consensus over the past 15 years of observational studies, albeit with some quantitative tensions (see discussion in S. Wuyts et al. 2011; J. S. Speagle et al. 2014; J. Leja et al. 2015, 2022; L. Sandles et al. 2022; P. Popesso et al. 2023; H. Fu et al.
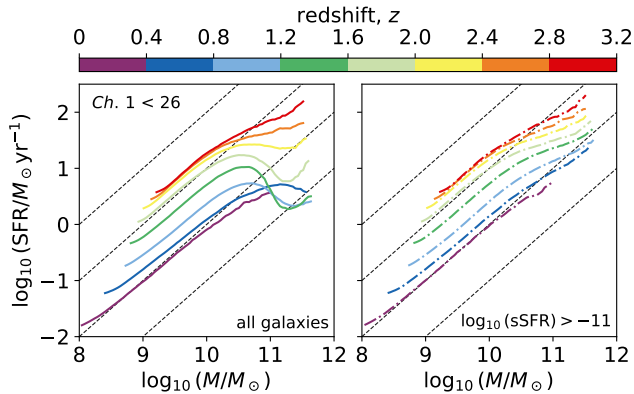
Figure 13. Median star forming sequence predictions from Figure 12 for **(left)** all galaxies, and **(right)** galaxies with $\log_{10}(\text{sSFR}) > -11$. The black dashed lines show constant sSFR; top to bottom these correspond to $\log_{10}(\text{sSFR/yr}^{-1}) = [-8, -9, -10, -11]$.

2024). Figure 12 shows the star-forming main sequence predicted by pop-cosmos for *Ch. 1 < 26*. In the lowest redshift bin, this is highly consistent with the results from A24 (whose results have shown good consistency at $z = 0$ with simulations; J. S. Bennett et al. 2024). At higher redshifts, we see that the 68% credible interval encompasses lower SFRs, particularly at the massive end, than were included previously. This broadening of the distribution is likely reflective of the greater completeness of the *Ch. 1 < 26* selection with respect to quiescent galaxies (J. R. Weaver et al. 2023b). In Figure 12 we also show the median and 68% interval for star-forming galaxies — selected based on $\log_{10}(\text{sSFR/yr}^{-1}) > -11$ (e.g., O. Ilbert et al. 2010, 2013; H. Domínguez Sánchez et al. 2011) — which shows the expected monotonic rise. The star-forming/quiescent split of pop-cosmos model galaxies, and the demographics of these samples will be explored in more detail by S. Deger et al. (in prep.).

Overlaid on Figure 12 is the double power-law SFS model constrained by J. Leja et al. (2022) using the same COSMOS2015 and 3D-HST sample as was used in the J. Leja et al. (2020) stellar mass function analysis, which agrees well with other literature results with regard to the shape and slope of the SFS (for an exhaustive review and comparison of these, see J. S. Speagle et al. 2014; P. Popesso et al. 2023). It is worth noting that J. Leja et al. (2022) found a consistently lower normalization at a given stellar mass than previous literature results; likely due to improved stellar mass and SFR estimates arising from their use of a non-parametric SFH (see A. C. Carnall et al. 2019; J. Leja et al. 2019b,a; S. Lower et al. 2020; L. Sandles et al. 2022). We show the J. Leja et al. (2022) result by evaluating the fitting formula in their table 1 at the central redshift of each bin in Figure 12. We find overall very good agreement with J. Leja et al. (2022) across the full redshift range

covered in Figure 12. At $z > 1$, we find a slightly steeper relation at the low mass end than J. Leja et al. (2022); this aligns well with recent empirical modeling results based on abundance matching (H. Fu et al. 2024), and predictions from cosmological simulations (E. Chaikin et al. 2025). For $z < 2$, we find that the low-mass slope remains fairly consistent, and that the median relation at low masses has a gradient fairly close to unity in this redshift range. At $1.6 < z < 2$, the median $\log_{10}(\text{SFR})$ vs. $\log_{10}(M/M_\odot)$ relation for $\log_{10}(M/M_\odot) \lesssim 10$ corresponds to sSFR $\sim 10^{-9}\,\text{yr}^{-1}$. By the lowest redshift bin at $0.0 < z < 0.4$, the sSFR implied by the median relation has dropped by almost an order of magnitude to sSFR $\sim 10^{-10}\,\text{yr}^{-1}$, with this being the median across the full mass range $\log_{10}(M/M_\odot) < 11$. For $z > 2$, the low-mass slope predicted by pop-cosmos falls somewhere between the shallower slope inferred by J. Leja et al. (2022), and the slightly steeper slope that would be implied by a constant sSFR $= 10^{-9}\,\text{yr}^{-1}$. The median $\log_{10}(\text{SFR})$ vs. $\log_{10}(M/M_\odot)$ relations from Figure 12 are shown in Figure 13, with lines of constant sSFR highlighted.

5.4. *Mass–Metallicity Relation*

Both the gas-phase (C. A. Tremonti et al. 2004) and stellar metallicity (A. Gallazzi et al. 2005) of galaxies have been observed (predominantly using spectroscopic indicators of these quantities) to correlate with stellar mass, with important implications for the chemical evolution of galaxies (for a review, see R. Maiolino & F. Mannucci 2019). In Figure 14, we show the stellar metallicity vs. stellar mass relation in increasing redshift bins. In each bin, we show the median metallicity as a function of stellar mass, as well as the 68 and 95% credible intervals. We overlay in cyan the gas-phase metallicity vs. stellar mass relation (median, 68%, and 95% credible intervals). As in A24, the gas-phase metallicity is consistently higher than the stellar metallicity at all redshifts. In all redshift bins, the gas-phase metallicity decreases less steeply towards low masses than the stellar metallicity. At the high mass end, we typically see an offset of $\sim +0.2$ dex between gas-phase and stellar metallicity, with a larger offset at the low-mass end.

Figure 15 shows the median stellar and gas-phase metallicity–mass relations from Figure 14. We show these for both the $r < 25$ and *Ch. 1 < 26* mock galaxy catalogs generated with pop-cosmos. We find a much stronger redshift evolution for the optically-selected sample compared to the MIR-selected one; potentially due to the MIR selection including a wider diversity of galaxy types at any given redshift. In both cases however, the evolution of the median relation is much smaller than the inferred scatter in any given redshift bin. For both selections, we observe a steepening of the relation with redshift. The results for the $r < 25$ sample are in good general agreement with the trends
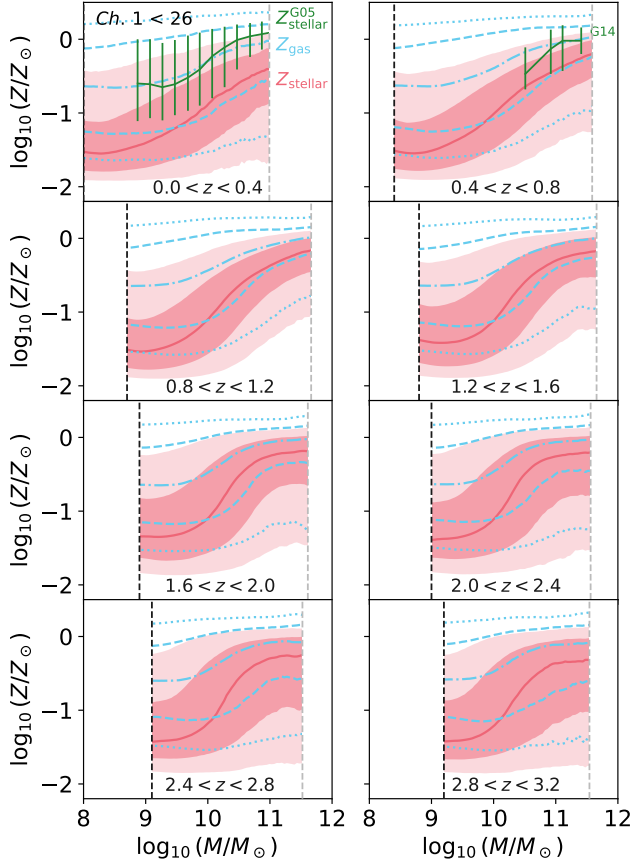
Figure 14. Stellar mass–metallicity relation (red) predicted by pop-cosmos in the redshift bins from Figure 12, for galaxies with *Ch. 1* < 26. The gas-phase metallicity vs. stellar mass relation is overlaid in cyan. Overlaid in green on the first panel, we show the A. Gallazzi et al. (2005, G05) stellar metallicity vs. mass relation (median and 68% interval) based on SDSS spectra. On the second panel, the green shows the *z* ~ 0.7 results from A. Gallazzi et al. (2014, G14).

Figure 15. Median mass–metallicity relations predicted by pop-cosmos in the redshift bins from Figure 14. We show the predictions for (**left**) the *r* < 25 selection, (**right**) the *Ch. 1* < 26 selection; and (**top**) the stellar metallicity, (**bottom**) the gas-phase metallicity.
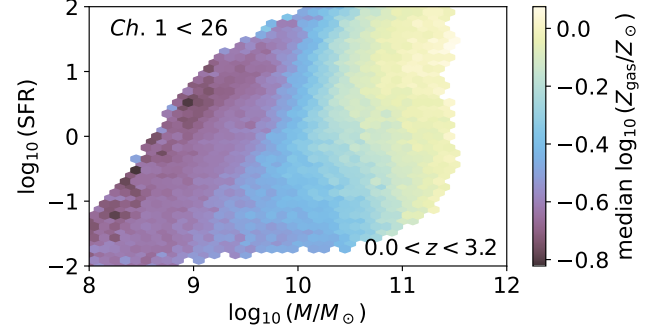
Figure 16. Fundamental metallicity relation for *Ch. 1* galaxies at $0.0 < z < 3.2$, visualized in the style of M. Curti et al. (2020). Mock galaxies drawn from the generative model are binned by stellar mass and SFR, with the cells shaded by the median gas-phase metallicity. Cells are shaded only if they contain > 100 model galaxies.

seen by A24, although we see less of an obvious convergence in metallicities at high stellar mass.

Comparing our $0.0 < z < 0.4$ results to the well known stellar mass vs. metallicity relation (shown in green and indicated G05 on Figure 14) from A. Gallazzi et al. (2005), our median relation is ~ 0.5 dex lower at the high-mass end, and ~ 1.0 dex lower at $\log_{10}(M/M_\odot) \approx 9$ (the lowest stellar mass reported by A. Gallazzi et al. 2005). The A. Gallazzi et al. (2005) analysis is based on SDSS galaxies (D. G. York et al. 2000; K. Abazajian et al. 2004) with $r < 17.77$; brighter than almost anything in the COSMOS field. In the $0.4 < z < 0.8$ redshift bin, the pop-cosmos median relation shows a smaller offset when compared to the $z \sim 0.7$ results from A. Gallazzi et al. (2014), which were based on a similar methodology to A. Gallazzi et al. (2005). Due to the selection involved in the A. Gallazzi et al. (2005, 2014) analyses, and the very different methodol-

ogy used to constrain stellar metallicity, interpretation of the offset is challenging. A. Gallazzi et al. (2005, 2014) base their constraints on comparison between a set of absorption line diagnostics (e.g., G. Worthey 1994) to model spectra from the G. Bruzual & S. Charlot (2003) stellar population models. Results at low-*z* from full spectral fitting (e.g., B. Panter et al. 2008; H. J. Zahid et al. 2017), or from measurements of individual supergiant stars (e.g., R. P. Kudritzki et al. 2016), have dif-

fered in the asymptote at high masses, and the predicted behavior at low masses[21].

High-z studies typically rely on different spectroscopic absorption (e.g., S. A. Rix et al. 2004; C. Halliday et al. 2008; V. Sommariva et al. 2012; A. Calabrò et al. 2021), or full spectral fitting (e.g., F. Cullen et al. 2019; D. Kashino et al. 2022; N. Chartab et al. 2024). The high-z spectroscopic constraints typically find much lower stellar metallicities than the low-z relations inferred by e.g., A. Gallazzi et al. (2005), R. P. Kudritzki et al. (2016), or H. J. Zahid et al. (2017). Results from F. Cullen et al. (2019, $2.5 \leq z \leq 5.0$), D. Kashino et al. (2022, $1.6 \leq z \leq 3.0$) and N. Chartab et al. (2024, $2 \lesssim z \lesssim 3$) suggest that the median stellar metallicity at a given mass is shifted downwards by ~ 0.5–1.0 dex relative to the relation at low-z, with a fairly similar slope between the low- and high-mass ends. This is in good qualitative agreement with the trend we predict for the $r < 25$ selection (which should be more reflective of the depth of the high-z spectroscopic surveys used; S. J. Lilly et al. 2007; L. Pentericci et al. 2018; R. J. McLure et al. 2018; A. B. Newman et al. 2020; D. Kashino et al. 2021). As shown by F. Cullen et al. (2019), the normalization of the stellar metallicity vs. mass relation predicted by simulations (e.g., X. Ma et al. 2016; R. Davé et al. 2019) is often lower than the spectroscopic constraints suggest, although predictions for the shape and redshift evolution are similar.

As it is traced by emission — rather than absorption — features, the stellar mass dependence of gas-phase metallicity has been studied more extensively with spectroscopic observations. Of particular interest has been the fundamental metallicity relation (FMR) between stellar mass, SFR, and gas-phase metallicity (F. Mannucci et al. 2010; M. A. Lara-López et al. 2010). We visualize the FMR implied by draws from the pop-cosmos model in Figure 16, by plotting SFR against stellar mass, with each cell shaded based on the median $\log_{10}(Z_{\mathrm{gas}}/Z_\odot)$ in that part of the $\log_{10}(\mathrm{SFR})$ vs. $\log_{10}(M/M_\odot)$ plane (inspired by figure 5 of M. Curti et al. 2020). In Figure 16, we include all pop-cosmos model draws with $z < 3.2$ and $Ch. 1 < 26$, motivated by observational (e.g., F. Mannucci et al. 2010; H. J. Zahid et al. 2014; L. Hunt et al. 2016; G. Cresci et al. 2019; M. Curti et al. 2024) and theoretical (e.g., S. J. Lilly et al. 2013) results suggesting approximate invariance of the FMR out to at least $z \sim 3$. We find that the pop-cosmos result in Figure 16 is fairly insensitive to the upper redshift limit chosen, or the choice of an $r < 25$ selection, although in future work a more detailed investigation of this would be worthwhile. We see good qualitative agreement between our Figure 16

[21] Additionally, studies of local dwarf galaxies have suggested extrapolation to extremely subsolar metallicities at very low mass (E. N. Kirby et al. 2013).
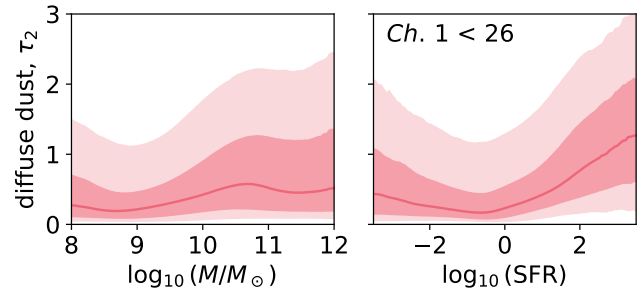
Figure 17. Diffuse dust attenuation τ_2 vs. **(left)** stellar mass, and **(right)** star formation rate, as predicted by pop-cosmos for galaxies with $Ch. 1 < 26$.

and figure 5 of M. Curti et al. (2020). The strongest trend is one of metal enrichment with increasing stellar mass at a fixed SFR. We also see the expected gradient from the upper left towards the lower right of the Figure, implying that the most star-forming galaxies at a given stellar mass have the lowest gas-phase metallicity (with a weaker SFR dependence at the high mass end, reflecting the saturation). As noted by M. Curti et al. (2020), quantitative comparisons of the FMR even amongst spectroscopic measurements can be challenging due to differing selection criteria and metallicity calibrations, so we forgo a more detailed comparison here.

5.5. Dust Attenuation

Modeling the dust attenuation of starlight (for a review see D. Calzetti 2001; S. Salim & D. Narayanan 2020) is crucial for the interpretation of other observables (see, e.g., S. Salim & M. Boquien 2019; S. Lower et al. 2022; G. Nagaraj et al. 2022a). Correlations between attenuation properties and other quantities such as stellar mass and SFR have been investigated in detail both observationally (e.g., T. Garn & P. N. Best 2010; H. J. Zahid et al. 2013; S. Salim et al. 2016, 2018; G. Nagaraj et al. 2022b) and theoretically (e.g., A. N. Witt & K. D. Gordon 2000; J. Chevallard et al. 2013; L. Sommovigo et al. 2025). Figure 17 shows the distribution of diffuse dust attenuation, as a function of stellar mass and star formation rate. The results here are consistent with A24, albeit with a slightly heavier tail towards high dust attenuation at a given stellar mass or star formation rate. Again, this reflects the change of the pop-cosmos training set to include extremely dusty galaxies that are too optically faint to pass an $r < 25$ selection. The upturn in the level of dust attenuation for $\mathrm{SFR} \gtrsim 1\ M_\odot\,\mathrm{yr}^{-1}$ reproduces the trends seen by, e.g., H. J. Zahid et al. (2013) and G. Nagaraj et al. (2022b). The pop-cosmos predictions of dust attenuation are evaluated further by F. Petri et al. (2025), who find that the median attenuation predicted by pop-cosmos is consistently lower than G. Nagaraj et al. (2022b) for all SFRs, but that the upturn at high SFR occurs in the same place and has a comparable slope. F. Petri et al.
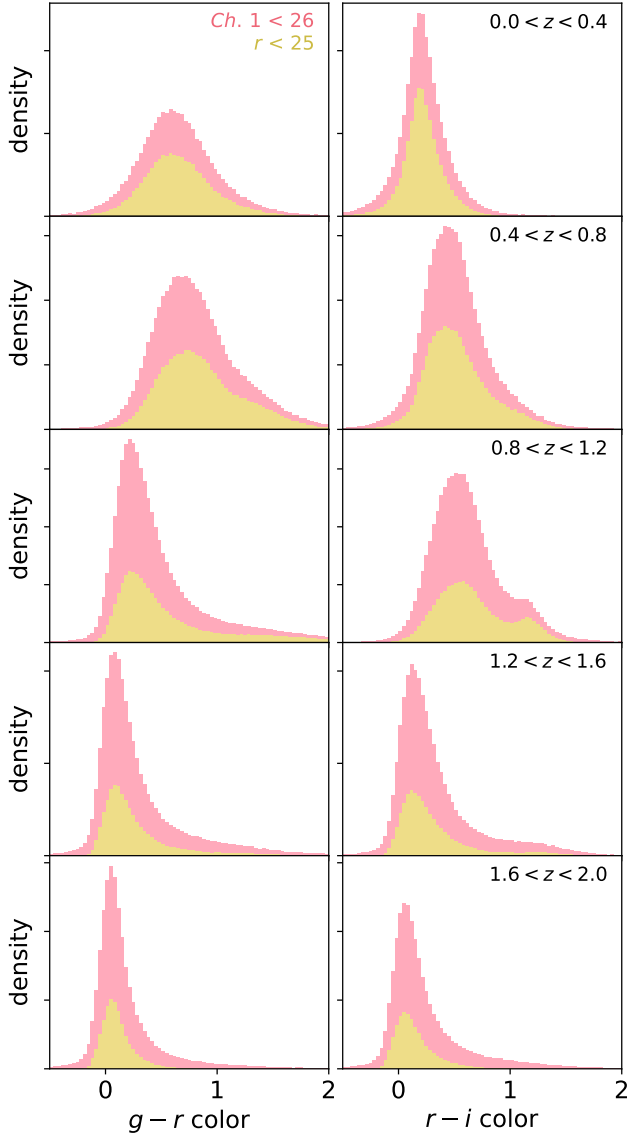
Figure 18. Redshift evolution of the observer-frame **(left)** $g - r$ and **(right)** $r - i$ color distributions predicted by `pop-cosmos`. Histograms are plotted for the *Ch. 1* < 26 mock catalog, with the $r < 25$ portion of this catalog shaded in yellow.

(2025) show that this has significant implications for the selection of Lyman-break galaxies, highlighting the importance of constraining this relationship at high-z.

5.6. *Observer-Frame Color–Redshift Relations*

As well as looking at the survey-wide distribution of observer-frame colors, as we did in Section 4.1, we can also examine the redshift evolution of these distributions and their associations with physical parameters. Figure 18 shows the observer-frame $g-r$ and $r-i$ color distributions predicted by `pop-cosmos` in a series of redshift bins for $z < 2.0$. Within this redshift range, the typical $g-r$

colors are reddest around $z \sim 0.6$. At $z \sim 1.0$ the $r-i$ color shows a noticeable bimodality — more pronounced for the $r < 25$ selection[22] — that is seen prominently in spectroscopic surveys that cover this range (e.g., J. A. Newman et al. 2013). The expected $g-r$ bimodality is less strongly visible.

In Figure 19, we show the $g-r$ and $r-i$ color vs. redshift relations predicted by `pop-cosmos` for $0 < z < 4$. The density (top row) at $z < 1.5$ shows the structure expected from deep spectroscopic surveys (e.g., J. A. Newman et al. 2013). At $z \approx 3.5$ in the $g-r$ color, we see a striking "pinch" point, coinciding with Ly-α 1215Å reaching the red edge of the g-band. In rows 2–4 of Figure 19, we shade each cell based on the median physical properties of galaxies in that cell. The correlation between $g-r$ and $r-i$ and SFR is strongest at $z \sim 1$. For $1 < z < 3$, the model predicts a strong correlation between optical color and mass, seen in the third row of Figure 19. The fourth row illustrates that the optical depth of diffuse dust strongly influences $r-i$ color at $z < 1.5$, and $g-r$ color at $z < 1$. From the fifth row, we see that there are small regions of color–redshift space with high median AGN fractions ($\gtrsim 10\%$) and blue colors at $z < 2$. The AGN model in `pop-cosmos` (M. Nenkova et al. 2008a,b; J. Leja et al. 2018) concerns re-radiation from a dust torus, so the association with blue color is not due to optical emission from an accretion disk. As was suggested by C. K. Jespersen et al. (2025a) recently, IR luminosity can be very tightly coupled to optical color, which could be influenced by feedback or AGN–host co-evolution. We now explore this result in greater detail.

5.7. *AGN Luminosity*

One potentially significant source of feedback in galaxies is AGN emission. We investigate this possibility in Figure 20, which shows the stellar mass vs. redshift relation from Figure 8, but with the right-hand panel shaded based on the median AGN bolometric luminosity fraction, $f_{\rm AGN}$, in each bin. In A24 and here (see the full SPS parameter distribution in Appendix G), we generally see a bimodal population distribution in $f_{\rm AGN}$, with the two modes broadly corresponding to negligible and non-negligible AGN contributions. We set the color scale in the right panel of Figure 20 to have its pivot point at 3×10^{-2}, approximately between the two modes of the population distribution[23] (c.f., Figures 7 and 34). This shows that there is a region with $9 \lesssim \log_{10}(M/M_\odot) \lesssim 11$ and $1 \lesssim z \lesssim 2$ where the median $f_{\rm AGN}$ falls within the non-negligible range.

---

[22] In Figure 18, we are plotting the *Ch. 1*-selected mock catalog, with the $r < 25$ portion of this highlighted in yellow.

[23] This corresponds to an AGN torus with a bolometric luminosity that contributes 3% as much as the stellar component.

Figure 20. Stellar mass vs. redshift for *Ch. 1* < 26. Cells are shaded by **(left)** model galaxy count, and **(right)** median AGN bolometric luminosity fraction, $f_{\rm AGN}$. Cells are shaded only if they contain > 10 model galaxies from the total sample of 2 million. The black and gray dashed lines show respectively the mass completeness limit and upper mass limit (see Figures 8, 27, and Appendix E). Solid, dashed, and dash-dotted yellow lines show respectively the 50th, 84th, and 97.5th percentiles of $\log_{10}(M/M_\odot)$ given z (same as Figure 8).

Figure 19. Observer-frame color vs. redshift relations, for **(left)** $g - r$ and **(right)** $r - i$, predicted by pop-cosmos for *Ch. 1* < 26. A cell in each panel is shaded only if it contains > 100 model galaxies (0.005% of the full mock catalog), and galaxies are only included above the mass completeness limit. Panels are shaded by: **(row 1)** model galaxy count; **(row 2)** median $\log_{10}(\mathrm{SFR}/M_\odot\mathrm{yr}^{-1})$; **(row 3)** median $\log_{10}(M/M_\odot)$; **(row 4)** diffuse dust optical depth at 5500Å, τ_2; and **(row 5)** median AGN bolometric luminosity fraction, $f_{\rm AGN}$.

In Figure 22, we show SFR vs. stellar mass, binned by redshift and shaded based on $f_{\rm AGN}$. As with Figure 20, we see the most pronounced AGN contributions around $1 \lesssim z \lesssim 2$, where the upper edge of the star forming main sequence correlates with a high median $f_{\rm AGN}$; i.e. the model predicts that galaxies with the highest sSFR have the strongest AGN-related emission in the IR. A non-stellar IR excess towards the upper edge of the star forming sequence was previously identified by

Figure 21. Same as row 5 of Figure 19, but showing observer-frame IRAC *Ch. 1 − Ch. 2* color vs. redshift.

J. U. Lange et al. (2016) for $0.5 \leq z \leq 1.5$, with a very similar tendency to the behavior we see in Figure 22. Observational works have showed that heavily-obscured IR-luminous AGN are ubiquitous within the galaxy population (e.g., S. Juneau et al. 2013; A. Kirkpatrick et al. 2012, 2015), and make up a substantial fraction of all AGN (e.g., M. Lacy et al. 2013; R. J. Assef et al. 2013, 2015, 2018; R. C. Hickox & D. M. Alexander 2018; C. M. Carroll et al. 2021), making them a plausible source of such a non-stellar excess (J. Leja et al. 2018). The IR signature of AGN tends to be reasonably distinguishable from that of star formation, due to the former's tendency to produce a power-law SED (identifiable at

$z \lesssim 2$ by very red *Ch. 1 − Ch. 2* colors; see, e.g., D. Stern et al. 2005, 2012; R. J. Assef et al. 2010; J. L. Donley et al. 2012; S. Juneau et al. 2013). In Figure 21, we show the pop-cosmos observer-frame *Ch. 1 − Ch. 2* vs. z relation, shaded by median f_{AGN}, highlighting that this color is a strong discriminant of AGN activity for $z \lesssim 2$ (c.f., D. Stern et al. 2005, figure 2). For $z \gtrsim 2$, the AGN IR power law is shifted too far redward in the observer frame for this color to be a strong diagnostic.

The redshift window in Figure 20 where we identify the strongest MIR signature correlates roughly with the redshift range where the cosmic SFR density and the black hole accretion rate density peak and begin to turn over (e.g., P. F. Hopkins et al. 2007; J. Aird et al. 2010; I. Delvecchio et al. 2014; P. Madau & M. Dickinson 2014; R. C. Hickox & D. M. Alexander 2018). In a unified picture of AGN–host co-evolution (see, e.g., T. M. Heckman & P. N. Best 2014; P. Padovani et al. 2017; C. M. Harrison 2017; D. M. Alexander et al. 2025), the presence of a gas reservoir would be favorable for both star formation and rapid black hole growth via radiatively efficient accretion modes, potentially correlating the two. Moreover, it is theoretically expected that the interplay between stellar (i.e. supernova) feedback and AGN feedback conspires to set a characteristic mass scale where non-linear black hole growth via rapid, efficient accretion will activate (see, e.g. Y. Dubois et al. 2015; R. G. Bower et al. 2017; S. McAlpine et al. 2017, 2018); this stellar mass scale of $\sim 10^{10.3} M_\odot$ aligns with our Figure 20. For galaxies much below this characteristic mass, stellar feedback is expected to dominate gas dynamics and prevent radiatively efficient accretion. For galaxies much above this mass, AGN feedback is sufficient to remove gas from the halo, regulating the accretion onto the black hole and shutting down star formation (see R. G. Bower et al. 2017; L. Lucie-Smith et al. 2025).

Obscured or IR-selected AGN have typically been found to correlate more strongly with SFR than other AGN types (e.g., R. C. Hickox et al. 2009; S. Juneau et al. 2013; M. Azadi et al. 2017), with this population lying typically on or above the star-forming main sequence (e.g., S. Juneau et al. 2013; S. L. Ellison et al. 2016; Y.-Y. Chang et al. 2017; C. Hatcher et al. 2021; P. A. C. Cunha et al. 2025); this aligns well with the trend we see in Figure 22. For X-ray AGN (which only have partial overlap with IR-detected sources; see, e.g., J. L. Donley et al. 2007, 2010, 2012; M. E. Eckart et al. 2010; S. M. LaMassa et al. 2019; C. M. Carroll et al. 2021), a wider diversity of hosts is observed (e.g., R. C. Hickox et al. 2009; J. Aird et al. 2012; J. R. Mullaney et al. 2015; M. Azadi et al. 2015, 2017; H. Suh et al. 2019), but some works have found a connection with enhanced SFR (e.g., J. Florez et al. 2020, 2021; G. Mountrichas et al. 2024), and many X-ray AGN lie on the main sequence (e.g., J. S. Kartaltepe et al. 2012; J. R. Mullaney et al. 2012a; H. Suh et al. 2019). Whilst some theoretical (e.g., P. F. Hopkins et al. 2006) and obser-
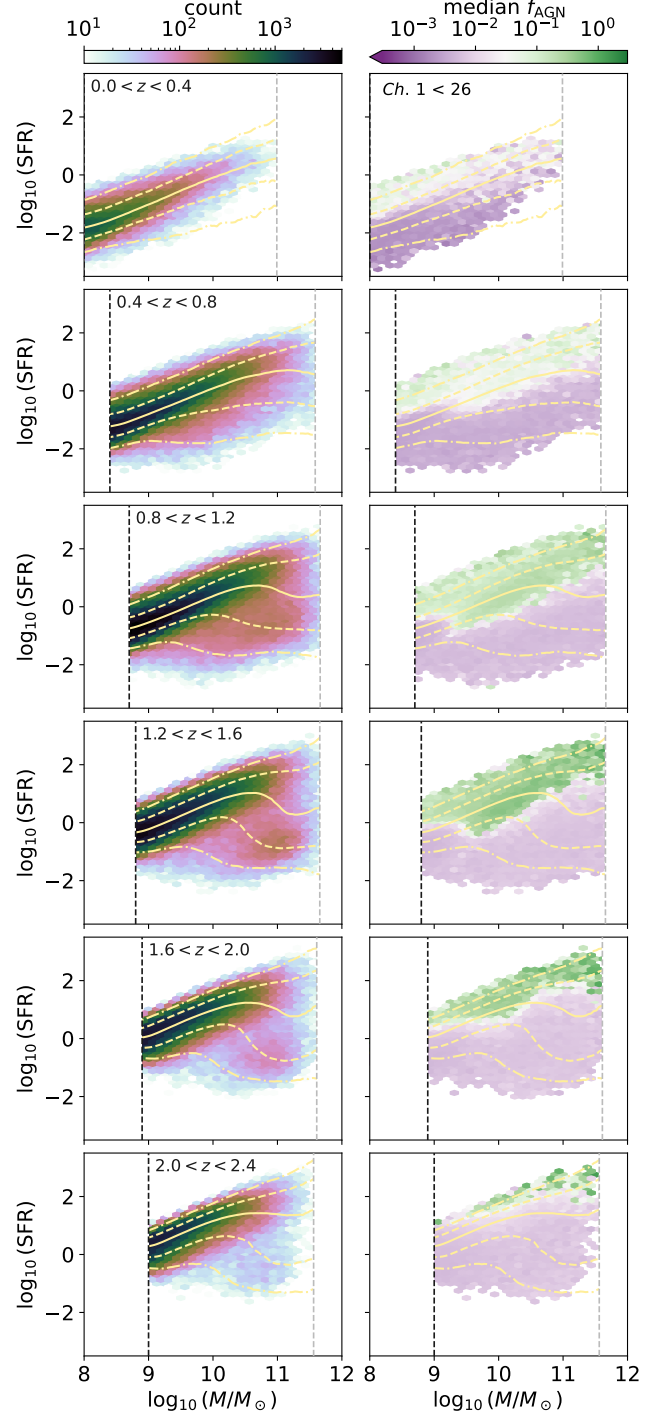
Figure 22. Star formation rate vs. mass binned by redshift and shaded by **(left)** model galaxy count, and **(right)** median AGN bolometric luminosity fraction, f_{AGN}. Cells are shaded only if they contain > 10 model galaxies. The black and gray dashed lines show respectively the mass completeness limit and upper mass limit. Solid yellow lines shows the median from Figure 12, and dashed (dash-dotted) yellow lines show 68 (95)% intervals.

vational (e.g., J. S. Kartaltepe et al. 2012; S. L. Ellison et al. 2016; M. E. Weston et al. 2017; A. D. Goulding et al. 2018; S. L. Ellison et al. 2019) studies have made a connection between galaxy interactions or mergers as driving AGN and star formation (for a review see R. C. Hickox & D. M. Alexander 2018), others have suggested that the correlation is driven by secular processes linked to a common gas reservoir (J. R. Mullaney et al. 2012b,a; T. M. Heckman & P. N. Best 2014; Y.-Y. Chang et al. 2017; Y. S. Dai et al. 2018).

With future work, we will be able to revisit the correlations seen in Figures 20 and 22 in further detail. As highlighted in Appendix H.3, it will be possible to take advantage of recent improvements to FSPS (B. Wang et al. 2025) that refine the joint AGN–galaxy SED model by incorporating UV–optical and broad-line emission from the accretion disk (via, e.g., the model of M. J. Temple et al. 2021b)[24]. Moreover, by folding in data from further in the IR (e.g., the *JWST F770W* photometry in COSMOS2025; M. Shuntov et al. 2025b), it may be possible to probe the rest-frame IR and to trace the relationship between SFR and AGN activity at higher-z.

6. DISCUSSION AND CONCLUSIONS

In this work we have demonstrated that an SPS-based forward model, trained on NUV–MIR photometry from COSMOS2020 with *Spitzer* IRAC *Ch. 1* < 26, is able to faithfully describe the redshift-evolving galaxy population over ~ 12.8 Gyr of cosmic time. To enable population fitting using such deep data, we have developed a new, more flexible, noise modeling approach that uses a conditional diffusion model to predict the photometric uncertainties of an object given its brightness. This approach can accurately represent the complicated, multi-modal uncertainty distributions that arise from variable depths in the COSMOS field. With the trained population model, we are able to rapidly synthesize large mock galaxy catalogs ($\sim 1,000,000$ galaxies/GPU-hr) with realistic multi-band photometry and physical properties. With an efficient SPS emulator, and hardware accelerated MCMC, we are able to perform rapid parameter inference (with a throughput of < 10 GPU-sec/galaxy for full MCMC chains) using the trained model as an informative prior, achieving superior photometric redshift accuracy for $0 < z < 6$.

We validate the learned generative model by comparing it to well-studied scaling relations within galaxy evolution, and explore correlations that are normally extremely difficult to access from photometry alone. Where literature results are available for comparison, we find generally very good agreement. Our stellar mass function agrees well with the expectation from other recent results (e.g. J. Leja et al. 2020; J. R. Weaver et al. 2023b; Euclid Collaboration et al. 2025b), albeit with a tail (comprising $\sim 0.5\%$ of the model galaxy population) at very high stellar masses $\gtrsim 10^{11} M_\odot$. There may be reasons to investigate the extrapolation of the model in this regime further. For instance, recent surveys have made tentative detections of such massive galaxies even at $z \gtrsim 5$. M. Xiao et al. (2024) report three candidate galaxies with masses in excess of $10^{11} M_\odot$ at $5 < z < 6$ in the 124 arcmin^2 FRESCO survey field (P. A. Oesch et al. 2023), one of which has been confirmed and studied extensively (GN10; see e.g., D. A. Riechers et al. 2020; F. Sun et al. 2024; T. Herard-Demanche et al. 2025; G. Lagache et al. 2025). This number is consistent with the expectation from re-scaling a COSMOS-sized mock galaxy catalog from pop-cosmos to a FRESCO-sized field. Our learned star-forming main sequence agrees well with the deep photometric study by J. Leja et al. (2022), with a similar normalization (i.e., relatively low median SFRs compared to earlier work based on parametric SFH models). Other galaxy evolution trends we explore in depth include metallicity scaling relations, dust attenuation properties, and the correlation between AGN activity and SFR.

Taken together, the high fidelity representation of astrophysics and the computational efficiency of the model gives us the ability to model the quantities needed for Stage IV cosmology, with the required level of realism, and at the required throughput. Because the pop-cosmos population model is learned as a distribution over SPS parameters, it is possible to deploy it as a generative model for *any* other survey which has a depth compatible with the COSMOS2020 training sample. In upcoming work, we will demonstrate this with two applications of the model to the KiDS survey (K. Kuijken et al. 2019): B. Leistedt et al. (in prep.) will show the use of the trained model's predictions to determine tomographic redshift distributions, whilst A. Halder et al. (in prep.) will show the use of the model as a prior for inferring KiDS galaxy properties and redshifts at scale.

Complementary approaches are being developed in parallel to meet related goals. C. Hahn et al. (2023a, 2024) and J. Li et al. (2024) have presented a forward modeling approach based on emulated SPS optimized for handling bright, low-redshift galaxies with spectro-photometry from the Dark Energy Spectroscopic Instrument (DESI; C. Hahn et al. 2023b; DESI Collaboration et al. 2024). Large-scale spectroscopic data, such as that provided at low-z by DESI, or that expected in future at high-z from programs such as MOONRISE (R. Maiolino et al. 2020) on the Very Large Telescope's Multi-Object Optical and NIR Spectrograph (MOONS; M. Cirasuolo et al. 2011, 2020), will be important for future extensions

---

[24] Recent studies of blue dust obscured galaxies at intermediate redshifts ("BlueDOGs"; e.g., R. J. Assef et al. 2016, 2020, 2022; A. Noboriguchi et al. 2019, 2022; S. Kim et al. 2025), and "little red dots" at high redshifts (see e.g., G. Barro et al. 2024; J. F. W. Baggen et al. 2024; D. J. Setton et al. 2024; J. Matthee et al. 2024; B. Wang et al. 2025; Y. Ma et al. 2025; D. D. Kocevski et al. 2025) have highlighted the challenges and importance of detailed modeling to separate AGN and SFR signatures (see extensive discussion in S. Kim et al. 2025).

to the `pop-cosmos` model. Such data will offer the possibility of further refining our modeling of emission lines (see, e.g., A. Khederlarian et al. 2024) and gas physics.

In future work, we plan to incorporate the galaxy–halo connection into `pop-cosmos` in order to further improve the utility of the model within cosmological analyses. Conditional abundance matching (for a review see R. H. Wechsler & J. L. Tinker 2018) offers one of the most attractive paths to achieving this. Recent work has shown that sub-halo abundance matching can be expressed elegantly as the solution to an optimal transport problem (S. Fischbacher et al. 2025a). Empirical models such as `UniverseMachine` (P. Behroozi et al. 2019) have demonstrated the predictive power that can be gained from modeling the galaxy–halo connection. The `Diffsky` forward model[25] builds on this foundational work, linking halo mass assembly (A. P. Hearin et al. 2021, 2022) with galaxies' SFHs (A. Alarcon et al. 2023) and SPS within a differentiable framework (A. P. Hearin et al. 2023).

Modeling galaxy sizes and morphologies simultaneously with their other physical parameters is a further extension to `pop-cosmos` that we plan to develop. We will take a featurized approach to this, rather than carrying out full image simulation. The `GalSBI` forward modeling framework (J. Herbel et al. 2017; B. Moser et al. 2024; S. Fischbacher et al. 2025b, 2024a) is a complementary effort that connects a fast simulator of full images (J. Bergé et al. 2013; S. Fischbacher et al. 2024b) to a parametric model for the galaxy population, and a template-based representation of galaxy SEDs. Recently, L. Tortorelli et al. (2025) demonstrated an SPS-based version of this model using an empirically calibrated population model.

In addition to the directions highlighted above, in future iterations of the `pop-cosmos` model, we will revisit our SPS modeling choices to take advantage of the newest advances made in this space (some of the most promising avenues are highlighted in depth in Appendix H). Improvements such as incorporating α-enhancement (e.g., C. M. Byrne et al. 2022, 2025; M. Park et al. 2024), emulated photoionization calculations (e.g., Y. Li et al. 2025; C. Morisset et al. 2025), binary stellar evolution pathways (for a review see J. J. Eldridge & E. R. Stanway 2019), and re-calibrated IGM attenuation (e.g., A. K. Inoue et al. 2014), will be important for meeting the challenges of large spectroscopic datasets (e.g., R. Maiolino et al. 2020; DESI Collaboration et al. 2024), or the redshift frontier being probed by the deepest *JWST* surveys (e.g., M. Shuntov et al. 2025b; for a review see, e.g., B. E. Robertson 2022; J. P. Gardner et al. 2023; E. Merlin et al. 2024).

---

[25] https://github.com/ArgonneCPAC/diffsky

DATA AVAILABILITY

AUTHOR CONTRIBUTIONS

We outline the different contributions below using keywords based on the Contribution Roles Taxonomy (CRediT; A. Brand et al. 2015). **ST:** conceptualization; methodology; software; validation; investigation; data curation; writing – original draft. **HVP:** Conceptualization; methodology; investigation; visualization; validation; supervision; project administration; funding acquisition; writing – review and editing. **GJ:** methodology; software; validation. **SD:** methodology; validation; investigation; writing – review and editing. **JA:** conceptualization; methodology; software; investigation. **BL:** methodology; validation; investigation; writing – review and editing. **DJM:** methodology; investigation; visualization; writing – review and editing. **AH:** validation. **JL:** methodology; writing – review and editing.

ACKNOWLEDGMENTS

---

[26] https://github.com/Cosmo-Pop

<div align="center">APPENDIX</div>

A. CALIBRATION PARAMETERS

Our model has a large number of calibration parameters which need to be included to ensure the galaxy parameters of interested are not biased. Table 5 lists the set of band-dependent parameters (f_b, α_{ZP}), and Table 6 lists the emission line-dependent parameters $(\beta_{EM}, \gamma_{EM})$. The tabulated values of α_{ZP}, β_{EM}, and γ_{EM} are optimized during the population model training process described in Section 3.5. The values of f_b are fixed and determined using the method outlined in Section 2. In Table 5, the f_b values are quoted in nano-maggies (i.e. $10^9 \times$ flux relative to a standard source; e.g., C. Stoughton et al. 2002; K. Abazajian et al. 2003), assuming an AB standard source with a flux of 3631 Jy. The effective depths these correspond to are logarithmic AB magnitudes, i.e. eff. depth $= 22.5 - 2.5 \log_{10}(f_b/\text{nMgy})$. The emission line strength corrections (β_{EM}) and standard deviations (γ_{EM}) in Table 6 are dimensionless fractional quantities.

In Figure 23, we show the effective change in emission line strength implied by the values of β_{EM} in Table 6, relative to the earlier calibrations by L23 and A24. The mean flux of any line will be given by $(1+\beta_{EM}) \times F_{EM}(\boldsymbol{\vartheta})$, where $F_{EM}(\boldsymbol{\vartheta})$ is the uncorrected line strength from the N. Byler et al. (2017) Cloudy model grid in FSPS. The pop-cosmos line strengths from this work relative to L23 will thus be given by $(1+\beta_{EM})/(1+\beta_{EM}^{L23})$, and similarly for this work relative to A24. For the 44 lines in Table 6, the effective change relative to L23 is within $\pm 3\%$. The pair of doubly-forbidden sulphur lines, [S III] 9071 and [S III] 9533, are omitted from Figure 23, as these are essentially negated by their inferred values of $\beta_{EM} = -1$ (for L23, A24, and this work).

Figure 24 shows the effective change in photometric flux implied by our zero-point corrections, $\boldsymbol{\alpha}_{ZP}$, relative to the inferences from L23 and A24. For all bands, our corrections leave the fluxes within $\pm 2\%$ of the L23 and A24 results. The absolute values of the zero-point corrections we infer in Table 5 are typically significantly smaller than the offsets estimated by R. E. Skelton et al. (2014), where our sets of passbands overlap, reflecting the progress that has been made in both data and modeling in the past decade.

B. AMBIGUOUS CROSS-MATCHES IN THE COSMOS SPECTROSCOPIC ARCHIVE

From the unique objects in the A. A. Khostovan et al. (2025) spectroscopic archive, we find 39,702 that have a CL $\geq 50\%$ redshift, and a COSMOS2020 Farmer ID that is within our $Ch. 1 < 26$ catalog. We find that these 39,588 sources correspond to 39,579 unique COS-

[27] https://github.com/justinalsing/affine
[28] https://github.com/Cosmo-Pop/flowfusion
[29] https://github.com/Cosmo-Pop/pop-cosmos
[30] https://github.com/stevet40/hist_contour
[31] https://github.com/jackgoffinet/ffjord-lite

Table 5. Band-dependent calibration parameters.

Band	f_b (nMgy)	Eff. Depth (mag)	α_{ZP}
u	0.001080	29.916840	1.009450
g	0.003085	28.776844	1.085712
r	0.005081	28.235106	1.075836
i	0.000879	30.139624	1.019943
z	0.002913	28.839199	1.023337
y	0.004059	28.478890	1.048863
Y	0.002624	28.952726	0.999337
J	0.004989	28.254908	1.004181
H	0.003334	28.692439	0.970161
K_S	0.001146	29.851584	1.042268
$Ch.\ 1$	0.003229	28.727212	0.960127
$Ch.\ 2$	0.003452	28.654855	0.921076
IB427	0.005263	28.196843	0.976578
IB464	0.003999	28.495201	0.972594
IA484	0.006620	27.947804	1.020441
IB505	0.004718	28.315492	1.008716
IA527	0.006163	28.025453	0.994083
IB574	0.001148	29.850477	0.929141
IA624	0.004297	28.417198	1.012976
IA679	0.006221	28.015279	1.131825
IB709	0.001822	29.348750	0.976940
IA738	0.004118	28.463242	0.969744
IA767	0.008231	27.711334	0.965354
IB827	0.008909	27.625397	0.920645
NB711	0.010834	27.413028	0.987654
NB816	0.007929	27.752018	0.929771

Table 6. Emission line-dependent calibration parameters.

Line	λ_{EM} (Å)	β_{EM}	γ_{EM}
C II] 2326	2326.11	1.336×10^{-3}	1.425×10^{-13}
[O III] 2321	2321.66	2.927×10^{-3}	1.424×10^{-13}
[O I] 6302	6302.05	8.403×10^{-3}	1.464×10^{-13}
[S II] 4070	4069.75	-8.081×10^{-3}	1.433×10^{-13}
H I (Ly-α)	1215.67	6.178×10^{-3}	1.455×10^{-13}
[Al II] 2670	2669.95	2.853×10^{-3}	1.425×10^{-13}
[Ar III] 7753	7753.19	-1.335×10^{-2}	1.427×10^{-13}
H I (Pa-7)	9017.80	-1.219×10^{-2}	1.425×10^{-13}
[Al II] 2660	2661.15	2.905×10^{-3}	1.425×10^{-13}
[S III] 6314	6313.81	4.801×10^{-3}	1.426×10^{-13}
H I (Pa-6)	9232.20	-1.025×10^{-2}	1.426×10^{-13}
[S III] 3723	3722.75	-4.322×10^{-3}	1.425×10^{-13}
Mg II 2800	2803.53	-6.834×10^{-3}	1.434×10^{-13}
H I (Pa-5)	9548.80	-8.112×10^{-3}	1.427×10^{-13}
He I 7065	7067.14	2.562×10^{-3}	1.427×10^{-13}
[N II] 6549	6549.86	1.167×10^{-2}	2.163×10^{-13}
[S II] 6732	6732.67	9.945×10^{-3}	3.600×10^{-13}
C III]	1908.73	1.850×10^{-3}	1.424×10^{-13}
He I 6680	6679.99	9.079×10^{-3}	1.437×10^{-13}
Mg II 2800	2796.35	-5.579×10^{-3}	1.460×10^{-13}
[S II] 6717	6718.29	8.977×10^{-3}	1.028×10^{-13}
[Ar III] 7138	7137.77	2.573×10^{-3}	1.490×10^{-13}
[C III]	1906.68	1.905×10^{-3}	1.425×10^{-13}
He I 4472	4472.73	4.407×10^{-3}	1.437×10^{-13}
[O III] 4364	4364.44	2.081×10^{-3}	1.425×10^{-13}
[N II] 6585	6585.27	-5.913×10^{-1}	1.000×10^{-13}
[S III] 9071	9071.10	-1.011	1.702×10^{-13}
H-8 3798	3798.99	-5.927×10^{-4}	1.465×10^{-13}
He I 3889	3889.75	-6.603×10^{-3}	1.500×10^{-13}
H-7 3835	3836.49	-1.916×10^{-3}	1.485×10^{-13}
[Ne III] 3968	3968.59	-1.005×10^{-2}	1.448×10^{-13}
He I 5877	5877.25	3.240×10^{-1}	1.555×10^{-13}
H-6 3889	3890.17	-7.012×10^{-3}	1.536×10^{-13}
[S III] 9533	9533.20	-1.008	2.017×10^{-13}
H-5 3970	3971.20	-1.504×10^{-1}	1.689×10^{-13}
[O II] 3726	3727.10	4.595×10^{-3}	1.000×10^{-13}
H-δ 4102	4102.89	-5.235×10^{-1}	1.912×10^{-13}
[O II] 3729	3729.86	2.642×10^{-1}	1.000×10^{-13}
[Ne III] 3870	3869.86	-8.795×10^{-2}	1.889×10^{-13}
H-γ 4340	4341.69	-3.297×10^{-1}	1.013×10^{-13}
[O III] 4960	4960.30	2.037×10^{-4}	1.000×10^{-13}
H-β 4861	4862.71	-5.577×10^{-1}	1.000×10^{-13}
H-α 6563	6564.60	-3.314×10^{-1}	1.000×10^{-13}
[O III] 5007	5008.24	1.160×10^{-1}	2.642×10^{-2}

MOS2020 `Farmer` IDs. There are thus 9 of our selected `Farmer` IDs (30647, 58379, 117181, 159826, 189699, 257074, 656499, 942331, 948097) that have two spectra associated with them in A. A. Khostovan et al. (2025). One of these (257074) has two spectra with distinct COSMOS2020 `Classic` IDs (576772, 577956) associated with them, but for the other eight there is one spectrum with a `Classic` source match and one with no `Classic` counterpart. The two spectra / `Classic` sources associated with `Farmer` ID 257074 have distinct spectroscopic redshifts (both from PRIMUS; A. L. Coil et al. 2011): $z = 0.62133$ and $z = 0.1717$, with CL = 80% and CL = 50% respectively. For `Farmer` ID 30647, the two spectroscopic redshifts are substantially different: $z = 1.51598$ (from DESI EDR; DESI Collaboration et al. 2024) and $z = 5.621$ (from 10k–DEIMOS; G. Hasinger et al. 2018), both with CL = 80%. Only the former of these has an associated `Classic` ID in A. A. Khostovan et al. (2025).

Figure 23. Effective change in predicted emission line strength for pop-cosmos relative to the line strengths inferred in L23 and A24.
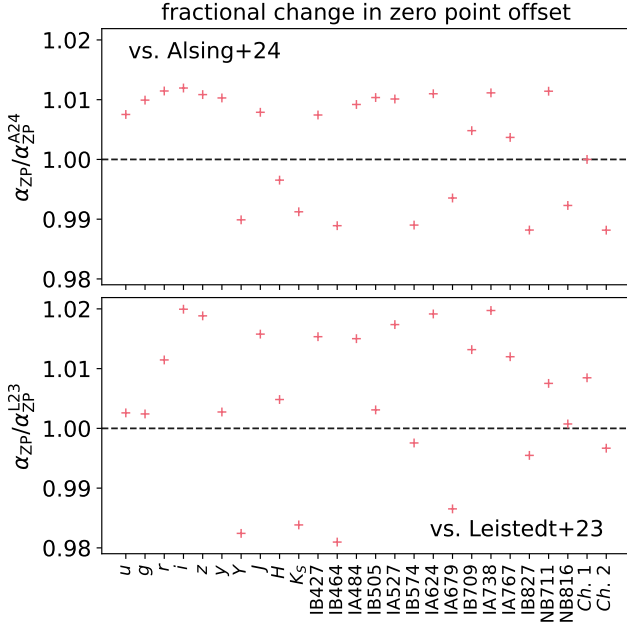
Figure 24. Effective change in photometric fluxes in the the 26 COSMOS bands, relative to the corrections inferred in L23 and A24.

C. DERIVED PARAMETERS

In this Appendix, we describe the computation of (specific) star formation rate and mass-weighted age given the base SPS parameters in the upper part of Table 2. For a model galaxy at redshift z, we compute the age of the Universe at that redshift, $t_{\mathrm{Univ}}(z)$, under a flat ΛCDM cosmology with Planck Collaboration et al.

(2020) parameters. The SFH for a galaxy is piecewise constant and follows the sequence of 7 lookback time bins from J. Leja et al. (2019a,b). The first two bin edges are fixed at $t_{\mathrm{edge},1} = 0$ Myr, and $t_{\mathrm{edge},2} = 30$ Myr, with the last placed at $t^8_{\mathrm{edge}} = t_{\mathrm{Univ}}(z)$. The five inner edges are spaced logarithmically (P. Ocvirk et al. 2006) between 100 Myr and $0.85 \times t_{\mathrm{Univ}}(z)$, so the ith edge is given by

$$t_{\mathrm{edge},i} = 100\,\mathrm{Myr} + \frac{i-3}{4} \times [0.85 \times t_{\mathrm{Univ}}(z) - 100\,\mathrm{Myr}] \tag{C1}$$

for $i = 3, \ldots, 7$.

The SFR in bin j, relative to the first (most recent) bin can be computed from the SFR ratios as

$$\log_{10}\left(\frac{\mathrm{SFR}_j}{\mathrm{SFR}_1}\right) = -\sum_{k=1}^{j-1} \Delta \log_{10}(\mathrm{SFR})_k \tag{C2}$$

for $j = 2, \ldots 7$. We can then define a normalized SFR for bin j,

$$\frac{\mathrm{SFR}_j}{M_{\mathrm{form}}} = \frac{\mathrm{SFR}_j/\mathrm{SFR}_1}{\Delta t_1 + \sum_{k=2}^{7}(\Delta t_k \times \mathrm{SFR}_k/\mathrm{SFR}_1)}$$
$$= \frac{\mathrm{SFR}_j}{\sum_{k=1}^{7}(\Delta t_k \times \mathrm{SFR}_k)}, \tag{C3}$$

where $\Delta t_j = t_{\mathrm{edge},j+1} - t_{\mathrm{edge},j}$ is the width of the jth bin, and M_{form} is the total mass the galaxy will form by redshift z. The fraction of stellar mass formed in bin j is then simply given by $M_{\mathrm{form},j}/M_{\mathrm{form}} = \Delta t_j \times \mathrm{SFR}_j/M_{\mathrm{form}}$.

From this, the SFR per unit mass formed, averaged over the most recent 100 Myr of a galaxy's life can be computed by summing the fractional mass formed over the first two bins of the SFH and dividing by 100 Myr:

$$\frac{\mathrm{SFR}}{M_{\mathrm{form}}} = \left(\frac{M_{\mathrm{form},1}}{M_{\mathrm{form}}} + \frac{M_{\mathrm{form},2}}{M_{\mathrm{form}}}\right)/100\,\mathrm{Myr}. \tag{C4}$$

We compute the SFR by multiplying Equation C4 by M_{form}. The sSFR values we quote throughout the paper follow the conventional definition of SFR per unit mass remaining. This is found by dividing Equation C4 by the fraction of stellar mass remaining, M/M_{form}, where this quantity is emulated as described in Section 3.2.

The normalized SFR defined in Equation C3 sets the fractional contribution to a galaxy's total stellar content of the stars formed at any given instant. In effect, it gives the weighting of all possible stellar ages within the galaxy's SFH. As the SFR is piecewise constant, taking the weighted mean of stellar age across the full SFH can be done by summing the weighted mean in each bin:

$$t_{\mathrm{age}} = \sum_{j=1}^{7}\left[\int_{t_{\mathrm{edge},j}}^{t_{\mathrm{edge},j+1}} \frac{\mathrm{SFR}_j}{M_{\mathrm{form}}} \times t\,\mathrm{d}t\right]$$
$$= \sum_{j=1}^{7}\left[\frac{1}{2}(t^2_{\mathrm{edge},j+1} - t^2_{\mathrm{edge},j}) \times \frac{\mathrm{SFR}_j}{M_{\mathrm{form}}}\right]. \tag{C5}$$
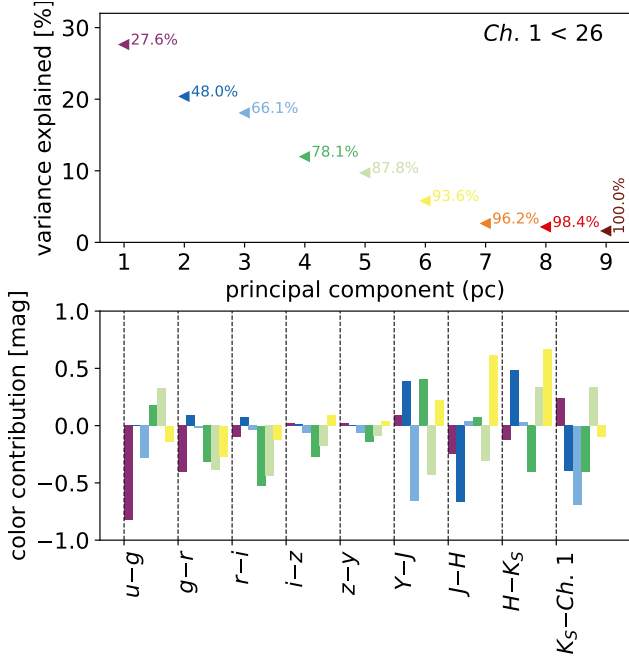
Figure 25. Principal components (PCs) of the broadband colors of the COSMOS2020 catalog with *Ch. 1 < 26*, showing **(upper panel)** the eigenvalues, and **(lower panel)** the principal eigenvectors. The vertical axis in the upper panel shows the fraction of variance explained by each PC, with the annotation showing the cumulative variance explained by each PC and those leftward of it in the plot. The lower panel shows the contribution of each color to the first 6 PCs, which cumulatively explain > 90% of the variance in COSMOS2020.

D. FURTHER PHOTOMETRIC VALIDATION

In Figure 25, we follow S. Thorp et al. (2025b) in performing a principal component analysis (PCA) of the 9 broadband colors defined by adjacent passbands. Shown are the principal components (PCs) of the COSMOS2020 catalog itself. The upper panel shows a scree plot (R. B. Cattell 1966) ranking the PCs by the fraction of total variance explained, and with labels showing the cumulative variance explained. We see that 6 PCs explains 93.6% of the total variance in the COSMOS2020 broadband colors, so we will focus on these going forward. In the lower panel of Figure 25, we visualize the principal eigenvectors corresponding to each PC, with a bar plot showing the contribution of each color to the first 6 eigenvectors. We see that the first PC (explaining $\sim 28\%$ of total variance) is dominated by $u - g$ and $g - r$, with the second and third (both explaining $\sim 20\%$ of variance each) being IR-dominated.

In Figure 26, we project the COSMOS2020 and `pop-cosmos` colors along the principal axes identified in Figure 25. We then perform a comparison of the projected data and model colors using quantile–quantile (Q–Q) and probability–probability (P–P) plots (M. B.
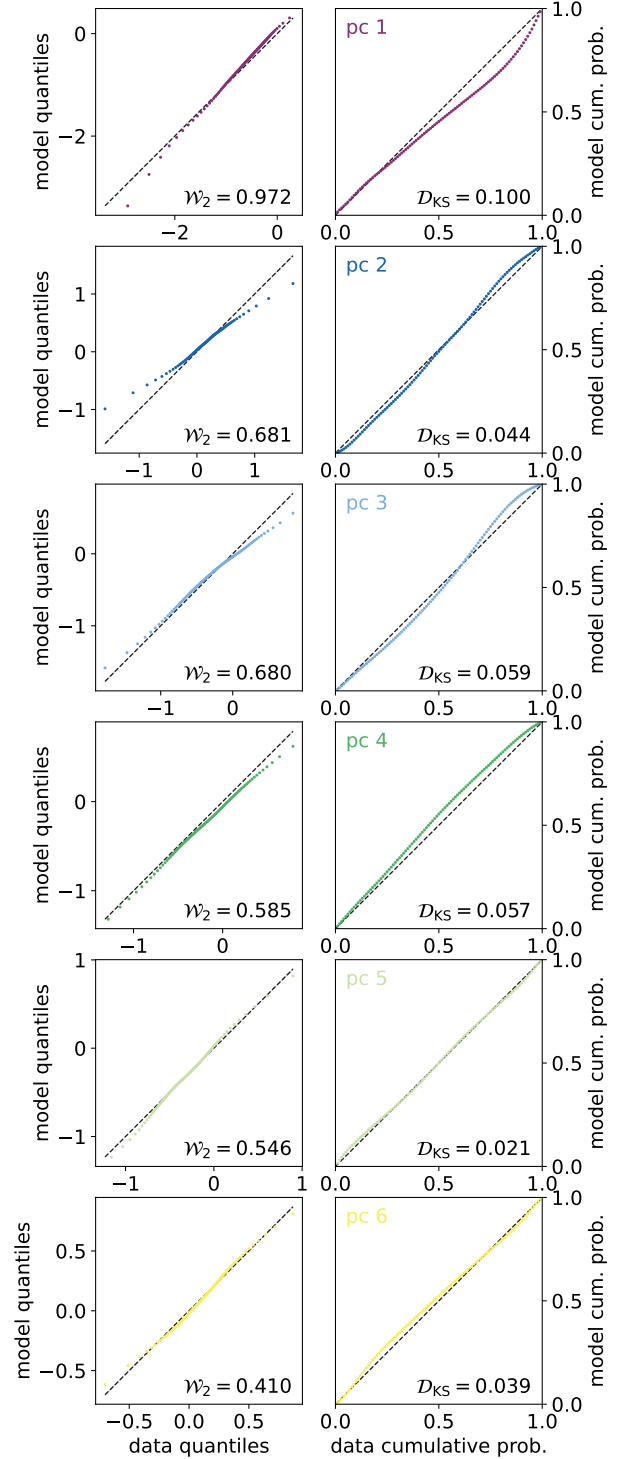
Figure 26. Quantile-based comparison — **(left)** Q–Q plots, and **(right)** P–P plots — of the model colors and COSMOS colors projected along the first six principal axes (93% of total variance) identified in Figure 25. On the lower right corners of the Q–Q plots, we show the marginal 2-Wasserstein distance $\mathcal{W}_2$; on the lower right corner of the P–P plots we show the K–S statistic $\mathcal{D}_{\mathrm{KS}}$.

Wilk & R. Gnanadesikan 1968). As a numerical summary of discrepancy in quantile space, we quote the 2-Wasserstein distance $\mathcal{W}_2$, equal to the Euclidian distance between the data and model quantile functions (see e.g. A. Munk & C. Czado 1998; E. Levina & P. Bickel 2001; A. Ramdas et al. 2017; G. Peyré & M. Cuturi 2019). In cumulative probability space, we use the K–S statistic $\mathcal{D}_{KS}$ as a numerical discrepancy measure, with this being equal to the maximal deviation in the P–P plot (for further discussion, see S. Thorp et al. 2025b, and references therein). The agreement between model and data is generally good along all six directions, but with some mismatches. For the first PC, we see in the P–P plot a mismatch in the right-hand part of the distribution, suggesting that the model has more probability density in this direction than the data, implying bluer $u-g$ and $g-r$ colors than are seen in COS-MOS2020 (see also Figure 2). The level of disagreement between model and data in the direction of these two colors was also largest in S. Thorp et al. (2025b); here the maximum deviation in the P–P plot is slightly larger ($\mathcal{D}_{KS} = 0.100$; compared to $\mathcal{D}_{KS} = 0.055$ there). The second and third PCs show similar mismatches to each other, suggestive of the model being under-dispersed relative to the data in these directions. For PC2, which is dominated by $Y - J$ and $K_S - Ch.\,1$, this underdispersion appears fairly symmetric, whereas for PC3 it is stronger in the right tail (which corresponds to red $Y - J$ and $H - K_S$, and blue $J - H$ and $K_S - Ch.\,1$). For PC2 the K–S statistic, $\mathcal{D}_{KS} = 0.044$, is relatively low compared to the 2-Wasserstein distance, $\mathcal{W}_2 = 0.681$, suggesting that the distribution is most discrepant in the tails (which the Q–Q plot/Wasserstein distance are most sensitive to). In the color marginals (Figure 2, the $K_S - Ch.\,1$ color is visibly very discrepant in the blue tail, as was the case with the A24 model. As discussed in Section 4.1, this color shows the largest systematic difference between the different photometric extraction methods deployed in J. R. Weaver et al. (2022).

The fourth, fifth, and sixth PCs each explain $\sim 10\%$ of the total variance in COSMOS2020. For PC4, we see a consistent small bias in the negative direction, stronger towards the right tail of the distribution (c.f. the skew-normal in fig. 3 of S. Thorp et al. 2025b). The closest match between model and data is seen for PC5, which corresponds to an even mixture of all colors that would imply a blue tilt in the SED between g- and H-band. A good match is also seen in PC6, which corresponds to a positive slope in the SED between the i- and K_S-bands.

E. REDSHIFT DEPENDENT MASS LIMITS

In this Appendix, we define fitting formulae for the upper (Section E.1) and lower (Section E.2) limits of stellar mass where we expect pop-cosmos to be most reliable. These limits can be used to select pop-cosmos model galaxies that ought to be well represented in the COSMOS data itself (based on the $Ch.\,1 < 26$ limit).
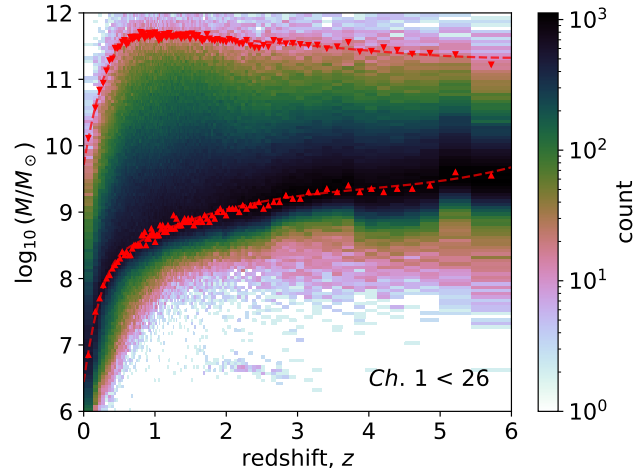
Figure 27. Histogram of stellar mass vs. redshift for pop-cosmos with $Ch.\,1 < 26$. Redshift binning is based on percentiles of z, so each vertical slice of the histogram contains 20,000 mock galaxies. Red markers show the redshift binned turnover of the mass function (upward pointing triangles), and the 99.5th percentile in stellar mass (downward pointing triangles). Red dashed lines show the fitting functions of the form $\mathcal{M}(z)$ based on the results in Section E.1–E.2.

The upper limit in Section E.1 is designed to mask very massive galaxies with $M \gtrsim 10^{11} M_\odot$. Due to the size of the COSMOS field, these galaxies are not well represented in the pop-cosmos training data. The lower limit in Section E.2 is based on the estimated completeness of our model in stellar mass.

For both mass limits, we define a fitting function, $\mathcal{M}(z) = \log_{10}[M_{\lim}(z)/M_\odot]$, using a three-knot cubic Hermite spline. This is defined by the redshifts (z_0, z_1, z_2), log stellar masses $(\mathcal{M}_0, \mathcal{M}_1, \mathcal{M}_2)$, and gradients (s_0, s_1, s_2) of $\mathcal{M}(z)$ at three positions. The general form of the fitting function for $z_k \leq z < z_{k+1}$ is

$$\begin{aligned}
\mathcal{M}(z) = \mathcal{M}_k a(t_k) &+ \mathcal{M}_{k+1} b(t_k) \\
&+ (z_{k+1} - z_k) s_k c(t_k) \\
&+ (z_{k+1} - z_k) s_{k+1} d(t_k)
\end{aligned} \tag{E6}$$

where $t_k = (z - z_k)/(z_{k+1} - z_k)$, and

$$a(t) = 2t^3 - 3t^2 + 1 \tag{E7}$$
$$b(t) = -2t^3 + 3t^2 \tag{E8}$$
$$c(t) = t^3 - 2t^2 + t \tag{E9}$$
$$d(t) = t^3 - t^2, \tag{E10}$$

are the Hermite basis functions. Galaxies that lie beyond the limits defined by our fitted $\mathcal{M}(z)$ are included in the public catalogs, but are flagged. The fits described in Section E.1–E.2 are shown graphically in Figure 27.
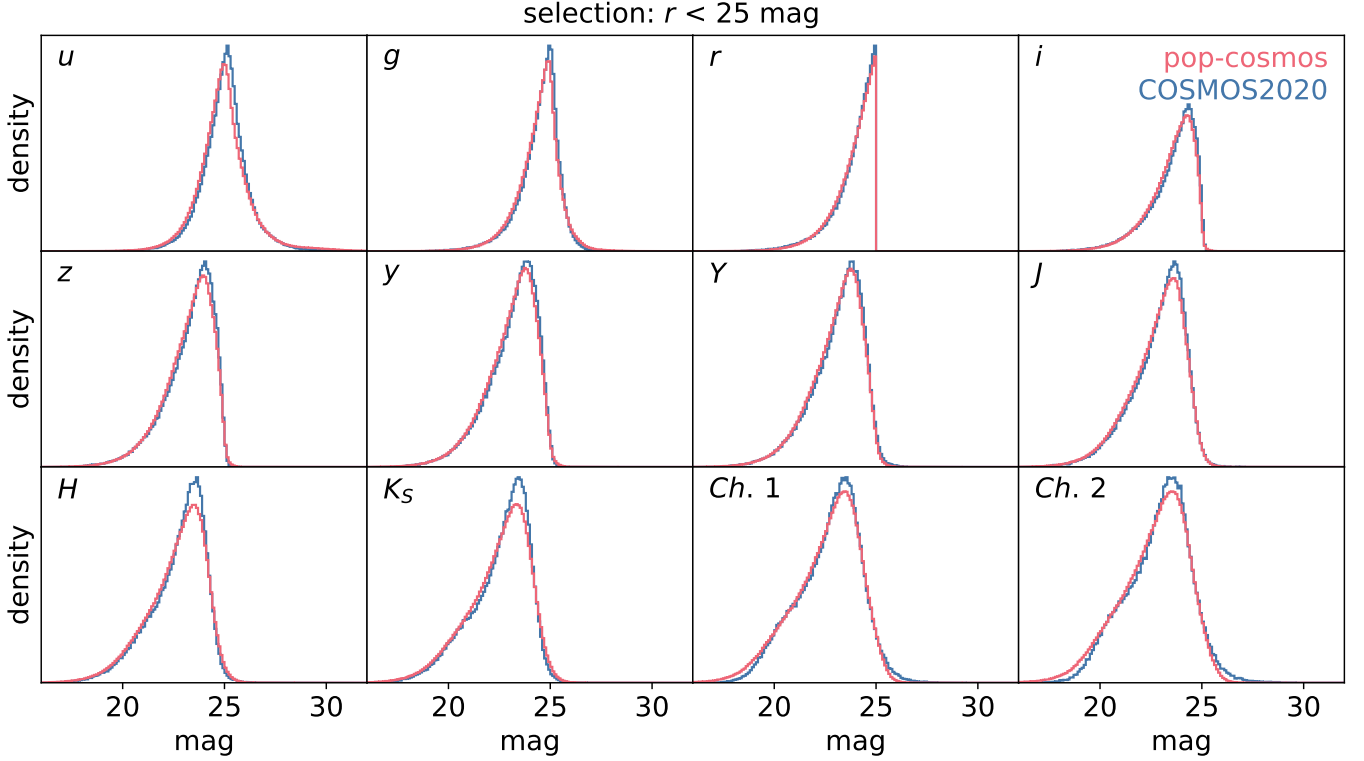
Figure 28. Broadband magnitudes (logarithmic) predicted by the new `pop-cosmos` generative model, compared to COS-MOS2020 (J. R. Weaver et al. 2022) for $r < 25$.

E.1. *Mask for Over-Massive Galaxies*

We find that in the extreme tails of the trained `pop-cosmos` model, there are mock galaxies more massive than almost any that are found in the COSMOS field (e.g., via the `LePhare` fits from J. R. Weaver et al. 2022, 2023b; the individual-galaxy posterior inferences from T24; or the updated individual-galaxy inferences in Section 4.4). We find that a mask for the `pop-cosmos` model draws based on the redshift-dependent 99.5th percentile of stellar mass is a reasonable cut for removing these. We divide our mock catalog of two million galaxies into 100 redshift bins each containing 1% of the catalog. For each bin, we compute the 99.5th percentile of $\log_{10}(M/M_\odot)$.

We fit a two-segment cubic Hermite spline $\mathcal{M}(z)$ to these binned 99.5th percentiles via least squares. We fix $z_0 \equiv 0$, $z_2 \equiv 6$, $s_1 \equiv s_2 \equiv 0$, and optimize the remaining parameters. The resulting fit has $z_1 = 0.7948$, $\mathcal{M}_0 = 9.7108$, $\mathcal{M}_1 = 11.6535$, $\mathcal{M}_\in = 11.3251$, and $s_0 = 6.4555$. This is shown as a gray dashed line on Figure 8, and the upper red dashed line on Figure 27.

E.2. *Mask for Mass Incompleteness*

We estimate a mass completeness limit for our mock galaxy catalogs by computing the rolling turnover of the stellar mass function in the same 100 redshift bins as in Section E.1. We estimate the turnover based on the mode of the mass distribution computed with a bin size

of $\Delta \log_{10}(M/M_\odot) = 0.05$. We perform a least squares fit of $\mathcal{M}(z)$ to this estimated turnover, with $z_0 \equiv 0$ and $z_2 \equiv 6$ fixed. The resulting fit has $z_1 = 0.7404$, $\mathcal{M}_0 = 6.5184$, $\mathcal{M}_1 = 8.5429$, $\mathcal{M}_\in = 9.6966$, $s_0 = 6.9873$, $s_1 = 0.5549$, and $s_2 = 0.2956$. The completeness function is overplotted on Figures 8 and 27 with these parameters. We find that the resulting curve for $0.2 \lesssim z \leq 6.0$ is within ± 0.3 dex of the total mass completeness curve, $-3.23 \times 10^7 (1+z) + 7.83 \times 10^7 (1+z)^2$, provided by J. R. Weaver et al. (2023b) for COSMOS2020. For $z \lesssim 0.2$, we find our stellar mass function turns over at lower masses than the J. R. Weaver et al. (2023b) curve by up to 0.8 dex.

F. FURTHER COLOR AND MAGNITUDE DISTRIBUTIONS

F.1. *Broadband Colors and Magnitudes for $r < 25$*

Figures 28, and 29 show respectively the broadband magnitudes and colors predicted by the trained model for $r < 25$. For this r-band limited mock catalog, the model magnitudes are in excellent agreement with the COSMOS2020 `Farmer` magnitudes. The $r < 25$ agrees better with COSMOS2020 for the NIR and MIR colors than for the optical colors. Figure 30 shows 2D color vs. color plots for the color pairs used by A24.
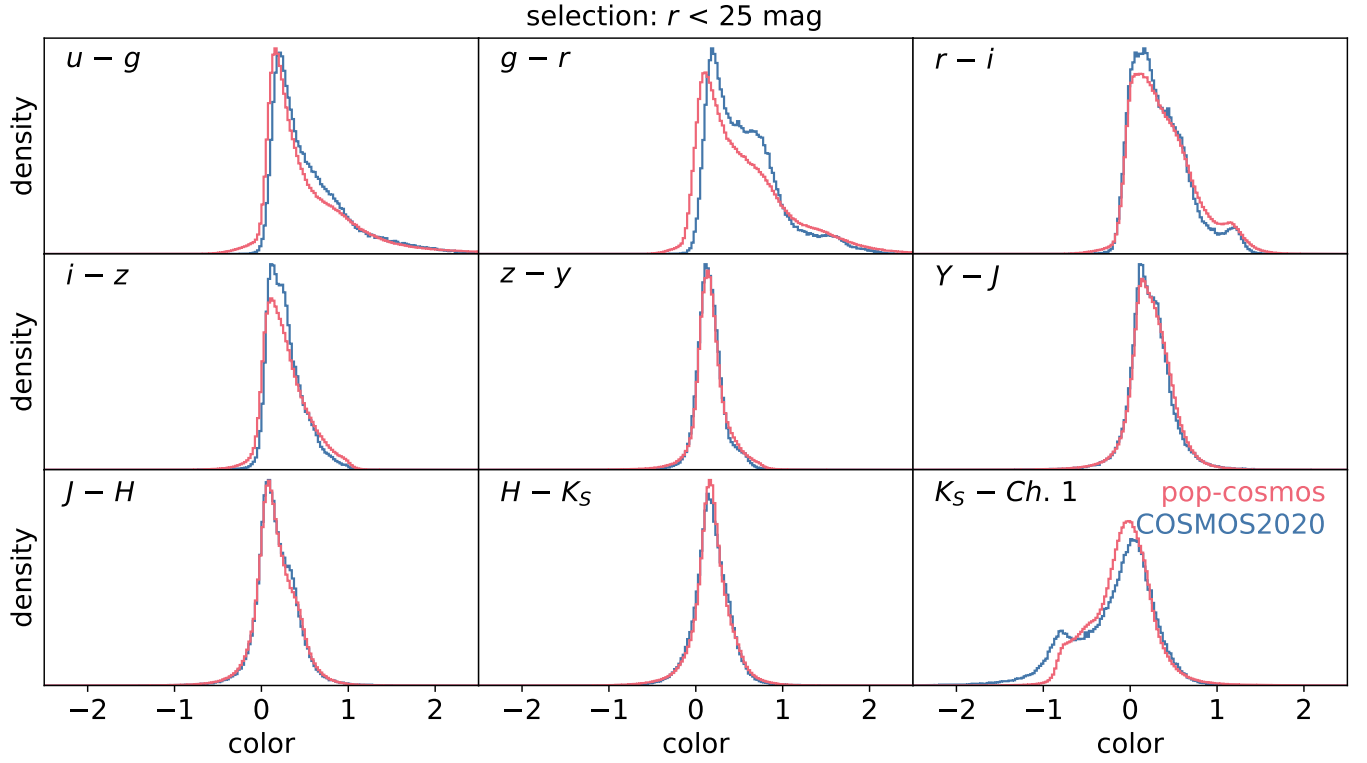
Figure 29. Broadband colors (logarithmic, adjacent bands) predicted by the new `pop-cosmos` generative model, compared to COSMOS2020 (J. R. Weaver et al. 2022) for $r < 25$.
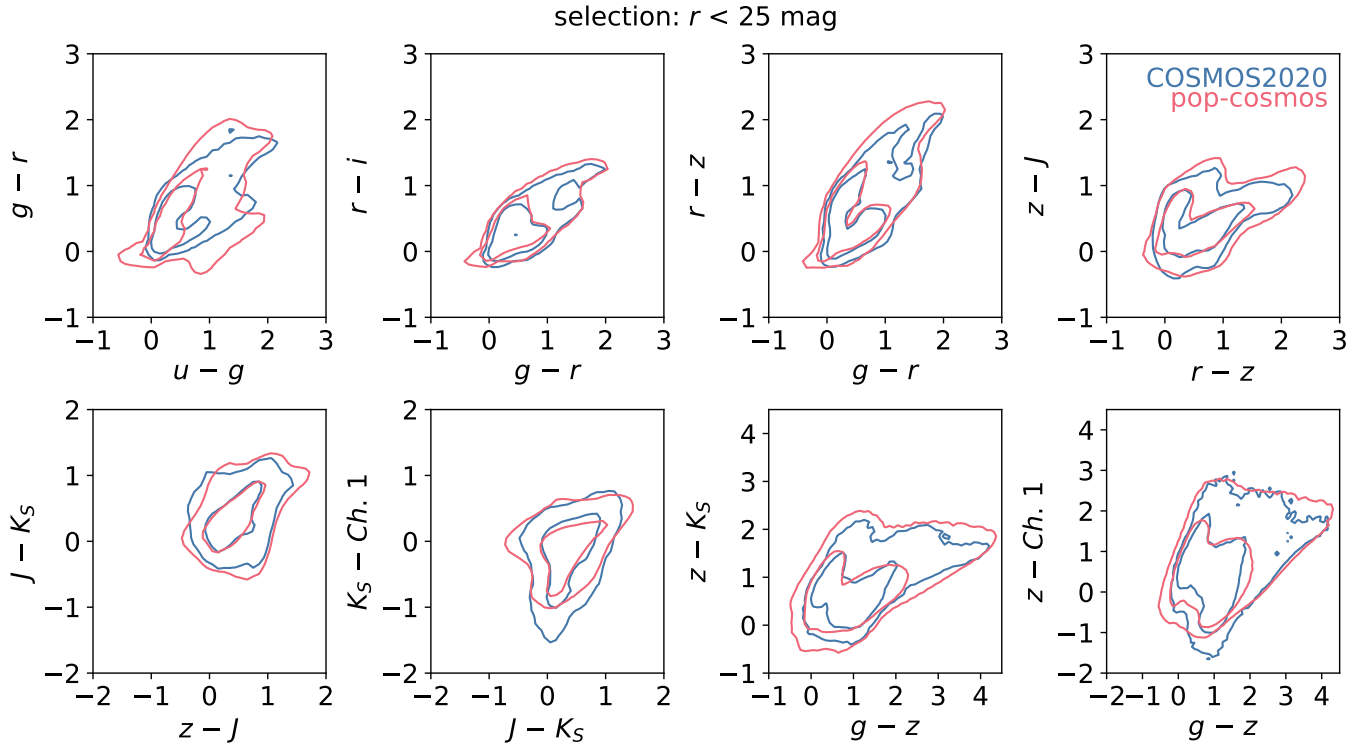
Figure 30. Broadband colors (logarithmic, A24 band combinations) predicted by the new `pop-cosmos` generative model, compared to COSMOS2020 (J. R. Weaver et al. 2022) for $r < 25$. Contours enclose the 68 and 95% highest density regions.
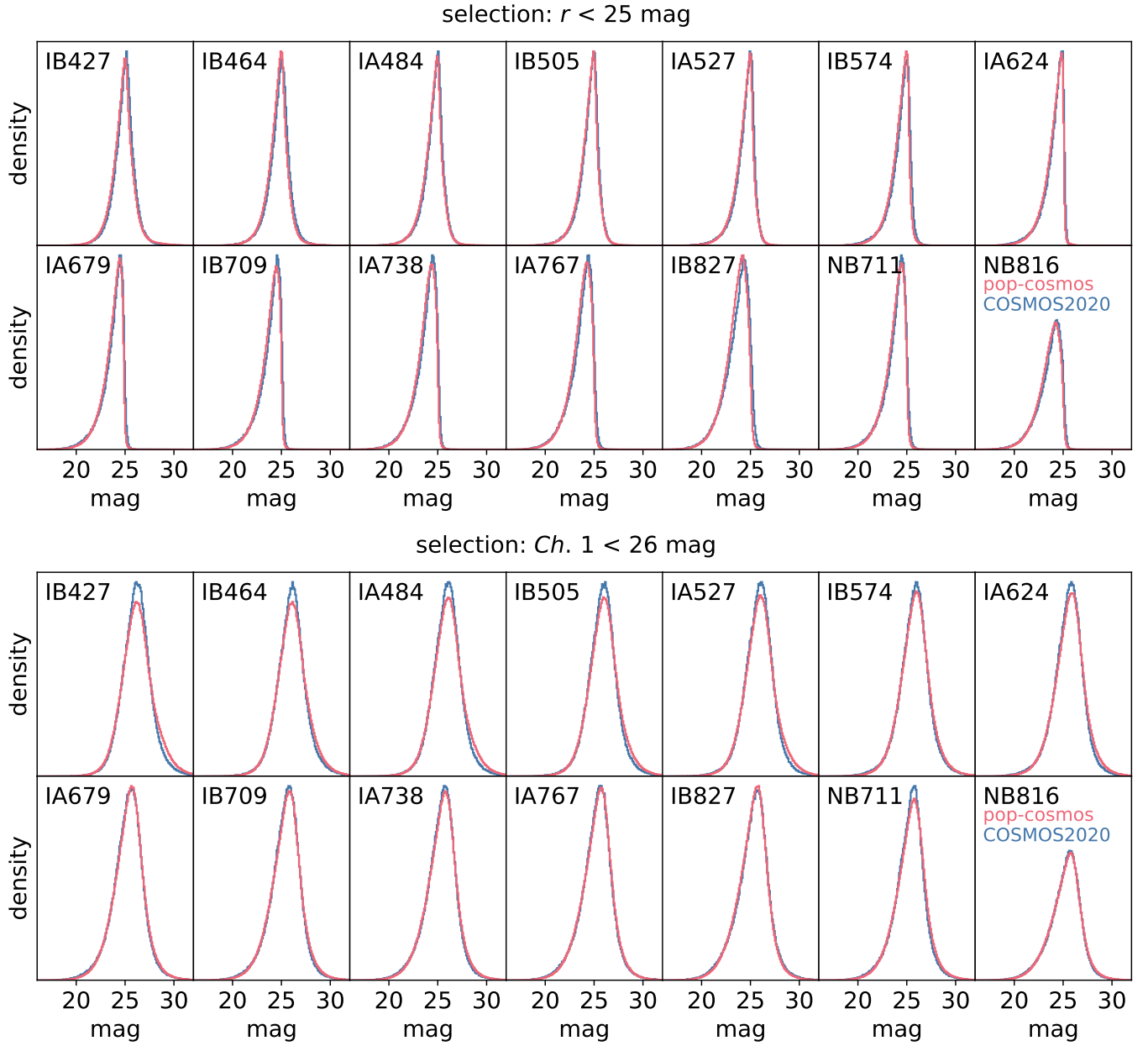
Figure 31. Narrow- and intermediate-band magnitudes from `pop-cosmos` vs. COSMOS2020 (J. R. Weaver et al. 2022).
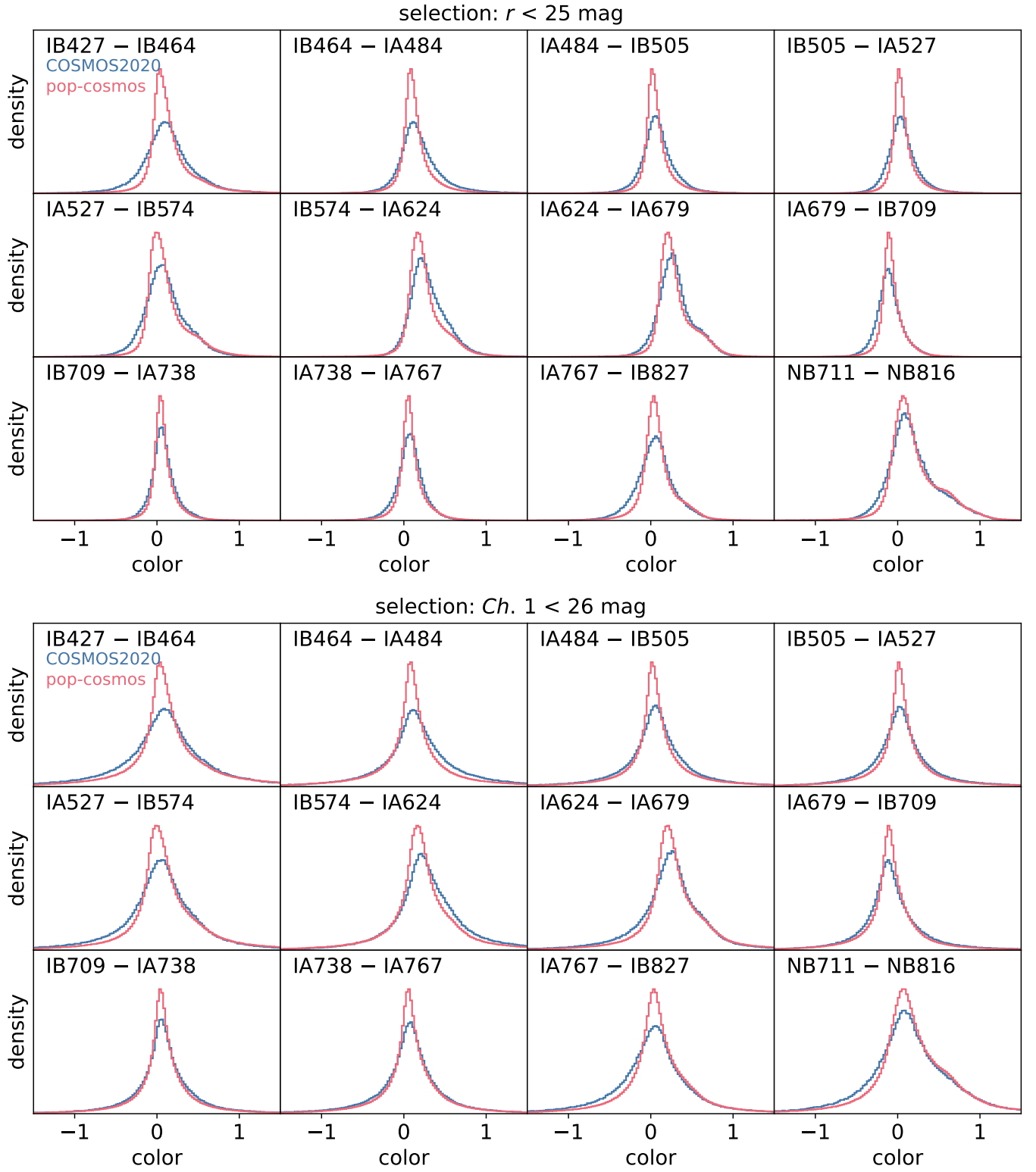
Figure 32. Narrow- and intermediate-band colors from `pop-cosmos` compared to COSMOS2020 (J. R. Weaver et al. 2022).

F.2. *Narrowband Colors and Magnitudes*

Figure 31 shows the marginal distributions of the 14 narrow- and intermediate-band magnitudes predicted by `pop-cosmos` for the *r*-band- and *Ch.1*-selected mock catalogs, compared to the COSMOS2020 catalog (J. R. Weaver et al. 2022). Figure 32 shows a set of 11 colors constructed from adjacent pairs of intermediate bands, and the pair of narrow bands. The magnitudes show excellent agreement for both selections. The `pop-cosmos` colors tend to be slightly under-dispersed relative to COSMOS2020, but the central tendency and widths of the distributions are in close agreement.

F.3. *Uncertainty Model Predictions for $r < 25$*

In Figure 33, we show the predictions made by the uncertainty model for the $r < 25$ mock galaxy catalog, compared to the COSMOS2020 `Farmer` catalog (J. R. Weaver et al. 2022). The agreement here is excellent for all of the broadbands. Despite the $r < 25$ selection being relatively shallow compared to the *Ch.1 < 26* cut, we still see substantial bimodality in the $YJHK_S$ bands from UltraVISTA, arising from the mix of deep and ultra-deep stripes (H. J. McCracken et al. 2012).

G. FULL SPS PARAMETER DISTRIBUTION

In Figure 34, we show 1D and 2D marginals of the SPS parameter distribution implied by the `pop-cosmos` model. In place of the SFR ratios between the bins of our non-parametric SFH, we show SFR averaged over the last 100 Myr, and mass weighted age.

H. POSSIBLE AUGMENTATIONS TO THE SPS MODEL

Our existing SPS configuration has proven to be very successful, both within `pop-cosmos`, and as the basis of the widely used `Prospector` model family (J. Leja et al. 2017, 2018, 2019b,a; B. D. Johnson et al. 2021a; B. Wang et al. 2023b). Nevertheless, these models are under active development, and we will continue to revisit the modeling choices, particularly in future extensions of the model to high-z data or spectroscopy. In the following subsections, we highlight some aspects of the SPS model that will be particularly interesting to revisit.

H.1. *Stellar Templates*

Thus far, we have made use of the MILES empirical stellar template library (P. Sánchez-Blázquez et al. 2006; J. Falcón-Barroso et al. 2011), combined with the MIST evolutionary tracks (A. Dotter 2016; J. Choi et al. 2016), and a G. Chabrier (2003) IMF as the basis of our SPS model. Whilst the MIST models — based on the MESA stellar evolution code (B. Paxton et al. 2011, 2013, 2015) — include stellar rotation, they do not incorporate the effect of binary stars which can have significantly different evolutionary pathways (see, e.g., the reviews by J. J. Eldridge & E. R. Stanway 2019,

2022) and which are expected to be ubiquitous (H. Sana et al. 2012). The incorporation of binary stars can boost the strength of the ionizing continuum, leading to bluer galaxy SEDs and stronger nebular emission (see, e.g., J. J. Eldridge & E. R. Stanway 2009, 2012; L. Xiao et al. 2018; Y. Götberg et al. 2017, 2019); this is necessary to explain spectroscopic observations at high redshifts (e.g., E. R. Stanway et al. 2014; C. C. Steidel et al. 2014, 2016; R. L. Larson et al. 2023). Future iterations of `pop-cosmos` could explore the impact of this using the Binary Population & Spectral Synthesis models (J. J. Eldridge & E. R. Stanway 2009, 2012; J. J. Eldridge et al. 2017; E. R. Stanway & J. J. Eldridge 2018; C. M. Byrne et al. 2022), which are included as template simple stellar populations (SSPs) in FSPS with `Cloudy`-based nebular emission (K. Garofali et al. 2024).

Another potential improvement would be the incorporation of α-enhancement (super-solar $[\alpha/\text{Fe}]$, where $\alpha = \{\text{C}, \text{O}, \text{Ne}, \text{Mg}, \text{S}, \text{Ar}\}$) into the underlying stellar libraries. This is observationally supported at $z \gtrsim 2$ (e.g., C. C. Steidel et al. 2016; A. L. Strom et al. 2017, 2022; F. Cullen et al. 2019; T. M. Stanton et al. 2024), and can have effects on line strengths and ratios, as well as broadband colours (see, e.g., M. Park et al. 2024). Self-consistent SSPs with α-enhancement are now available for the MIST isochrones (M. Park et al. 2024) and single stars from BPASS (C. M. Byrne et al. 2022, 2025).

H.2. *Nebular Emission*

As already noted in Section H.1, binary stellar evolution and the presence of α-enhancement can modify the nebular emission properties of galaxies. Contributions from AGN (see Section H.3) can also introduce additional emission lines with variable widths and strengths that are not captured well by a single universal template (see, e.g., A. Marshall et al. 2022). Beyond improving the underlying stellar and AGN templates, we can incorporate more flexible emission line corrections into our forward-modeling to handle these effects. Currently, we use the approach introduced by L23, whereby the emission line strengths are based on FSPS (N. Byler et al. 2017) and treated as delta functions pushed through approximations of the bandpasses. A more sophisticated model would allow for variation of the line widths as well as their amplitudes. Rather than inferring empirical corrections to a base set of precomputed nebular emission models, direct on-the-fly emulation of the photoionization calculations is a promising alternative route to more flexible emission line treatment (e.g., Y. Li et al. 2024, 2025; R. P. Naidu et al. 2025; C. Morisset et al. 2025). A more detailed line model would be best calibrated directly with spectroscopic data, as has been explored recently by A. Khederlarian et al. (2024) using DESI Bright Galaxy Survey data (C. Hahn et al. 2023b).
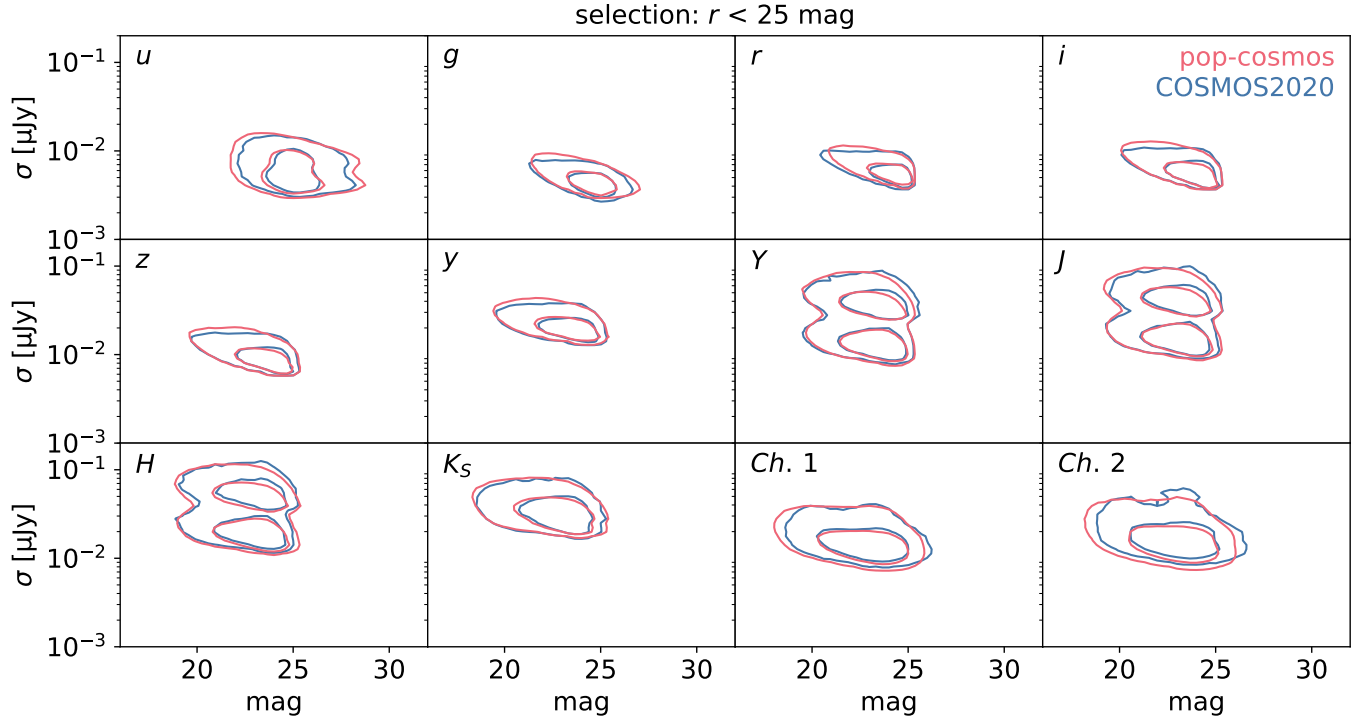
Figure 33. Flux uncertainty (micro Janskys, logarithmic axes) vs. magnitude (logarithmic) for the COSMOS broad bands, compared to the flux errors reported in COSMOS2020 for $r < 25$. Contours enclose the 68 and 95% highest density regions.

H.3. *AGN Modeling*

Our existing AGN treatment follows J. Leja et al. (2018) in using the M. Nenkova et al. (2008a,b) templates for emission from a dusty torus. As discussed in J. Leja et al. (2018), the re-radiated emission from the AGN torus serves primarily to enhance the total MIR emission from a galaxy. In the absence of an AGN, MIR luminosity typically ascribed to the re-radiation of starlight absorbed at UV and optical wavelengths, with a tight correlation between UV attenuation and MIR emission following from a balancing of the energy budget (e.g., M. Rowan-Robinson 1992; V. Buat & C. Xu 1996; A. Efstathiou et al. 2000; A. N. Witt & K. D. Gordon 2000; V. Buat et al. 2002, 2005; D. Burgarella et al. 2005; E. da Cunha et al. 2008). A dust-shrouded AGN can be invoked to explain MIR excess, with inhomogeneity or clumpiness in the torus (as in the M. Nenkova et al. 2008a,b models) being required to explain the diversity of emission scenarios (see also M. Rowan-Robinson 1995; M. Nenkova et al. 2002; R. Mor et al. 2009; S. F. Hönig et al. 2010; S. F. Hönig & M. Kishimoto 2010, 2017; R. Siebenmorgen et al. 2015; M. Stalevski et al. 2016; A. Efstathiou et al. 2013, 2022).

Significant emission in the MIR from the dust around AGN is highly prevalent ($\sim$ 20–40% of galaxies; see, e.g., A. Kirkpatrick et al. 2012, 2015; S. Juneau et al. 2013), and it has been shown to be a crucial component in accurate modeling of star formation histories (L. Ciesla et al. 2015; S. Salim et al. 2016; J. Leja et al. 2018). How-

ever, depending on the orientation of the AGN and the amount of dust surrounding it (see, e.g., R. Antonucci 1993; C. M. Urry & P. Padovani 1995), we may see significant flux from the accretion disk directly, which radiates primarily at UV and optical wavelengths (see, e.g., C. M. Harrison 2016; P. Padovani et al. 2017) with a UV–optical continuum that follows a power law (e.g., A. T. Koski 1978). The impact of this on a galaxy SED can be non-negligible (e.g., L. S. M. Cardoso et al. 2017), and a variety of templates exist incorporating radiation from the disk, emission by hot dust, and broad and narrow emission lines (e.g., N. Maddox & P. C. Hewett 2006; G. T. Richards et al. 2006; A. Kubota & C. Done 2018; M. J. Temple et al. 2021b).

Although these templates have been incorporated alongside SPS-based galaxy models in SED fitting frameworks (e.g., G. Calistro Rivera et al. 2016; L. S. M. Cardoso et al. 2017; M. Boquien et al. 2019; A. Marshall et al. 2022; L. N. Martínez-Ramírez et al. 2024; B. Wang et al. 2025), we have not yet included them within our forward model. Incorporating a more flexible and complete AGN model will be particularly important in future work to model spectroscopic data, where there will be increased sensitivity to the contributions of AGN-driven emission lines (which are highly diverse and dependent on black hole physics; e.g., G. T. Richards et al. 2011, 2021; J. D. Timlin et al. 2020; A. L. Rankine et al. 2020; A. Marshall et al. 2022; M. J. Temple et al. 2021a,b, 2023), as well as the suppression of stellar absorption lines due to the AGN UV continuum (e.g.,
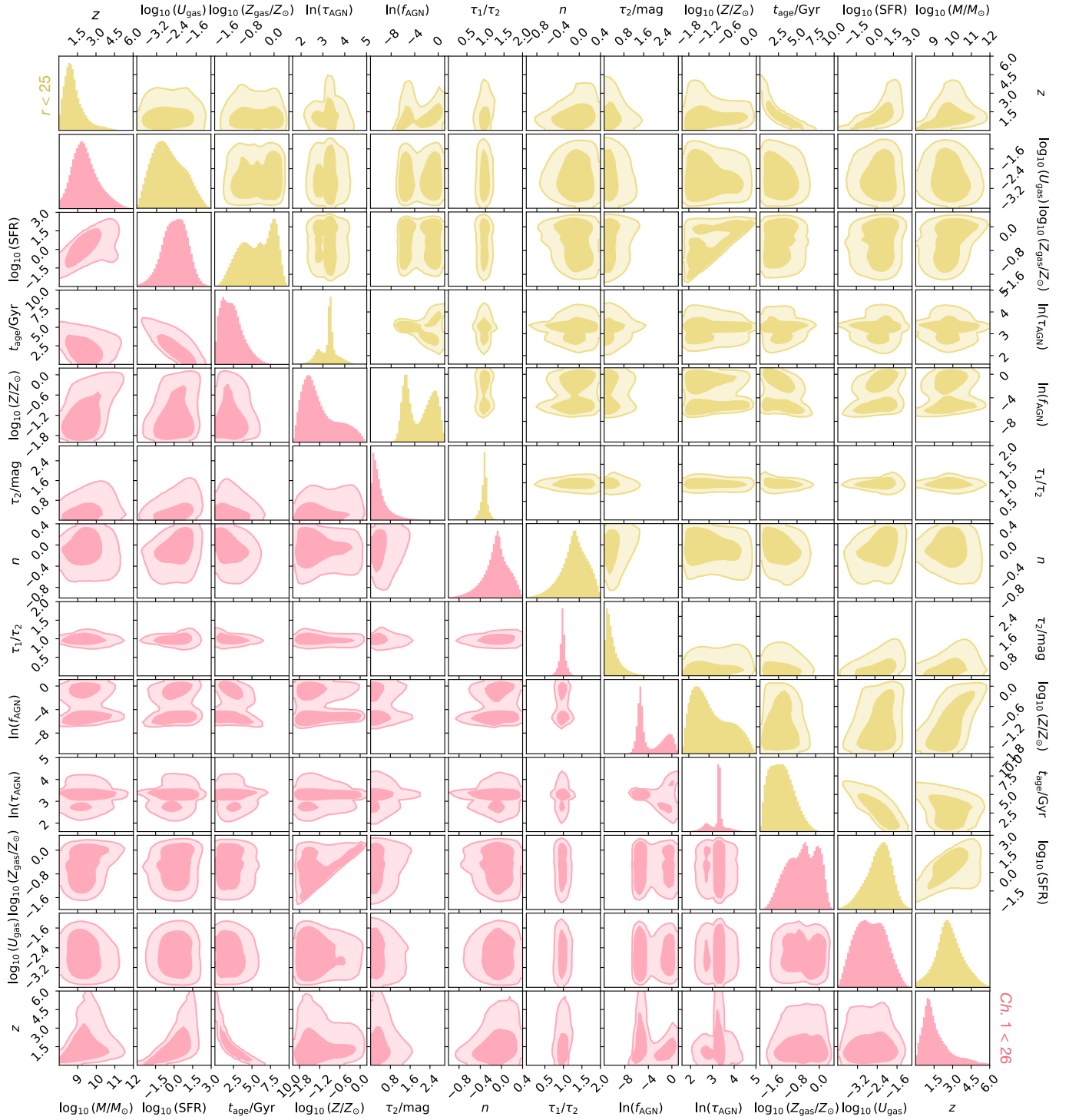
Figure 34. Marginals of the `pop-cosmos` population distribution over SPS parameters for **(top triangle, yellow)** $r < 25$, and **(bottom triangle, red)** *Ch. 1* < 26. Contours enclose the 68 and 95% highest probability density regions.

A. T. Koski 1978; R. Cid Fernandes et al. 1998; L. R. Vega et al. 2009; L. S. M. Cardoso et al. 2017).

L. Tortorelli et al. (2024) recently explored the impact on the broadband colors of galaxies drawn from the Prospector-β prior (B. Wang et al. 2023b), when adopting an AGN spectral template based on M. J. Temple et al. (2021b) in place of the M. Nenkova et al. (2008a,b) torus model that is standard in FSPS. They found that the choice of AGN model had a moderate impact on observer-frame galaxy colors, and some difference of tomographic redshift bin assignments based on a self-organising map (D. Masters et al. 2015; J. McCullough et al. 2024). They also showed that the optical colors $(g-r, r-i, i-z)$ of SDSS quasars (B. W. Lyke et al. 2020) are more closely matched by galaxies including the M. J. Temple et al. (2021b) AGN models than those including the M. Nenkova et al. (2008a,b). This is somewhat unsurprising given that (a) the M. J. Temple et al. (2021b) models are calibrated against the B. W. Lyke et al. (2020) catalog, and (b) the M. Nenkova et al. (2008a,b) models only aim to model the MIR emission due to dust. Nevertheless, incorporating an AGN disk model alongside a torus model (as was done recently by B. Wang et al. 2025 in Prospector-β) would be a worthwhile extension of the pop-cosmos framework.

H.4. *IGM Absorption*

We have so far used the P. Madau (1995) model for the attenuation of SEDs by the IGM (J. E. Gunn & B. A. Peterson 1965). As the population of Ly-α absorbers in the Universe has become better understood, refined models for their impact have been developed (see, e.g., M. A. Bershady et al. 1999; A. Meiksin 2006; T. Tepper-García & U. Fritze 2008; A. K. Inoue & I. Iwata 2008; G. D. Becker et al. 2013; A. K. Inoue et al. 2014; K. Kakiichi & M. Dijkstra 2018). The A. K. Inoue et al. (2014) update to the P. Madau (1995) model predicts an IGM transmission function that agrees better with modern observations; particularly at $z \gtrsim 5$ where the Ly-α optical depth increases very rapidly close to the epoch of reionization (e.g., R. H. Becker et al. 2001; R. Cen & P. McDonald 2002; X. Fan et al. 2002, 2006; R. L. White et al. 2003; P. Paschos & M. L. Norman 2005). Although galaxies with $z \gtrsim 4$ make up a fairly small fraction of our current training set, a post-P. Madau (1995) IGM absorption model (e.g., A. K. Inoue et al. 2014) may be needed in future work to extrapolate pop-cosmos to higher redshifts. Moreover, there is considerable variance observed in the amount of IGM absorption along different sightlines towards high-z galaxies (see, e.g., R. Thomas et al. 2017, 2020, 2021), which can in principle be accounted for in FSPS/Prospector (see B. D. Johnson et al. 2021a). Other factors such as the damping wing that attenuates light redward of the Ly-α line center (see, e.g., J. Miralda-Escudé 1998; P. Madau & M. J. Rees 2000; T. Totani et al. 2006; J. Schroeder et al. 2013; K. Bach & H.-W. Lee 2015; D. Mortlock 2016; D. Ďurovčíková et al. 2024) may merit inclusion when extending pop-cosmos to $z \gtrsim 6$.

REFERENCES

Abazajian, K., Adelman-McCarthy, J. K., Agüeros, M. A., et al. 2003, AJ, 126, 2081, doi: 10.1086/378165

—. 2004, AJ, 128, 502, doi: 10.1086/421365

Abramson, L. E., Gladders, M. D., Dressler, A., et al. 2015, ApJL, 801, L12, doi: 10.1088/2041-8205/801/1/L12

Adams, N. J., Bowler, R. A. A., Jarvis, M. J., Häußler, B., & Lagos, C. D. P. 2021, MNRAS, 506, 4933, doi: 10.1093/mnras/stab1956

Aihara, H., AlSayyad, Y., Ando, M., et al. 2019, PASJ, 71, 114, doi: 10.1093/pasj/psz103

—. 2022, PASJ, 74, 247, doi: 10.1093/pasj/psab122

Aird, J., Nandra, K., Laird, E. S., et al. 2010, MNRAS, 401, 2531, doi: 10.1111/j.1365-2966.2009.15829.x

Aird, J., Coil, A. L., Moustakas, J., et al. 2012, ApJ, 746, 90, doi: 10.1088/0004-637X/746/1/90

Alarcon, A., Hearin, A. P., Becker, M. R., & Chaves-Montero, J. 2023, MNRAS, 518, 562, doi: 10.1093/mnras/stac3118

Alexander, D. M., Hickox, R. C., Aird, J., et al. 2025, NewAR, 101, 101733, doi: 10.1016/j.newar.2025.101733

Alsing, J., Peiris, H., Mortlock, D., Leja, J., & Leistedt, B. 2023, ApJS, 264, 29, doi: 10.3847/1538-4365/ac9583

Alsing, J., Thorp, S., Deger, S., et al. 2024, ApJS, 274, 12, doi: 10.3847/1538-4365/ad5c69

Alsing, J., Peiris, H., Leja, J., et al. 2020, ApJS, 249, 5, doi: 10.3847/1538-4365/ab917f

Antonucci, R. 1993, ARA&A, 31, 473, doi: 10.1146/annurev.aa.31.090193.002353

Arnouts, S., Cristiani, S., Moscardini, L., et al. 1999, MNRAS, 310, 540, doi: 10.1046/j.1365-8711.1999.02978.x

Ashby, M. L. N., Willner, S. P., Fazio, G. G., et al. 2013, ApJ, 769, 80, doi: 10.1088/0004-637X/769/1/80

—. 2015, ApJS, 218, 33, doi: 10.1088/0067-0049/218/2/33

Ashby, M. L. N., Caputi, K. I., Cowley, W., et al. 2018, ApJS, 237, 39, doi: 10.3847/1538-4365/aad4fb

Assef, R. J., Stern, D., Noirot, G., et al. 2018, ApJS, 234, 23, doi: 10.3847/1538-4365/aaa00a

Assef, R. J., Kochanek, C. S., Brodwin, M., et al. 2010, ApJ, 713, 970, doi: 10.1088/0004-637X/713/2/970

Assef, R. J., Stern, D., Kochanek, C. S., et al. 2013, ApJ, 772, 26, doi: 10.1088/0004-637X/772/1/26

Assef, R. J., Eisenhardt, P. R. M., Stern, D., et al. 2015, ApJ, 804, 27, doi: 10.1088/0004-637X/804/1/27

Assef, R. J., Walton, D. J., Brightman, M., et al. 2016, ApJ, 819, 111, doi: 10.3847/0004-637X/819/2/111

Assef, R. J., Brightman, M., Walton, D. J., et al. 2020, ApJ, 897, 112, doi: 10.3847/1538-4357/ab9814

Assef, R. J., Bauer, F. E., Blain, A. W., et al. 2022, ApJ, 934, 101, doi: 10.3847/1538-4357/ac77fc

Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33, doi: 10.1051/0004-6361/201322068

Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, AJ, 156, 123, doi: 10.3847/1538-3881/aabc4f

Astropy Collaboration, Price-Whelan, A. M., Lim, P. L., et al. 2022, ApJ, 935, 167, doi: 10.3847/1538-4357/ac7c74

Autenrieth, M., Wright, A. H., Trotta, R., et al. 2024, MNRAS, 534, 3808, doi: 10.1093/mnras/stae2243

Azadi, M., Aird, J., Coil, A. L., et al. 2015, ApJ, 806, 187, doi: 10.1088/0004-637X/806/2/187

Azadi, M., Coil, A. L., Aird, J., et al. 2017, ApJ, 835, 27, doi: 10.3847/1538-4357/835/1/27

Bach, K., & Lee, H.-W. 2015, MNRAS, 446, 264, doi: 10.1093/mnras/stu2030

Baggen, J. F. W., van Dokkum, P., Brammer, G., et al. 2024, ApJL, 977, L13, doi: 10.3847/2041-8213/ad90b8

Barro, G., Pérez-González, P. G., Kocevski, D. D., et al. 2024, ApJ, 963, 128, doi: 10.3847/1538-4357/ad167e

Becker, G. D., Hewett, P. C., Worseck, G., & Prochaska, J. X. 2013, MNRAS, 430, 2067, doi: 10.1093/mnras/stt031

Becker, R. H., Fan, X., White, R. L., et al. 2001, AJ, 122, 2850, doi: 10.1086/324231

Behroozi, P., Wechsler, R. H., Hearin, A. P., & Conroy, C. 2019, MNRAS, 488, 3143, doi: 10.1093/mnras/stz1182

Bennett, J. S., Smith, M. C., Fielding, D. B., et al. 2024, arXiv e-prints, arXiv:2410.12909, doi: 10.48550/arXiv.2410.12909

Bergé, J., Gamper, L., Réfrégier, A., & Amara, A. 2013, Astron. Comput., 1, 23, doi: 10.1016/j.ascom.2013.01.001

Bershady, M. A., Charlton, J. C., & Geoffroy, J. M. 1999, ApJ, 518, 103, doi: 10.1086/307257

Bertin, E., & Arnouts, S. 1996, A&AS, 117, 393, doi: 10.1051/aas:1996164

Biewald, L. 2020, Experiment Tracking with Weights and Biases. https://www.wandb.com/

Bishop, C. M. 2006, Pattern Recognition and Machine Learning (Springer)

Boquien, M., Burgarella, D., Roehlly, Y., et al. 2019, A&A, 622, A103, doi: 10.1051/0004-6361/201834156

Bower, R. G., Schaye, J., Frenk, C. S., et al. 2017, MNRAS, 465, 32, doi: 10.1093/mnras/stw2735

Brammer, G. B., van Dokkum, P. G., & Coppi, P. 2008, ApJ, 686, 1503, doi: 10.1086/591786

Brand, A., Allen, L., Altman, M., Hlava, M., & Scott, J. 2015, Learned Publishing, 28, 151, doi: 10.1087/20150211

Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000, doi: 10.1046/j.1365-8711.2003.06897.x

Buat, V., Boselli, A., Gavazzi, G., & Bonfanti, C. 2002, A&A, 383, 801, doi: 10.1051/0004-6361:20011832

Buat, V., & Xu, C. 1996, A&A, 306, 61

Buat, V., Iglesias-Páramo, J., Seibert, M., et al. 2005, ApJL, 619, L51, doi: 10.1086/423241

Burgarella, D., Buat, V., & Iglesias-Páramo, J. 2005, MNRAS, 360, 1413, doi: 10.1111/j.1365-2966.2005.09131.x

Byler, N., Dalcanton, J. J., Conroy, C., & Johnson, B. D. 2017, ApJ, 840, 44, doi: 10.3847/1538-4357/aa6c66

Byrne, C. M., Eldridge, J. J., & Stanway, E. R. 2025, MNRAS, 537, 2433, doi: 10.1093/mnras/staf178

Byrne, C. M., Stanway, E. R., Eldridge, J. J., McSwiney, L., & Townsend, O. T. 2022, MNRAS, 512, 5329, doi: 10.1093/mnras/stac807

Calabrò, A., Castellano, M., Pentericci, L., et al. 2021, A&A, 646, A39, doi: 10.1051/0004-6361/202039244

Calistro Rivera, G., Lusso, E., Hennawi, J. F., & Hogg, D. W. 2016, ApJ, 833, 98, doi: 10.3847/1538-4357/833/1/98

Calzetti, D. 2001, PASP, 113, 1449, doi: 10.1086/324269

Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, ApJ, 533, 682, doi: 10.1086/308692

Caputi, K. I., Ilbert, O., Laigle, C., et al. 2015, ApJ, 810, 73, doi: 10.1088/0004-637X/810/1/73

Cardoso, L. S. M., Gomes, J. M., & Papaderos, P. 2017, A&A, 604, A99, doi: 10.1051/0004-6361/201630378

Carnall, A. C., Leja, J., Johnson, B. D., et al. 2019, ApJ, 873, 44, doi: 10.3847/1538-4357/ab04a2

Carroll, C. M., Hickox, R. C., Masini, A., et al. 2021, ApJ, 908, 185, doi: 10.3847/1538-4357/abd185

Cattell, R. B. 1966, Multivariate Behavioral Res., 1, 245, doi: 10.1207/s15327906mbr0102_10

Cen, R., & McDonald, P. 2002, ApJ, 570, 457, doi: 10.1086/339723

Chabrier, G. 2003, PASP, 115, 763, doi: 10.1086/376392

Chaikin, E., Schaye, J., Schaller, M., et al. 2025, arXiv e-prints, arXiv:2509.04067, doi: 10.48550/arXiv.2509.04067

Chang, Y.-Y., Le Floc'h, E., Juneau, S., et al. 2017, ApJS, 233, 19, doi: 10.3847/1538-4365/aa97da

Charlot, S., & Fall, S. M. 2000, ApJ, 539, 718,
doi: 10.1086/309250

Chartab, N., Newman, A. B., Rudie, G. C., Blanc, G. A., &
Kelson, D. D. 2024, ApJ, 960, 73,
doi: 10.3847/1538-4357/ad0554

Chartab, N., Mobasher, B., Cooray, A. R., et al. 2023, ApJ,
942, 91, doi: 10.3847/1538-4357/acacf5

Chen, R. T. Q., Rubanova, Y., Bettencourt, J., &
Duvenaud, D. K. 2018, in Advances in Neural
Information Processing Systems, ed. S. Bengio,
H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi,
& R. Garnett, Vol. 31 (Curran Associates, Inc.),
6572–6583. https://arxiv.org/abs/1806.07366

Chérouvrier, D., Macías-Pérez, J. F., Désert, F. X., et al.
2025, A&A, 700, A30, doi: 10.1051/0004-6361/202554450

Chevallard, J., Charlot, S., Wandelt, B., & Wild, V. 2013,
MNRAS, 432, 2061, doi: 10.1093/mnras/stt523

Chiang, Y.-K., Overzier, R. A., Gebhardt, K., et al. 2015,
ApJ, 808, 37, doi: 10.1088/0004-637X/808/1/37

Childress, M. J., Wolf, C., & Zahid, H. J. 2014, MNRAS,
445, 1898, doi: 10.1093/mnras/stu1892

Choi, J., Dotter, A., Conroy, C., et al. 2016, ApJ, 823, 102,
doi: 10.3847/0004-637X/823/2/102

Cid Fernandes, Jr., R., Storchi-Bergmann, T., & Schmitt,
H. R. 1998, MNRAS, 297, 579,
doi: 10.1046/j.1365-8711.1998.01519.x

Ciesla, L., Charmandaris, V., Georgakakis, A., et al. 2015,
A&A, 576, A10, doi: 10.1051/0004-6361/201425252

Cirasuolo, M., Afonso, J., Bender, R., et al. 2011, The
Messenger, 145, 11

Cirasuolo, M., Fairley, A., Rees, P., et al. 2020, The
Messenger, 180, 10, doi: 10.18727/0722-6691/5195

Civano, F., Marchesi, S., Comastri, A., et al. 2016, ApJ,
819, 62, doi: 10.3847/0004-637X/819/1/62

Coil, A. L., Blanton, M. R., Burles, S. M., et al. 2011, ApJ,
741, 8, doi: 10.1088/0004-637X/741/1/8

Conroy, C. 2013, ARA&A, 51, 393,
doi: 10.1146/annurev-astro-082812-141017

Conroy, C., & Gunn, J. E. 2010, ApJ, 712, 833,
doi: 10.1088/0004-637X/712/2/833

Conroy, C., Gunn, J. E., & White, M. 2009, ApJ, 699, 486,
doi: 10.1088/0004-637X/699/1/486

Conroy, C., White, M., & Gunn, J. E. 2010, ApJ, 708, 58,
doi: 10.1088/0004-637X/708/1/58

Crenshaw, J. F., Kalmbach, J. B., Gagliano, A., et al. 2024,
AJ, 168, 80, doi: 10.3847/1538-3881/ad54bf

Cresci, G., Mannucci, F., & Curti, M. 2019, A&A, 627,
A42, doi: 10.1051/0004-6361/201834637

Cucciati, O., Lemaux, B. C., Zamorani, G., et al. 2018,
A&A, 619, A49, doi: 10.1051/0004-6361/201833655

Cullen, F., McLure, R. J., Dunlop, J. S., et al. 2019,
MNRAS, 487, 2038, doi: 10.1093/mnras/stz1402

Cunha, P. A. C., Humphrey, A., Brinchmann, J., et al.
2025, A&A, 696, A110,
doi: 10.1051/0004-6361/202453068

Curti, M., Mannucci, F., Cresci, G., & Maiolino, R. 2020,
MNRAS, 491, 944, doi: 10.1093/mnras/stz2910

Curti, M., Maiolino, R., Curtis-Lake, E., et al. 2024, A&A,
684, A75, doi: 10.1051/0004-6361/202346698

da Costa-Luis, C., Larroque, S. K., Altendorf, K., et al.
2024, tqdm: A fast, Extensible Progress Bar for Python
and CLI, v4.66.4, Zenodo, doi: 10.5281/zenodo.595120

da Cunha, E., Charlot, S., & Elbaz, D. 2008, MNRAS, 388,
1595, doi: 10.1111/j.1365-2966.2008.13535.x

Daddi, E., Dickinson, M., Morrison, G., et al. 2007, ApJ,
670, 156, doi: 10.1086/521818

Dai, Y. S., Wilkes, B. J., Bergeron, J., et al. 2018, MNRAS,
478, 4238, doi: 10.1093/mnras/sty1341

Dale, D. A., & Helou, G. 2002, ApJ, 576, 159,
doi: 10.1086/341632

Davé, R., Anglés-Alcázar, D., Narayanan, D., et al. 2019,
MNRAS, 486, 2827, doi: 10.1093/mnras/stz937

Davidzon, I., Ilbert, O., Laigle, C., et al. 2017, A&A, 605,
A70, doi: 10.1051/0004-6361/201730419

Delvecchio, I., Gruppioni, C., Pozzi, F., et al. 2014,
MNRAS, 439, 2736, doi: 10.1093/mnras/stu130

DESI Collaboration, Adame, A. G., Aguilar, J., et al. 2024,
AJ, 168, 58, doi: 10.3847/1538-3881/ad3217

Diener, C., Lilly, S. J., Ledoux, C., et al. 2015, ApJ, 802,
31, doi: 10.1088/0004-637X/802/1/31

Domínguez Sánchez, H., Pozzi, F., Gruppioni, C., et al.
2011, MNRAS, 417, 900,
doi: 10.1111/j.1365-2966.2011.19263.x

Donley, J. L., Rieke, G. H., Alexander, D. M., Egami, E., &
Pérez-González, P. G. 2010, ApJ, 719, 1393,
doi: 10.1088/0004-637X/719/2/1393

Donley, J. L., Rieke, G. H., Pérez-González, P. G., Rigby,
J. R., & Alonso-Herrero, A. 2007, ApJ, 660, 167,
doi: 10.1086/512798

Donley, J. L., Koekemoer, A. M., Brusa, M., et al. 2012,
ApJ, 748, 142, doi: 10.1088/0004-637X/748/2/142

Dotter, A. 2016, ApJS, 222, 8,
doi: 10.3847/0067-0049/222/1/8

Draine, B. T., & Li, A. 2007, ApJ, 657, 810,
doi: 10.1086/511055

Driver, S. P., Bellstedt, S., Robotham, A. S. G., et al. 2022,
MNRAS, 513, 439, doi: 10.1093/mnras/stac472

Drlica-Wagner, A., Sevilla-Noarbe, I., Rykoff, E. S., et al.
2018, ApJS, 235, 33, doi: 10.3847/1538-4365/aab4f5

Dubois, Y., Volonteri, M., Silk, J., et al. 2015, MNRAS, 452, 1502, doi: 10.1093/mnras/stv1416

Dunlop, J. S. 2016, UltraVISTA – Ultra Deep Survey with VISTA, DR 4.1.1, European Southern Observatory (ESO), doi: 10.18727/archive/52

Ďurovčíková, D., Eilers, A.-C., Chen, H., et al. 2024, ApJ, 969, 162, doi: 10.3847/1538-4357/ad4888

Eckart, M. E., McGreer, I. D., Stern, D., Harrison, F. A., & Helfand, D. J. 2010, ApJ, 708, 584, doi: 10.1088/0004-637X/708/1/584

Eddington, A. S. 1913, MNRAS, 73, 359, doi: 10.1093/mnras/73.5.359

—. 1940, MNRAS, 100, 354, doi: 10.1093/mnras/100.5.354

Efstathiou, A., Christopher, N., Verma, A., & Siebenmorgen, R. 2013, MNRAS, 436, 1873, doi: 10.1093/mnras/stt1695

Efstathiou, A., Rowan-Robinson, M., & Siebenmorgen, R. 2000, MNRAS, 313, 734, doi: 10.1046/j.1365-8711.2000.03269.x

Efstathiou, A., Farrah, D., Afonso, J., et al. 2022, MNRAS, 512, 5183, doi: 10.1093/mnras/stab3642

Eldridge, J. J., & Stanway, E. R. 2009, MNRAS, 400, 1019, doi: 10.1111/j.1365-2966.2009.15514.x

—. 2012, MNRAS, 419, 479, doi: 10.1111/j.1365-2966.2011.19713.x

—. 2019, in Cambridge Astrophysics, Vol. 54, The Impact of Binary Stars on Stellar Evolution, ed. G. Beccari & H. M. J. Boffin (Cambridge University Press), 307–328, doi: 10.1017/9781108553070.023

—. 2022, ARA&A, 60, 455, doi: 10.1146/annurev-astro-052920-100646

Eldridge, J. J., Stanway, E. R., Xiao, L., et al. 2017, PASA, 34, e058, doi: 10.1017/pasa.2017.51

Ellison, S. L., Teimoorinia, H., Rosario, D. J., & Mendel, J. T. 2016, MNRAS, 458, L34, doi: 10.1093/mnrasl/slw012

Ellison, S. L., Viswanathan, A., Patton, D. R., et al. 2019, MNRAS, 487, 2491, doi: 10.1093/mnras/stz1431

Euclid Collaboration, Moneti, A., McCracken, H. J., et al. 2022, A&A, 658, A126, doi: 10.1051/0004-6361/202142361

Euclid Collaboration, Mellier, Y., Abdurro'uf, et al. 2025a, A&A, 697, A1, doi: 10.1051/0004-6361/202450810

Euclid Collaboration, Zalesky, L., Weaver, J. R., et al. 2025b, arXiv e-prints, arXiv:2504.17867, doi: 10.48550/arXiv.2504.17867

Falcón-Barroso, J., Sánchez-Blázquez, P., Vazdekis, A., et al. 2011, A&A, 532, A95, doi: 10.1051/0004-6361/201116842

Fan, X., Narayanan, V. K., Strauss, M. A., et al. 2002, AJ, 123, 1247, doi: 10.1086/339030

Fan, X., Strauss, M. A., Becker, R. H., et al. 2006, AJ, 132, 117, doi: 10.1086/504836

Ferland, G. J., Porter, R. L., van Hoof, P. A. M., et al. 2013, RMxAA, 49, 137, doi: 10.48550/arXiv.1302.4485

Fischbacher, S., Kacprzak, T., Machado Poletti Valle, L. F., & Refregier, A. 2025a, MNRAS, 542, L53, doi: 10.1093/mnrasl/slaf062

Fischbacher, S., Kacprzak, T., Tortorelli, L., et al. 2025b, JCAP, 2025, 007, doi: 10.1088/1475-7516/2025/06/007

Fischbacher, S., Moser, B., Kacprzak, T., et al. 2024a, arXiv e-prints, arXiv:2412.08722, doi: 10.48550/arXiv.2412.08722

—. 2024b, arXiv e-prints, arXiv:2412.08716, doi: 10.48550/arXiv.2412.08716

Florez, J., Jogee, S., Sherman, S., et al. 2020, MNRAS, 497, 3273, doi: 10.1093/mnras/staa2200

Florez, J., Jogee, S., Guo, Y., et al. 2021, MNRAS, 508, 762, doi: 10.1093/mnras/stab2593

Foreman-Mackey, D. 2016, J. Open Source Software, 1, 24, doi: 10.21105/joss.00024

Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306, doi: 10.1086/670067

Forrest, B., Lemaux, B. C., Shah, E., et al. 2023, MNRAS, 526, L56, doi: 10.1093/mnrasl/slad114

Fu, H., Shankar, F., Ayromlou, M., et al. 2024, MNRAS, 532, 177, doi: 10.1093/mnras/stae1492

Fukugita, M., Ichikawa, T., Gunn, J. E., et al. 1996, AJ, 111, 1748, doi: 10.1086/117915

Gallazzi, A., Bell, E. F., Zibetti, S., Brinchmann, J., & Kelson, D. D. 2014, ApJ, 788, 72, doi: 10.1088/0004-637X/788/1/72

Gallazzi, A., Charlot, S., Brinchmann, J., White, S. D. M., & Tremonti, C. A. 2005, MNRAS, 362, 41, doi: 10.1111/j.1365-2966.2005.09321.x

Gardner, J. P., Mather, J. C., Abbott, R., et al. 2023, PASP, 135, 068001, doi: 10.1088/1538-3873/acd1b5

Garn, T., & Best, P. N. 2010, MNRAS, 409, 421, doi: 10.1111/j.1365-2966.2010.17321.x

Garofali, K., Basu-Zych, A. R., Johnson, B. D., et al. 2024, ApJ, 960, 13, doi: 10.3847/1538-4357/ad0a6a

Gilda, S., Lower, S., & Narayanan, D. 2021, ApJ, 916, 43, doi: 10.3847/1538-4357/ac0058

Goodman, J., & Weare, J. 2010, Communications in Applied Math. & Comput. Sci., 5, 65, doi: 10.2140/camcos.2010.5.65

Götberg, Y., de Mink, S. E., & Groh, J. H. 2017, A&A, 608, A11, doi: 10.1051/0004-6361/201730472

Götberg, Y., de Mink, S. E., Groh, J. H., Leitherer, C., & Norman, C. 2019, A&A, 629, A134, doi: 10.1051/0004-6361/201834525

Goulding, A. D., Greene, J. E., Bezanson, R., et al. 2018, PASJ, 70, S37, doi: 10.1093/pasj/psx135

Grathwohl, W., Chen, R. T. Q., Bettencourt, J., Sutskever, I., & Duvenaud, D. 2018, arXiv e-prints, arXiv:1810.01367, doi: 10.48550/arXiv.1810.01367

Grazian, A., Fontana, A., Santini, P., et al. 2015, A&A, 575, A96, doi: 10.1051/0004-6361/201424750

Green, D. A. 2011, Bull. Astron. Soc. India, 39, 289, doi: 10.48550/arXiv.1108.5083

Grogin, N. A., Kocevski, D. D., Faber, S. M., et al. 2011, ApJS, 197, 35, doi: 10.1088/0067-0049/197/2/35

Gunn, J. E., & Peterson, B. A. 1965, ApJ, 142, 1633, doi: 10.1086/148444

Hahn, C., & Melchior, P. 2022, ApJ, 938, 11, doi: 10.3847/1538-4357/ac7b84

Hahn, C., Kwon, K. J., Tojeiro, R., et al. 2023a, ApJ, 945, 16, doi: 10.3847/1538-4357/ac8983

Hahn, C., Wilson, M. J., Ruiz-Macias, O., et al. 2023b, AJ, 165, 253, doi: 10.3847/1538-3881/accff8

Hahn, C., Aguilar, J. N., Alam, S., et al. 2024, ApJ, 963, 56, doi: 10.3847/1538-4357/ad19c8

Halliday, C., Daddi, E., Cimatti, A., et al. 2008, A&A, 479, 417, doi: 10.1051/0004-6361:20078673

Han, Y., Fan, L., Zheng, X. Z., Bai, J.-M., & Han, Z. 2023, ApJS, 269, 39, doi: 10.3847/1538-4365/acfc3a

Han, Y., & Han, Z. 2012, ApJ, 749, 123, doi: 10.1088/0004-637X/749/2/123

—. 2014, ApJS, 215, 2, doi: 10.1088/0067-0049/215/1/2

—. 2019, ApJS, 240, 3, doi: 10.3847/1538-4365/aaeffa

Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, Nature, 585, 357, doi: 10.1038/s41586-020-2649-2

Harrison, C. M. 2016, Observational Constraints on the Influence of Active Galactic Nuclei on the Evolution of Galaxies, Springer Theses (Springer Cham), doi: 10.1007/978-3-319-28454-5

—. 2017, Nature Astronomy, 1, 0165, doi: 10.1038/s41550-017-0165

Hasinger, G., Capak, P., Salvato, M., et al. 2018, ApJ, 858, 77, doi: 10.3847/1538-4357/aabacf

Hatcher, C., Kirkpatrick, A., Fornasini, F., et al. 2021, AJ, 162, 65, doi: 10.3847/1538-3881/ac0530

Hearin, A. P., Chaves-Montero, J., Alarcon, A., Becker, M. R., & Benson, A. 2023, MNRAS, 521, 1741, doi: 10.1093/mnras/stad456

Hearin, A. P., Chaves-Montero, J., Becker, M. R., & Alarcon, A. 2021, Open J. Astrophys., 4, 7, doi: 10.21105/astro.2105.05859

Hearin, A. P., Ramachandra, N., Becker, M. R., & DeRose, J. 2022, Open J. Astrophys., 5, 3, doi: 10.21105/astro.2112.08423

Heckman, T. M., & Best, P. N. 2014, ARA&A, 52, 589, doi: 10.1146/annurev-astro-081913-035722

Herard-Demanche, T., Bouwens, R. J., Oesch, P. A., et al. 2025, MNRAS, 537, 788, doi: 10.1093/mnras/staf030

Herbel, J., Kacprzak, T., Amara, A., et al. 2017, JCAP, 2017, 035, doi: 10.1088/1475-7516/2017/08/035

Hickox, R. C., & Alexander, D. M. 2018, ARA&A, 56, 625, doi: 10.1146/annurev-astro-081817-051803

Hickox, R. C., Jones, C., Forman, W. R., et al. 2009, ApJ, 696, 891, doi: 10.1088/0004-637X/696/1/891

Ho, J., Jain, A., & Abbeel, P. 2020, in Advances in Neural Information Processing Systems 33, ed. H. Larochelle, M. Ranzato, R. Hadsell, M. Balcan, & H. Lin. https://arxiv.org/abs/2006.11239

Hönig, S. F., & Kishimoto, M. 2010, A&A, 523, A27, doi: 10.1051/0004-6361/200912676

—. 2017, ApJL, 838, L20, doi: 10.3847/2041-8213/aa6838

Hönig, S. F., Kishimoto, M., Gandhi, P., et al. 2010, A&A, 515, A23, doi: 10.1051/0004-6361/200913742

Hopkins, P. F., Hernquist, L., Cox, T. J., et al. 2006, ApJS, 163, 1, doi: 10.1086/499298

Hopkins, P. F., Richards, G. T., & Hernquist, L. 2007, ApJ, 654, 731, doi: 10.1086/509629

Hunt, L., Dayal, P., Magrini, L., & Ferrara, A. 2016, MNRAS, 463, 2002, doi: 10.1093/mnras/stw1993

Hunter, J. D. 2007, Comput. Sci. Eng., 9, 90, doi: 10.1109/MCSE.2007.55

Hyvärinen, A. 2005, J. Machine Learning Res., 6, 695. http://jmlr.org/papers/v6/hyvarinen05a.html

Ilbert, O., Arnouts, S., McCracken, H. J., et al. 2006, A&A, 457, 841, doi: 10.1051/0004-6361:20065138

Ilbert, O., Capak, P., Salvato, M., et al. 2009, ApJ, 690, 1236, doi: 10.1088/0004-637X/690/2/1236

Ilbert, O., Salvato, M., Le Floc'h, E., et al. 2010, ApJ, 709, 644, doi: 10.1088/0004-637X/709/2/644

Ilbert, O., McCracken, H. J., Le Fèvre, O., et al. 2013, A&A, 556, A55, doi: 10.1051/0004-6361/201321100

Inoue, A. K., & Iwata, I. 2008, MNRAS, 387, 1681, doi: 10.1111/j.1365-2966.2008.13350.x

Inoue, A. K., Shimizu, I., Iwata, I., & Tanaka, M. 2014, MNRAS, 442, 1805, doi: 10.1093/mnras/stu936

Ito, K., Tanaka, M., Valentino, F., et al. 2023, ApJL, 945, L9, doi: 10.3847/2041-8213/acb49b

Ivezić, Ž., Kahn, S. M., Tyson, J. A., et al. 2019, ApJ, 873, 111, doi: 10.3847/1538-4357/ab042c

Iyer, K. G., Pacifici, C., Calistro-Rivera, G., & Lovell, C. C. 2025, arXiv e-prints, arXiv:2502.17680, doi: 10.48550/arXiv.2502.17680

Jespersen, C. K., Melchior, P., Spergel, D. N., et al. 2025a, arXiv e-prints, arXiv:2503.03816, doi: 10.48550/arXiv.2503.03816

Jespersen, C. K., Steinhardt, C. L., Somerville, R. S., & Lovell, C. C. 2025b, ApJ, 982, 23, doi: 10.3847/1538-4357/adb422

Joachimi, B., Cacciato, M., Kitching, T. D., et al. 2015, SSRv, 193, 1, doi: 10.1007/s11214-015-0177-4

Johnson, B. D. 2021, bd-j/sedpy: sedpy, v0.2.0, Zenodo, doi: 10.5281/zenodo.4582723

Johnson, B. D., Leja, J., Conroy, C., & Speagle, J. S. 2021a, ApJS, 254, 22, doi: 10.3847/1538-4365/abef67

Johnson, B. D., Foreman-Mackey, D., Sick, J., et al. 2021b, dfm/python-fsps: python-fsps, v0.4.1rc1, Zenodo, doi: 10.5281/zenodo.4737461

Jones, D. O., McGill, P., Manning, T. A., et al. 2024, arXiv e-prints, arXiv:2410.17322, doi: 10.48550/arXiv.2410.17322

Juneau, S., Dickinson, M., Bournaud, F., et al. 2013, ApJ, 764, 176, doi: 10.1088/0004-637X/764/2/176

Kakiichi, K., & Dijkstra, M. 2018, MNRAS, 480, 5140, doi: 10.1093/mnras/sty2214

Kantorovich, L., & Rubinstein, G. S. 1958, Vestnik Leningrad Univ., 13, 52

Karchev, K., Trotta, R., & Jimenez, R. 2025, arXiv e-prints, arXiv:2508.15899, doi: 10.48550/arXiv.2508.15899

Kartaltepe, J. S., Dickinson, M., Alexander, D. M., et al. 2012, ApJ, 757, 23, doi: 10.1088/0004-637X/757/1/23

Kashino, D., Lilly, S. J., Silverman, J. D., et al. 2021, ApJ, 909, 213, doi: 10.3847/1538-4357/abdf62

Kashino, D., Lilly, S. J., Renzini, A., et al. 2022, ApJ, 925, 82, doi: 10.3847/1538-4357/ac399e

Khederlarian, A., Newman, J. A., Andrews, B. H., et al. 2024, MNRAS, 531, 1454, doi: 10.1093/mnras/stae1189

Khostovan, A. A., Kartaltepe, J. S., Salvato, M., et al. 2025, arXiv e-prints, arXiv:2503.00120, doi: 10.48550/arXiv.2503.00120

Kim, S., Jeong, W.-S., Kim, M., et al. 2025, arXiv e-prints, arXiv:2508.19618, doi: 10.48550/arXiv.2508.19618

Kingma, D. P., & Ba, J. 2015, in 3rd International Conference on Learning Representations, ed. Y. Bengio & Y. LeCun, San Diego, CA, USA. https://arxiv.org/abs/1412.6980

Kirby, E. N., Cohen, J. G., Guhathakurta, P., et al. 2013, ApJ, 779, 102, doi: 10.1088/0004-637X/779/2/102

Kirkpatrick, A., Pope, A., Sajina, A., et al. 2015, ApJ, 814, 9, doi: 10.1088/0004-637X/814/1/9

Kirkpatrick, A., Pope, A., Alexander, D. M., et al. 2012, ApJ, 759, 139, doi: 10.1088/0004-637X/759/2/139

Kocevski, D. D., Finkelstein, S. L., Barro, G., et al. 2025, ApJ, 986, 126, doi: 10.3847/1538-4357/adbc7d

Koekemoer, A. M., Faber, S. M., Ferguson, H. C., et al. 2011, ApJS, 197, 36, doi: 10.1088/0067-0049/197/2/36

Kolmogorov, A. 1933, G. Istituto Ital. Attuari, 4, 83

Koski, A. T. 1978, ApJ, 223, 56, doi: 10.1086/156235

Kriek, M., & Conroy, C. 2013, ApJL, 775, L16, doi: 10.1088/2041-8205/775/1/L16

Kubota, A., & Done, C. 2018, MNRAS, 480, 1247, doi: 10.1093/mnras/sty1890

Kudritzki, R. P., Castro, N., Urbaneja, M. A., et al. 2016, ApJ, 829, 70, doi: 10.3847/0004-637X/829/2/70

Kuijken, K., Heymans, C., Dvornik, A., et al. 2019, A&A, 625, A2, doi: 10.1051/0004-6361/201834918

Kwon, K. J., Hahn, C., & Alsing, J. 2023, ApJS, 265, 23, doi: 10.3847/1538-4365/acba14

Lacy, M., Ridgway, S. E., Gates, E. L., et al. 2013, ApJS, 208, 24, doi: 10.1088/0067-0049/208/2/24

Lagache, G., Xiao, M., Beelen, A., et al. 2025, arXiv e-prints, arXiv:2506.15322, doi: 10.48550/arXiv.2506.15322

Lagos, C. d. P., Tobar, R. J., Robotham, A. S. G., et al. 2018, MNRAS, 481, 3573, doi: 10.1093/mnras/sty2440

Lagos, C. d. P., Robotham, A. S. G., Trayford, J. W., et al. 2019, MNRAS, 489, 4196, doi: 10.1093/mnras/stz2427

Laigle, C., McCracken, H. J., Ilbert, O., et al. 2016, ApJS, 224, 24, doi: 10.3847/0067-0049/224/2/24

LaMassa, S. M., Georgakakis, A., Vivek, M., et al. 2019, ApJ, 876, 50, doi: 10.3847/1538-4357/ab108b

Lange, J. U., van Dokkum, P. G., Momcheva, I. G., et al. 2016, ApJL, 819, L4, doi: 10.3847/2041-8205/819/1/L4

Lara-López, M. A., Cepa, J., Bongiovanni, A., et al. 2010, A&A, 521, L53, doi: 10.1051/0004-6361/201014803

Larson, R. L., Hutchison, T. A., Bagley, M., et al. 2023, ApJ, 958, 141, doi: 10.3847/1538-4357/acfed4

Leistedt, B., Alsing, J., Peiris, H., Mortlock, D., & Leja, J. 2023, ApJS, 264, 23, doi: 10.3847/1538-4365/ac9d99

Leistedt, B., & Hogg, D. W. 2017, ApJ, 838, 5, doi: 10.3847/1538-4357/aa6332

Leistedt, B., Hogg, D. W., Wechsler, R. H., & DeRose, J. 2019, ApJ, 881, 80, doi: 10.3847/1538-4357/ab2d29

Leistedt, B., Mortlock, D. J., & Peiris, H. V. 2016, MNRAS, 460, 4258, doi: 10.1093/mnras/stw1304

Leitner, S. N. 2012, ApJ, 745, 149, doi: 10.1088/0004-637X/745/2/149

Leja, J., Carnall, A. C., Johnson, B. D., Conroy, C., & Speagle, J. S. 2019a, ApJ, 876, 3, doi: 10.3847/1538-4357/ab133c

Leja, J., Johnson, B. D., Conroy, C., & van Dokkum, P. 2018, ApJ, 854, 62, doi: 10.3847/1538-4357/aaa8db

Leja, J., Johnson, B. D., Conroy, C., van Dokkum, P. G., & Byler, N. 2017, ApJ, 837, 170, doi: 10.3847/1538-4357/aa5ffe

Leja, J., Speagle, J. S., Johnson, B. D., et al. 2020, ApJ, 893, 111, doi: 10.3847/1538-4357/ab7e27

Leja, J., van Dokkum, P. G., Franx, M., & Whitaker, K. E. 2015, ApJ, 798, 115, doi: 10.1088/0004-637X/798/2/115

Leja, J., Johnson, B. D., Conroy, C., et al. 2019b, ApJ, 877, 140, doi: 10.3847/1538-4357/ab1d5a

Leja, J., Speagle, J. S., Ting, Y.-S., et al. 2022, ApJ, 936, 165, doi: 10.3847/1538-4357/ac887d

Levina, E., & Bickel, P. 2001, in Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001, Vol. 2, 251–256, doi: 10.1109/ICCV.2001.937632

Li, J., Melchior, P., Hahn, C., & Huang, S. 2024, AJ, 167, 16, doi: 10.3847/1538-3881/ad0be4

Li, Y., Leja, J., Johnson, B. D., et al. 2025, ApJ, 986, 9, doi: 10.3847/1538-4357/adcab4

Li, Y., Leja, J., Johnson, B. D., Tacchella, S., & Naidu, R. P. 2024, ApJL, 969, L5, doi: 10.3847/2041-8213/ad5280

Liaw, R., Liang, E., Nishihara, R., et al. 2018, arXiv e-prints, arXiv:1807.05118, doi: 10.48550/arXiv.1807.05118

Lilly, S. J., Carollo, C. M., Pipino, A., Renzini, A., & Peng, Y. 2013, ApJ, 772, 119, doi: 10.1088/0004-637X/772/2/119

Lilly, S. J., Le Fèvre, O., Renzini, A., et al. 2007, ApJS, 172, 70, doi: 10.1086/516589

Lovell, C. C., Acquaviva, V., Thomas, P. A., et al. 2019, MNRAS, 490, 5503, doi: 10.1093/mnras/stz2851

Lower, S., Narayanan, D., Leja, J., et al. 2020, ApJ, 904, 33, doi: 10.3847/1538-4357/abbfa7

—. 2022, ApJ, 931, 14, doi: 10.3847/1538-4357/ac6959

Lucie-Smith, L., Peiris, H. V., Pontzen, A., et al. 2025, arXiv e-prints, arXiv:2505.18258, doi: 10.48550/arXiv.2505.18258

Lupton, R. H., Gunn, J. E., & Szalay, A. S. 1999, AJ, 118, 1406, doi: 10.1086/301004

Lyke, B. W., Higley, A. N., McLane, J. N., et al. 2020, ApJS, 250, 8, doi: 10.3847/1538-4365/aba623

Ma, X., Hopkins, P. F., Faucher-Giguère, C.-A., et al. 2016, MNRAS, 456, 2140, doi: 10.1093/mnras/stv2659

Ma, Y., Greene, J. E., Setton, D. J., et al. 2025, ApJ, 981, 191, doi: 10.3847/1538-4357/ada613

Madau, P. 1995, ApJ, 441, 18, doi: 10.1086/175332

Madau, P., & Dickinson, M. 2014, ARA&A, 52, 415, doi: 10.1146/annurev-astro-081811-125615

Madau, P., & Rees, M. J. 2000, ApJL, 542, L69, doi: 10.1086/312934

Maddox, N., & Hewett, P. C. 2006, MNRAS, 367, 717, doi: 10.1111/j.1365-2966.2005.09984.x

Maiolino, R., & Mannucci, F. 2019, A&A Rv, 27, 3, doi: 10.1007/s00159-018-0112-2

Maiolino, R., Cirasuolo, M., Afonso, J., et al. 2020, The Messenger, 180, 24, doi: 10.18727/0722-6691/5197

Malz, A. I. 2021, PhRvD, 103, 083502, doi: 10.1103/PhysRevD.103.083502

Malz, A. I., & Hogg, D. W. 2022, ApJ, 928, 127, doi: 10.3847/1538-4357/ac062f

Mandelbaum, R. 2018, ARA&A, 56, 393, doi: 10.1146/annurev-astro-081817-051928

Mandelbaum, R., Hirata, C. M., Ishak, M., Seljak, U., & Brinkmann, J. 2006, MNRAS, 367, 611, doi: 10.1111/j.1365-2966.2005.09946.x

Mandelbaum, R., Blake, C., Bridle, S., et al. 2011, MNRAS, 410, 844, doi: 10.1111/j.1365-2966.2010.17485.x

Mannucci, F., Cresci, G., Maiolino, R., Marconi, A., & Gnerucci, A. 2010, MNRAS, 408, 2115, doi: 10.1111/j.1365-2966.2010.17291.x

Marshall, A., Auger-Williams, M. W., Banerji, M., Maiolino, R., & Bowler, R. 2022, MNRAS, 515, 5617, doi: 10.1093/mnras/stac1619

Martínez-Ramírez, L. N., Calistro Rivera, G., Lusso, E., et al. 2024, A&A, 688, A46, doi: 10.1051/0004-6361/202449329

Masters, D., Capak, P., Stern, D., et al. 2015, ApJ, 813, 53, doi: 10.1088/0004-637X/813/1/53

Mathews, E. P., Leja, J., Speagle, J. S., et al. 2023, ApJ, 954, 132, doi: 10.3847/1538-4357/ace720

Matthee, J., Naidu, R. P., Brammer, G., et al. 2024, ApJ, 963, 129, doi: 10.3847/1538-4357/ad2345

McAlpine, S., Bower, R. G., Harrison, C. M., et al. 2017, MNRAS, 468, 3395, doi: 10.1093/mnras/stx658

McAlpine, S., Bower, R. G., Rosario, D. J., et al. 2018, MNRAS, 481, 3118, doi: 10.1093/mnras/sty2489

McCarthy, I. G., Schaye, J., Bird, S., & Le Brun, A. M. C. 2017, MNRAS, 465, 2936, doi: 10.1093/mnras/stw2792

McConachie, I., Wilson, G., Forrest, B., et al. 2022, ApJ, 926, 37, doi: 10.3847/1538-4357/ac2b9f

—. 2025, ApJ, 978, 17, doi: 10.3847/1538-4357/ad8f36

McCracken, H. J., Milvang-Jensen, B., Dunlop, J., et al. 2012, A&A, 544, A156, doi: 10.1051/0004-6361/201219507

McCullough, J., Gruen, D., Amon, A., et al. 2024, MNRAS, 531, 2582, doi: 10.1093/mnras/stae1316

McLure, R. J., Pentericci, L., Cimatti, A., et al. 2018, MNRAS, 479, 25, doi: 10.1093/mnras/sty1213

Meiksin, A. 2006, MNRAS, 365, 807,
    doi: 10.1111/j.1365-2966.2005.09756.x

Merlin, E., Santini, P., Paris, D., et al. 2024, A&A, 691,
    A240, doi: 10.1051/0004-6361/202451409

Miralda-Escudé, J. 1998, ApJ, 501, 15, doi: 10.1086/305799

Mobasher, B., Dahlen, T., Ferguson, H. C., et al. 2015,
    ApJ, 808, 101, doi: 10.1088/0004-637X/808/1/101

Momcheva, I. G., Brammer, G. B., van Dokkum, P. G.,
    et al. 2016, ApJS, 225, 27,
    doi: 10.3847/0067-0049/225/2/27

Mor, R., Netzer, H., & Elitzur, M. 2009, ApJ, 705, 298,
    doi: 10.1088/0004-637X/705/1/298

Morisset, C., Charlot, S., Sánchez, S. F., et al. 2025,
    MNRAS, 538, 1884, doi: 10.1093/mnras/staf143

Moritz, P., Nishihara, R., Wang, S., et al. 2018, in
    Proceedings of the 13th USENIX Conference on
    Operating Systems Design and Implementation, OSDI'18
    (USENIX Association), 561–577.
    https://arxiv.org/abs/1712.05889

Mortlock, D. 2016, in Astrophys. Space Sci. Libr., Vol. 423,
    Understanding the Epoch of Cosmic Reionization:
    Challenges and Progress, ed. A. Mesinger, 187,
    doi: 10.1007/978-3-319-21957-8_7

Moser, B., Kacprzak, T., Fischbacher, S., et al. 2024,
    JCAP, 2024, 049, doi: 10.1088/1475-7516/2024/05/049

Moster, B. P., Somerville, R. S., Newman, J. A., & Rix,
    H.-W. 2011, ApJ, 731, 113,
    doi: 10.1088/0004-637X/731/2/113

Mountrichas, G., Masoura, V. A., Corral, A., & Carrera,
    F. J. 2024, A&A, 683, A143,
    doi: 10.1051/0004-6361/202348952

Mullaney, J. R., Pannella, M., Daddi, E., et al. 2012a,
    MNRAS, 419, 95, doi: 10.1111/j.1365-2966.2011.19675.x

Mullaney, J. R., Daddi, E., Béthermin, M., et al. 2012b,
    ApJL, 753, L30, doi: 10.1088/2041-8205/753/2/L30

Mullaney, J. R., Alexander, D. M., Aird, J., et al. 2015,
    MNRAS, 453, L83, doi: 10.1093/mnrasl/slv110

Munk, A., & Czado, C. 1998, J. R. Statistical Soc. Series B,
    60, 223, doi: 10.1111/1467-9868.00121

Nagaraj, G., Forbes, J. C., Leja, J., Foreman-Mackey, D.,
    & Hayward, C. C. 2022a, ApJ, 939, 29,
    doi: 10.3847/1538-4357/ac9477

—. 2022b, ApJ, 932, 54, doi: 10.3847/1538-4357/ac6c80

Naidu, R. P., Oesch, P. A., Brammer, G., et al. 2025, arXiv
    e-prints, arXiv:2505.11263,
    doi: 10.48550/arXiv.2505.11263

Nayyeri, H., Hemmati, S., Mobasher, B., et al. 2017, ApJS,
    228, 7, doi: 10.3847/1538-4365/228/1/7

Nenkova, M., Ivezić, Ž., & Elitzur, M. 2002, ApJL, 570, L9,
    doi: 10.1086/340857

Nenkova, M., Sirocky, M. M., Ivezić, Ž., & Elitzur, M.
    2008a, ApJ, 685, 147, doi: 10.1086/590482

Nenkova, M., Sirocky, M. M., Nikutta, R., Ivezić, Ž., &
    Elitzur, M. 2008b, ApJ, 685, 160, doi: 10.1086/590483

Newman, A. B., Rudie, G. C., Blanc, G. A., et al. 2020,
    ApJ, 891, 147, doi: 10.3847/1538-4357/ab75ee

Newman, J. A., & Gruen, D. 2022, ARA&A, 60, 363,
    doi: 10.1146/annurev-astro-032122-014611

Newman, J. A., Cooper, M. C., Davis, M., et al. 2013,
    ApJS, 208, 5, doi: 10.1088/0067-0049/208/1/5

Noboriguchi, A., Nagao, T., Toba, Y., et al. 2019, ApJ, 876,
    132, doi: 10.3847/1538-4357/ab1754

—. 2022, ApJ, 941, 195, doi: 10.3847/1538-4357/aca403

Noll, S., Pierini, D., Cimatti, A., et al. 2009, A&A, 499, 69,
    doi: 10.1051/0004-6361/200811526

Ocvirk, P., Pichon, C., Lançon, A., & Thiébaut, E. 2006,
    MNRAS, 365, 46, doi: 10.1111/j.1365-2966.2005.09182.x

Oesch, P. A., Brammer, G., Naidu, R. P., et al. 2023,
    MNRAS, 525, 2864, doi: 10.1093/mnras/stad2411

Oke, J. B., & Gunn, J. E. 1983, ApJ, 266, 713,
    doi: 10.1086/160817

Onodera, M., Renzini, A., Carollo, M., et al. 2012, ApJ,
    755, 26, doi: 10.1088/0004-637X/755/1/26

Padovani, P., Alexander, D. M., Assef, R. J., et al. 2017,
    A&A Rv, 25, 2, doi: 10.1007/s00159-017-0102-9

Panter, B., Jimenez, R., Heavens, A. F., & Charlot, S.
    2008, MNRAS, 391, 1117,
    doi: 10.1111/j.1365-2966.2008.13981.x

Park, M., Conroy, C., Johnson, B. D., et al. 2024, arXiv
    e-prints, arXiv:2410.21375,
    doi: 10.48550/arXiv.2410.21375

Paschos, P., & Norman, M. L. 2005, ApJ, 631, 59,
    doi: 10.1086/431787

Paszke, A., Gross, S., Massa, F., et al. 2019, in Advances in
    Neural Information Processing Systems, ed. H. Wallach,
    H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, &
    R. Garnett, Vol. 32 (Curran Associates, Inc.), 8024–8035.
    https://arxiv.org/abs/1912.01703

Paxton, B., Bildsten, L., Dotter, A., et al. 2011, ApJS, 192,
    3, doi: 10.1088/0067-0049/192/1/3

Paxton, B., Cantiello, M., Arras, P., et al. 2013, ApJS, 208,
    4, doi: 10.1088/0067-0049/208/1/4

Paxton, B., Marchant, P., Schwab, J., et al. 2015, ApJS,
    220, 15, doi: 10.1088/0067-0049/220/1/15

Peng, Y.-J., Lilly, S. J., Kovač, K., et al. 2010, ApJ, 721,
    193, doi: 10.1088/0004-637X/721/1/193

Pentericci, L., McLure, R. J., Garilli, B., et al. 2018, A&A,
    616, A174, doi: 10.1051/0004-6361/201833047

Petri, F., Leistedt, B., Mortlock, D. J., et al. 2025, arXiv e-prints, arXiv:2506.06475, doi: 10.48550/arXiv.2506.06475

Peyré, G., & Cuturi, M. 2019, Foundations Trends Machine Learning, 11, 355, doi: 10.1561/2200000073

Planck Collaboration, Aghanim, N., Akrami, Y., et al. 2020, A&A, 641, A6, doi: 10.1051/0004-6361/201833910

Pogson, N. 1856, MNRAS, 17, 12, doi: 10.1093/mnras/17.1.12

Polletta, M., Tajer, M., Maraschi, L., et al. 2007, ApJ, 663, 81, doi: 10.1086/518113

Popesso, P., Concas, A., Cresci, G., et al. 2023, MNRAS, 519, 1526, doi: 10.1093/mnras/stac3214

Pozzetti, L., Bolzonella, M., Zucca, E., et al. 2010, A&A, 523, A13, doi: 10.1051/0004-6361/200913020

Qiu, Y., & Kang, X. 2022, ApJ, 930, 66, doi: 10.3847/1538-4357/ac63a1

Ramdas, A., Trillos, N. G., & Cuturi, M. 2017, Entropy, 19, 47, doi: 10.3390/e19020047

Rankine, A. L., Hewett, P. C., Banerji, M., & Richards, G. T. 2020, MNRAS, 492, 4553, doi: 10.1093/mnras/staa130

Rau, M. M., Wilson, S., & Mandelbaum, R. 2020, MNRAS, 491, 4768, doi: 10.1093/mnras/stz3295

Rau, M. M., Dalal, R., Zhang, T., et al. 2023, MNRAS, 524, 5109, doi: 10.1093/mnras/stad1962

Richards, G. T., McCaffrey, T. V., Kimball, A., et al. 2021, AJ, 162, 270, doi: 10.3847/1538-3881/ac283b

Richards, G. T., Strauss, M. A., Fan, X., et al. 2006, AJ, 131, 2766, doi: 10.1086/503559

Richards, G. T., Kruczek, N. E., Gallagher, S. C., et al. 2011, AJ, 141, 167, doi: 10.1088/0004-6256/141/5/167

Riechers, D. A., Hodge, J. A., Pavesi, R., et al. 2020, ApJ, 895, 81, doi: 10.3847/1538-4357/ab8c48

Rix, S. A., Pettini, M., Leitherer, C., et al. 2004, ApJ, 615, 98, doi: 10.1086/424031

Robertson, B. E. 2022, ARA&A, 60, 121, doi: 10.1146/annurev-astro-120221-044656

Rowan-Robinson, M. 1992, MNRAS, 258, 787, doi: 10.1093/mnras/258.4.787

—. 1995, MNRAS, 272, 737, doi: 10.1093/mnras/272.4.737

Salim, S., & Boquien, M. 2019, ApJ, 872, 23, doi: 10.3847/1538-4357/aaf88a

Salim, S., Boquien, M., & Lee, J. C. 2018, ApJ, 859, 11, doi: 10.3847/1538-4357/aabf3c

Salim, S., & Narayanan, D. 2020, ARA&A, 58, 529, doi: 10.1146/annurev-astro-032620-021933

Salim, S., Lee, J. C., Janowiecki, S., et al. 2016, ApJS, 227, 2, doi: 10.3847/0067-0049/227/1/2

Sana, H., de Mink, S. E., de Koter, A., et al. 2012, Science, 337, 444, doi: 10.1126/science.1223344

Sánchez, C., & Bernstein, G. M. 2019, MNRAS, 483, 2801, doi: 10.1093/mnras/sty3222

Sánchez-Blázquez, P., Peletier, R. F., Jiménez-Vicente, J., et al. 2006, MNRAS, 371, 703, doi: 10.1111/j.1365-2966.2006.10699.x

Sandles, L., Curtis-Lake, E., Charlot, S., Chevallard, J., & Maiolino, R. 2022, MNRAS, 515, 2951, doi: 10.1093/mnras/stac1999

Sawicki, M., Arnouts, S., Huang, J., et al. 2019, MNRAS, 489, 5202, doi: 10.1093/mnras/stz2522

Schaye, J., Kugel, R., Schaller, M., et al. 2023, MNRAS, 526, 4978, doi: 10.1093/mnras/stad2419

Schroeder, J., Mesinger, A., & Haiman, Z. 2013, MNRAS, 428, 3058, doi: 10.1093/mnras/sts253

Scoville, N., Aussel, H., Brusa, M., et al. 2007a, ApJS, 172, 1, doi: 10.1086/516585

Scoville, N., Aussel, H., Benson, A., et al. 2007b, ApJS, 172, 150, doi: 10.1086/516751

Setton, D. J., Greene, J. E., de Graaff, A., et al. 2024, arXiv e-prints, arXiv:2411.03424, doi: 10.48550/arXiv.2411.03424

Shuntov, M., Ilbert, O., Toft, S., et al. 2025a, A&A, 695, A20, doi: 10.1051/0004-6361/202452570

Shuntov, M., Akins, H. B., Paquereau, L., et al. 2025b, arXiv e-prints, arXiv:2506.03243, doi: 10.48550/arXiv.2506.03243

Siebenmorgen, R., Heymann, F., & Efstathiou, A. 2015, A&A, 583, A120, doi: 10.1051/0004-6361/201526034

Sikorski, D., Forrest, B., Lemaux, B. C., et al. 2025, arXiv e-prints, arXiv:2509.02714. https://arxiv.org/abs/2509.02714

Silva, L., Granato, G. L., Bressan, A., & Danese, L. 1998, ApJ, 509, 103, doi: 10.1086/306476

Skelton, R. E., Whitaker, K. E., Momcheva, I. G., et al. 2014, ApJS, 214, 24, doi: 10.1088/0067-0049/214/2/24

Smirnov, N. 1948, Ann. Math. Statistics, 19, 279 , doi: 10.1214/aoms/1177730256

Söchting, I. K., Coldwell, G. V., Clowes, R. G., Campusano, L. E., & Graham, M. J. 2012, MNRAS, 423, 2436, doi: 10.1111/j.1365-2966.2012.21050.x

Sohl-Dickstein, J., Weiss, E., Maheswaranathan, N., & Ganguli, S. 2015, in Proceedings of Machine Learning Research, Vol. 37, Proceedings of the 32nd International Conference on Machine Learning, ed. F. Bach & D. Blei (PMLR), 2256–2265. https://arxiv.org/abs/1503.03585

Somerville, R. S., & Davé, R. 2015, ARA&A, 53, 51, doi: 10.1146/annurev-astro-082812-140951

Sommariva, V., Mannucci, F., Cresci, G., et al. 2012, A&A, 539, A136, doi: 10.1051/0004-6361/201118134

Sommovigo, L., Cochrane, R. K., Somerville, R. S., et al. 2025, ApJ, 990, 114, doi: 10.3847/1538-4357/addec1

Song, J., Meng, C., & Ermon, S. 2021, in 9th International Conference on Learning Representations. https://arxiv.org/abs/2010.02502

Song, Y., & Ermon, S. 2019, in Advances in Neural Information Processing Systems, ed. H. M. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. B. Fox, & R. Garnett, Vol. 32, 11895–11907. https://arxiv.org/abs/1907.05600

Song, Y., Sohl-Dickstein, J., Kingma, D. P., et al. 2021, in 9th International Conference on Learning Representations. https://arxiv.org/abs/2011.13456

Speagle, J. S., Steinhardt, C. L., Capak, P. L., & Silverman, J. D. 2014, ApJS, 214, 15, doi: 10.1088/0067-0049/214/2/15

Spergel, D., Gehrels, N., Baltay, C., et al. 2015, arXiv e-prints, arXiv:1503.03757, doi: 10.48550/arXiv.1503.03757

Stalevski, M., Ricci, C., Ueda, Y., et al. 2016, MNRAS, 458, 2288, doi: 10.1093/mnras/stw444

Stanton, T. M., Cullen, F., McLure, R. J., et al. 2024, MNRAS, 532, 3102, doi: 10.1093/mnras/stae1705

Stanway, E. R., & Eldridge, J. J. 2018, MNRAS, 479, 75, doi: 10.1093/mnras/sty1353

Stanway, E. R., Eldridge, J. J., Greis, S. M. L., et al. 2014, MNRAS, 444, 3466, doi: 10.1093/mnras/stu1682

Steidel, C. C., Strom, A. L., Pettini, M., et al. 2016, ApJ, 826, 159, doi: 10.3847/0004-637X/826/2/159

Steidel, C. C., Rudie, G. C., Strom, A. L., et al. 2014, ApJ, 795, 165, doi: 10.1088/0004-637X/795/2/165

Steinhardt, C. L., Speagle, J. S., Capak, P., et al. 2014, ApJL, 791, L25, doi: 10.1088/2041-8205/791/2/L25

Stern, D., Eisenhardt, P., Gorjian, V., et al. 2005, ApJ, 631, 163, doi: 10.1086/432523

Stern, D., Assef, R. J., Benford, D. J., et al. 2012, ApJ, 753, 30, doi: 10.1088/0004-637X/753/1/30

Stoughton, C., Lupton, R. H., Bernardi, M., et al. 2002, AJ, 123, 485, doi: 10.1086/324741

Strom, A. L., Rudie, G. C., Steidel, C. C., & Trainor, R. F. 2022, ApJ, 925, 116, doi: 10.3847/1538-4357/ac38a3

Strom, A. L., Steidel, C. C., Rudie, G. C., et al. 2017, ApJ, 836, 164, doi: 10.3847/1538-4357/836/2/164

Suh, H., Civano, F., Hasinger, G., et al. 2019, ApJ, 872, 168, doi: 10.3847/1538-4357/ab01fb

Sun, F., Helton, J. M., Egami, E., et al. 2024, ApJ, 961, 69, doi: 10.3847/1538-4357/ad07e3

Szalay, A. S., Connolly, A. J., & Szokoly, G. P. 1999, AJ, 117, 68, doi: 10.1086/300689

Tanaka, M. 2015, ApJ, 801, 20, doi: 10.1088/0004-637X/801/1/20

Taniguchi, Y., Scoville, N., Murayama, T., et al. 2007, ApJS, 172, 9, doi: 10.1086/516596

Taniguchi, Y., Kajisawa, M., Kobayashi, M. A. R., et al. 2015, PASJ, 67, 104, doi: 10.1093/pasj/psv106

Temple, M. J., Ferland, G. J., Rankine, A. L., Chatzikos, M., & Hewett, P. C. 2021a, MNRAS, 505, 3247, doi: 10.1093/mnras/stab1610

Temple, M. J., Hewett, P. C., & Banerji, M. 2021b, MNRAS, 508, 737, doi: 10.1093/mnras/stab2586

Temple, M. J., Matthews, J. H., Hewett, P. C., et al. 2023, MNRAS, 523, 646, doi: 10.1093/mnras/stad1448

Tepper-García, T., & Fritze, U. 2008, MNRAS, 383, 1671, doi: 10.1111/j.1365-2966.2007.12691.x

Thomas, R., Le Fèvre, O., Le Brun, V., et al. 2017, A&A, 597, A88, doi: 10.1051/0004-6361/201425342

Thomas, R., Pentericci, L., Le Fèvre, O., et al. 2020, A&A, 634, A110, doi: 10.1051/0004-6361/201935925

—. 2021, A&A, 650, A63, doi: 10.1051/0004-6361/202038438

Thorp, S., Alsing, J., Peiris, H. V., et al. 2024a, ApJ, 975, 145, doi: 10.3847/1538-4357/ad7736

—. 2025a, pop-cosmos: Galaxy property and redshift catalog for COSMOS2020, v1.3.0, Zenodo, doi: 10.5281/zenodo.14832248

Thorp, S., Peiris, H. V., Mortlock, D. J., et al. 2025b, ApJS, 276, 5, doi: 10.3847/1538-4365/ad8ebd

—. 2024b, stevet40/quantile_utilities, v1.0.0, Zenodo, doi: 10.5281/zenodo.14018635

Thorp, S., Peiris, H. V., Jagwani, G., et al. 2025c, pop-cosmos: COSMOS-like mock galaxy catalogs, v1.0.1, Zenodo, doi: 10.5281/zenodo.15757717

—. 2025d, pop-cosmos: Galaxy property and redshift catalog for COSMOS2020, v2.1.1, Zenodo, doi: 10.5281/zenodo.15850536

Timlin, J. D., Brandt, W. N., Ni, Q., et al. 2020, MNRAS, 492, 719, doi: 10.1093/mnras/stz3433

Tinsley, B. M. 1980, FCPh, 5, 287, doi: 10.48550/arXiv.2203.02041

Tol, P. 2021, Colour Schemes, Tech. Rep. SRON/EPS/TN/09-002, Issue 3.2. https://sronpersonalpages.nl/~pault/data/colourschemes.pdf

Toni, G., Gozaliasl, G., Maturi, M., et al. 2025, A&A, 697, A197, doi: 10.1051/0004-6361/202553759

Tortorelli, L., Fischbacher, S., Grün, D., et al. 2025, arXiv e-prints, arXiv:2505.21610, doi: 10.48550/arXiv.2505.21610

Tortorelli, L., McCullough, J., & Gruen, D. 2024, A&A, 689, A144, doi: 10.1051/0004-6361/202450694

Totani, T., Kawai, N., Kosugi, G., et al. 2006, PASJ, 58, 485, doi: 10.1093/pasj/58.3.485

Tremonti, C. A., Heckman, T. M., Kauffmann, G., et al. 2004, ApJ, 613, 898, doi: 10.1086/423264

Troxel, M. A., & Ishak, M. 2015, PhR, 558, 1, doi: 10.1016/j.physrep.2014.11.001

Urry, C. M., & Padovani, P. 1995, PASP, 107, 803, doi: 10.1086/133630

Vaserstein, L. N. 1969, Problemy Peredachi Inf., 5, 64

Vega, L. R., Asari, N. V., Cid Fernandes, R., et al. 2009, MNRAS, 393, 846, doi: 10.1111/j.1365-2966.2008.13969.x

Villaume, A., Conroy, C., & Johnson, B. D. 2015, ApJ, 806, 82, doi: 10.1088/0004-637X/806/1/82

Vincent, P. 2011, Neural Comput., 23, 1661, doi: 10.1162/NECO_a_00142

Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, Nature Methods, 17, 261, doi: 10.1038/s41592-019-0686-2

Wang, B., Leja, J., Villar, V. A., & Speagle, J. S. 2023a, ApJL, 952, L10, doi: 10.3847/2041-8213/ace361

Wang, B., Leja, J., Bezanson, R., et al. 2023b, ApJL, 944, L58, doi: 10.3847/2041-8213/acba99

Wang, B., de Graaff, A., Davies, R. L., et al. 2025, ApJ, 984, 121, doi: 10.3847/1538-4357/adc1ca

Weaver, J. R., Kauffmann, O. B., Ilbert, O., et al. 2022, ApJS, 258, 11, doi: 10.3847/1538-4365/ac3078

Weaver, J. R., Zalesky, L., Kokorev, V., et al. 2023a, ApJS, 269, 20, doi: 10.3847/1538-4365/acf850

Weaver, J. R., Davidzon, I., Toft, S., et al. 2023b, A&A, 677, A184, doi: 10.1051/0004-6361/202245581

Wechsler, R. H., & Tinker, J. L. 2018, ARA&A, 56, 435, doi: 10.1146/annurev-astro-081817-051756

Weibel, A., Oesch, P. A., Barrufet, L., et al. 2024, MNRAS, 533, 1808, doi: 10.1093/mnras/stae1891

Weston, M. E., McIntosh, D. H., Brodwin, M., et al. 2017, MNRAS, 464, 3882, doi: 10.1093/mnras/stw2620

Whitaker, K. E., van Dokkum, P. G., Brammer, G., & Franx, M. 2012, ApJL, 754, L29, doi: 10.1088/2041-8205/754/2/L29

Whitaker, K. E., Franx, M., Leja, J., et al. 2014, ApJ, 795, 104, doi: 10.1088/0004-637X/795/2/104

White, R. L., Becker, R. H., Fan, X., & Strauss, M. A. 2003, AJ, 126, 1, doi: 10.1086/375547

Wilk, M. B., & Gnanadesikan, R. 1968, Biometrika, 55, 1, doi: 10.1093/biomet/55.1.1

Wiseman, P., Sullivan, M., Smith, M., & Popovic, B. 2023, MNRAS, 520, 6214, doi: 10.1093/mnras/stad488

Wiseman, P., Vincenzi, M., Sullivan, M., et al. 2022, MNRAS, 515, 4587, doi: 10.1093/mnras/stac1984

Witt, A. N., & Gordon, K. D. 2000, ApJ, 528, 799, doi: 10.1086/308197

Worthey, G. 1994, ApJS, 95, 107, doi: 10.1086/192096

Wright, A. H., Kuijken, K., Hildebrandt, H., et al. 2024, A&A, 686, A170, doi: 10.1051/0004-6361/202346730

Wuyts, S., Förster Schreiber, N. M., Lutz, D., et al. 2011, ApJ, 738, 106, doi: 10.1088/0004-637X/738/1/106

Xiao, L., Stanway, E. R., & Eldridge, J. J. 2018, MNRAS, 477, 904, doi: 10.1093/mnras/sty646

Xiao, M., Oesch, P. A., Elbaz, D., et al. 2024, Nature, 635, 311, doi: 10.1038/s41586-024-08094-5

York, D. G., Adelman, J., Anderson, Jr., J. E., et al. 2000, AJ, 120, 1579, doi: 10.1086/301513

Zahid, H. J., Dima, G. I., Kudritzki, R.-P., et al. 2014, ApJ, 791, 130, doi: 10.1088/0004-637X/791/2/130

Zahid, H. J., Kudritzki, R.-P., Conroy, C., Andrews, B., & Ho, I. T. 2017, ApJ, 847, 18, doi: 10.3847/1538-4357/aa88ae

Zahid, H. J., Yates, R. M., Kewley, L. J., & Kudritzki, R. P. 2013, ApJ, 763, 92, doi: 10.1088/0004-637X/763/2/92