Inferring More from Less: Prospector as a Photometric Redshift Engine in the Era of JWST

Bingjie Wang (王冰洁)[1,2,3] , Joel Leja[1,2,3] , Rachel Bezanson[4] , Benjamin D. Johnson[5] , Gourav Khullar[4] , Ivo Labbé[6] ,
Sedona H. Price[4] , John R. Weaver[7] , and Katherine E. Whitaker[7,8]

[1] Department of Astronomy & Astrophysics, The Pennsylvania State University, University Park, PA 16802, USA; bwang@psu.edu
[2] Institute for Computational & Data Sciences, The Pennsylvania State University, University Park, PA 16802, USA
[3] Institute for Gravitation and the Cosmos, The Pennsylvania State University, University Park, PA 16802, USA
[4] Department of Physics & Astronomy and PITT PACC, University of Pittsburgh, Pittsburgh, PA 15260, USA
[5] Center for Astrophysics | Harvard & Smithsonian, Cambridge, MA 02138, USA
[6] Centre for Astrophysics and Supercomputing, Swinburne University of Technology, Melbourne, VIC 3122, Australia
[7] Department of Astronomy, University of Massachusetts, Amherst, MA 01003, USA
[8] Cosmic Dawn Center (DAWN), Niels Bohr Institute, University of Copenhagen, Jagtvej 128, København N, DK-2200, Denmark
Received 2022 November 28; revised 2023 February 8; accepted 2023 February 9; published 2023 February 23

Abstract

The advent of the James Webb Space Telescope (JWST) signals a new era in exploring galaxies in the high-z universe. Current and upcoming JWST imaging will potentially detect galaxies at $z \sim 20$, creating a new urgency in the quest to infer accurate photometric redshifts (photo-z) for individual galaxies from their spectral energy distributions, as well as masses, ages, and star formation rates. Here we illustrate the utility of informed priors encoding previous observations of galaxies across cosmic time in achieving these goals. We construct three joint priors encoding empirical constraints of redshifts, masses, and star formation histories in the galaxy population within the `Prospector` Bayesian inference framework. In contrast with uniform priors, our model breaks an age–mass–redshift degeneracy, and thus reduces the mean bias error in masses from 0.3 to 0.1 dex, and in ages from 0.6 to 0.2 dex in tests done on mock JWST observations. Notably, our model recovers redshifts at least as accurately as the state-of-the-art photo-z code `EAzY` in deep JWST fields, but with two advantages: tailoring a model based on a particular survey is rendered mostly unnecessary given well-motivated priors; obtaining joint posteriors describing stellar, active galactic nuclei, gas, and dust contributions becomes possible. We can now confidently use the joint distribution to propagate full non-Gaussian redshift uncertainties into inferred properties of the galaxy population. This model, "`Prospector`-β," is intended for fitting galaxy photometry where the redshift is unknown, and will be instrumental in ensuring the maximum science return from forthcoming photometric surveys with JWST. The code is made publicly available online as a part of `Prospector` [9].

Unified Astronomy Thesaurus concepts: Bayesian statistics (1900); Computational astronomy (293); Galaxy evolution (594); Galaxy formation (595); Redshift surveys (1378); Spectrophotometry (1556); Spectral energy distribution (2129)

1. Introduction

Within months of the first data releases, JWST has already begun to revolutionize our view of galaxy formation (e.g., Finkelstein et al. 2022a; Treu et al. 2022; Windhorst et al. 2023). Fitting spectral energy distributions (SEDs) plays a key role in grounding new observations to theories of galaxy evolution. State-of-the-art Bayesian codes, including `BAG-PIPES` (Carnall et al. 2018), `BEAGLE` (Chevallard & Charlot 2016), and `Prospector` (Johnson et al. 2021), have been applied extensively in this context. One of the most advanced models in this family is `Prospector`-α (Leja et al. 2017). It incorporates galaxy properties including nonparametric star formation histories (SFHs; Leja et al. 2019), self-consistent nebular emission (Byler et al. 2017), and a sophisticated dust model (e.g., Lower et al. 2022), and

critically, fits all parameters at once with Markov Chain Monte Carlo or nested sampling to produce joint high-dimensional constraints. `Prospector`-α has been designed specifically to extract the most information from high signal-to-noise photometry or spectra with good wavelength coverage. Therefore, it has been traditionally applied to well-sampled SEDs using redshifts measured externally, e.g., from spectra or photo-z codes.

With regard to photo-z codes, considerable efforts have been devoted to various developments (see Salvato et al. 2019; Newman & Gruen 2022 for recent reviews, and also Alsing et al. 2023; Leistedt et al. 2023 for general discussions on photo-z frameworks). The most common algorithm is inferring redshift by comparing observations to SED templates, as is used in `LePhare` (Arnouts et al. 1999; Ilbert et al. 2006), `BPZ` (Benítez 2000), `ZEBRA` (Feldmann et al. 2006), `EAzY` (Brammer et al. 2008), and others. This method suffers from the fact that template colors are frequently degenerate with redshift. To mitigate this problem, some codes add on a magnitude-dependent redshift prior (e.g., `BPZ`, `EAzY`), the probability distribution of which is determined by fitting redshift distributions of existing catalogs in magnitude bins. However, few codes have adopted a full Bayesian approach, in which priors can be assigned to every parameter. Conversely,

---

[9] The version used in this work corresponds to the state of the Git repository at commit https://github.com/bd-j/prospector/commit/820ad72363a1f9c22cf03610bfe6e361213385cd.

codes that are capable of mapping the multidimensional posterior probability distributions for redshift and galaxy properties simultaneously via advanced sampling techniques, as mentioned earlier, are not optimized for fitting SEDs when the redshift is unknown.

The era of JWST creates an urgent new opportunity to measure the highly uncertain early phases of galaxy formation, and, for the first time, necessitates allowing redshift solutions over the range of $0 < z \lesssim 20$. Thus there is an increasing need for accurately inferring joint, correlated constraints between redshift and important galaxy parameters like mass, star formation rate (SFR), and dust properties where exquisite wavelength coverage and/or external spectra are lacking.

In this work, we establish a framework in which empirical constraints of galaxy evolution are encoded directly into the inference process as informative priors, so that redshift and galaxy properties can be constrained simultaneously and accurately. It is a known challenge that the photo-z accuracy depends strongly on the assumptions that go into the code (e.g., Kodra et al. 2023). Here we propose that the most accurate assumptions are our previous observations. A common argument for assuming uniform priors when fitting for high-z objects is that the epoch in question is largely unexplored and hence uniform priors are the least sinful (e.g., Finkelstein et al. 2022b). Here we argue for the opposite: priors represent our belief about which solutions we think are more probable, and it is thus always preferable to use well-motivated priors if available.

We present a new model, Prospector-β, optimized to recover photo-z in deep JWST fields, while taking full advantage of the capability of Prospector to produce a high-dimensional SED-model and obtain joint constraints on all inferred physical parameters. This means that the full probability distribution can be used to propagate full non-Gaussian redshift uncertainties into inferred properties of the galaxy population. Doing so will significantly enhance our confidence in the inferred properties, and will thus maximize the information returned from JWST.

We devise three new priors: a mass function prior, a galaxy number density prior, and a dynamic nonparametric SFH prior that reflects the consistent observational finding that massive galaxies form much earlier than low-mass galaxies (Cowie et al. 1996; Thomas et al. 2005). Our SFH prior also respects the observed cosmic star formation rate density by encouraging rising histories early in the universe, and falling histories late in the universe.

Moreover, we identify and characterize an age–mass–redshift degeneracy that contaminates the results of standard uniform priors. We show that our model is able to break this degeneracy, while recovering redshifts at least as accurately as EAzY in JWST surveys.

The Letter is structured as follows. Section 2 states the simulated and the observed photometry on which we assess our model performance. Section 3 details the construction of the three observationally informed priors. Section 4 describes the procedure of SED fitting. Section 5 reports the results. Section 6 summarizes our findings and concludes with implications for future works. When applicable, we adopt the best-fit cosmological parameters from the 9 yr results from the Wilkinson Microwave Anisotropy Probe mission: $H_0 = 69.32$ km s^{-1} Mpc^{-1}, $\Omega_M = 0.2865$, and $\Omega_\Lambda = 0.7135$ (Bennett et al. 2013).

2. Data

While the framework devised in this Letter is applicable to any data, we scope the current work to JWST surveys in light of the immediate need to explore the parameter space that has just been opened by JWST. With this in mind, we construct our simulation imitating a JWST survey to quantify the effect of the proposed priors.

The mock catalog is based on JAGUAR, which describes the evolution of galaxies across $0 \leqslant z \leqslant 15$ and $7 \leqslant \log M_\star / M_\odot \leqslant 13$ (Williams et al. 2018). It uses observed stellar mass and UV luminosity functions at $0 < z < 10$ (Bouwens et al. 2015; Oesch et al. 2018) to model the evolution of stellar mass, UV absolute magnitude, UV continuum slope, and galaxy number counts. The relevant parameters then become the inputs to BEAGLE (Chevallard & Charlot 2016) to generate SEDs in JAGUAR.

We note that BEAGLE and Prospector incorporate different stellar population synthesis: the former uses Bruzual & Charlot (2003), while the latter calls FSPS (Conroy & Gunn 2010; Choi et al. 2016; Dotter 2016). The underlying nebular emission, stellar templates, and isochrones all differ. In order to isolate the effect of the prior and avoid concerns about template differences, we reconstruct the mock catalog using identical inputs for FSPS. The stellar populations are modeled with a Chabrier initial mass function (Chabrier 2003), and a two-component dust model (Charlot & Fall 2000). The SFH is parameterized by a delayed-τ model.

We add noise typical for a deep field JWST + HST survey, then impose a $z \geqslant 0.5$ cut and a signal-to-noise $\geqslant 5$ cut based on the flux in the F444W band, which result in about 56,000 mock galaxies. Further details on the mock catalog can be found in the Appendix.

In addition to the simulations, we test our model on JWST data as well. The observations are recently acquired in a Director's Discretionary (DD) program (JWST-DD-2756, PI: Chen). We crossmatch the sources to collect spectroscopic redshifts from the NASA/IPAC Extragalactic Database, Treu et al. (2015), and Richard et al. (2021). Data reduction follows Weaver et al. (2023). The observed fluxes in the F444W band span roughly from 16.5 to 30.5 AB mag, with a 5σ depth of 25.5 AB mag. Star-forming galaxies constitute the majority of this sample.

3. Galaxy Evolution Priors

In this section, we detail the construction of three observationally informed priors on galaxy stellar mass, the observable number density, and galaxy star formation history SFH(M,z). We contrast each with the respective uniform prior in Figure 1.

3.1. Stellar Mass Function

Galaxy stellar mass can be strongly constrained from photometry when the distance is known (Bell & de Jong 2001; Pforr et al. 2012; Conroy et al. 2013; Mobasher et al. 2015; Li & Leja 2022); however, when the distance is unknown, inferring mass becomes challenging (mass refers to stellar mass hereafter). An informative prior can be helpful in this case. In nature, low-mass galaxies are far more numerous than high-mass galaxies, so a mass function prior allows us to avoid spurious high-mass, high-z solutions. For a given z, we draw mass from $P(M|z)$, which is just a mass function, Φ, at that
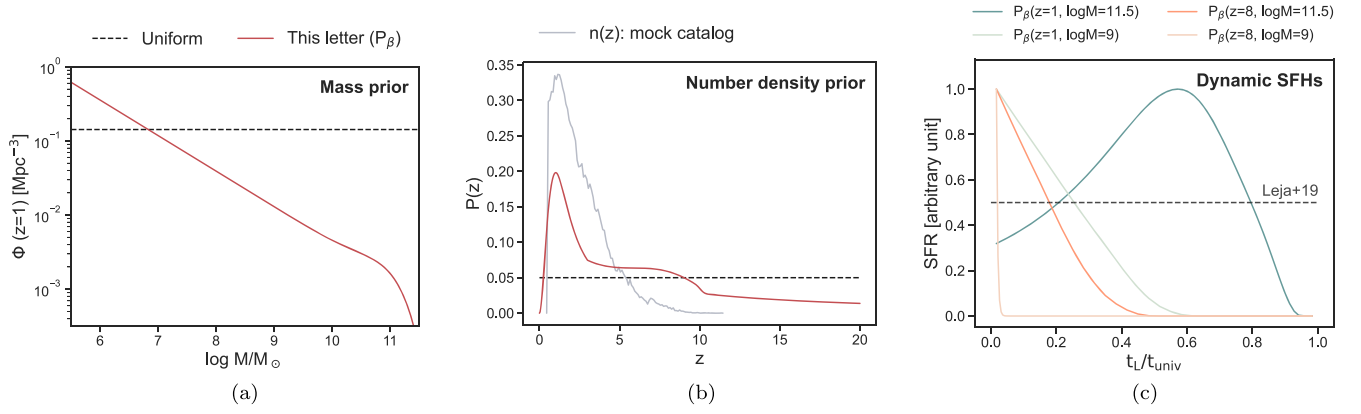
Figure 1. The three panels illustrate the difference between the proposed priors on mass, number density, and SFH of this Letter, and uniform priors (black dashed lines). (a) Mass function at $z = 1$ (Leja et al. 2020) is shown in red. (b) Observable galaxy number density (i.e., M_c cut applied), as a function of redshift is plotted in red. Also shown in gray is the number density directly calculated from the mock catalog, which differs from our $P(z)$ due to the different assumptions built in; in particular, mass functions at $z \gtrsim 5$. Nevertheless, we are able to improve the parameter inference by assuming a reasonable approximation of the mass functions used in the mock catalog. (c) Four SFHs from our prior are shown as colored lines. The x-axis is lookback time, t_L, normalized by the age of the universe, t_{univ}, at the respective z, and the y-axis is also normalized so that all SFRs can be shown on the same scale.

redshift normalized such that

$$\int_{M_{\min}}^{M_{\max}} \Phi(z) \, dM = 1, \qquad (1)$$

where $M_{\min}$ and $M_{\max}$ are the minimum and maximum masses. We set $M_{\min} = 10^6 M_\odot$ and $M_{\max} = 10^{12} M_\odot$ in this work.

In this Letter, we take the mass functions from Leja et al. (2020), which are continuous in redshift and are derived using the same stellar population model, `Prospector-α`. We note that they are only defined between $0.2 \leqslant z \leqslant 3.0$. For $z < 0.2$ and $z > 3$, we adopt a nearest-neighbor solution, i.e., the $z = 0.2$ and $z = 3$ mass functions. The nearest-neighbor choice may be favored in the absence of reliable high-resolution rest-frame optical selected mass functions at $z > 3$. Importantly, it allots a conservatively high probability for yet-to-be-discovered populations of high-mass, high-z galaxies, hints of which have already been observed in Labbe et al. (2022).

Conversely, if the goal is to decrease as much as possible the contamination of interlopers posing as extremely luminous galaxies at $z \gtrsim 10$ (Furlanetto & Mirocha 2022), one may consider transitioning over to theoretical mass functions. To this end, we additionally provide an option in the public release of our code: the mass functions can be switched to those from Tacchella et al. (2018) between $4 < z < 12$, with the $3 < z < 4$ transition from Leja et al. (2020) managed with a smoothly varying average in number density space. These can easily be replaced with observed mass functions at high z once they become available. While at this stage the high-z mass functions remain uncertain, we note that even our current estimate is inaccurate, including a mass function prior, which is a better choice than a flat prior. As can be seen from Figure 1(a), the number density at $10^9 M_\odot$ at $z = 1$ already deviates from the uniform prior by over 1 order of magnitude.

3.2. Galaxy Number Density

The probability of finding a galaxy at redshift z can be estimated from

$$P(z) = N(z) \, V_{co}(z) = \int_{M_c}^{\infty} \Phi(M, z) \, dM \, V_{co}(z), \qquad (2)$$

where $N(z)$ is the expected total number density of galaxies, $V_{co}(z)$ is the differential comoving volume, and M_c is the mass

completion limit. We normalize $P(z)$ by letting $\int P(z) = 1$. We show $P(z)$ adopted for this work in Figure 1(b).

As shown in Figure 2(a), the number density, and hence $P(z)$, is sensitive to the mass-completeness limit of the data. This means that a careful analysis must be carried out to estimate $M_c(z)$. In practice one would likely need to obtain M_c from using SED-modeling heuristics based on the flux-completeness limits in a given catalog. Many subtleties exist in such translation due to the complexities of stellar populations (e.g., Pozzetti et al. 2010; Leja et al. 2020), the discussion of which is beyond the scope of this work. For simplicity, we extract M_c directly from our mock catalog as the mass above which 90% of the galaxies in a given redshift bin is included.

3.3. Dynamic Star Formation History

An advantage of `Prospector-α` is the ability to adapt flexible SFHs. The SFH is set by the mass formed in seven logarithmically spaced time bins, assuming a continuity prior that ensures smooth transitions between bins (Leja et al. 2019). While the flat SFH in `Prospector-α` is a reasonable null expectation, it is inconsistent with expectations from real galaxies: they have rising SFHs in the early universe and falling SFHs at late times, and also show systematic trends with mass (e.g., Leitner 2012; Pacifici et al. 2016). As pointed out in Tacchella et al. (2022a) and Whitler et al. (2023), the inferred early formation history of high-z galaxies heavily depends on the SFH prior.

We therefore modify the flat SFH assumption by matching the expectation value in each bin in lookback time to the cosmic star formation rate densities (SFRD) in Behroozi et al. (2019), while still keeping the distribution about the mean identical to the Student's-t distribution in `Prospector-α`. Taking a step further, we introduce a mass dependence to the SFH prior, since it is well established that galaxy SFHs depend both on mass and cosmic time (Behroozi et al. 2019). A difficulty, though, is that SFH constraints require highly informative spectroscopy (e.g., Carnall et al. 2019; Leja et al. 2019; Tacchella et al. 2022b; Park et al. 2022). In the spirit of keeping the priors minimalistic, we opt to model the mass dependence by shifting the start of the age bins as

$$\log(t_{\text{start}}/\text{Gyr}) = \log(t_{\text{univ}}(z)/\text{Gyr}) + \delta_m, \qquad (3)$$
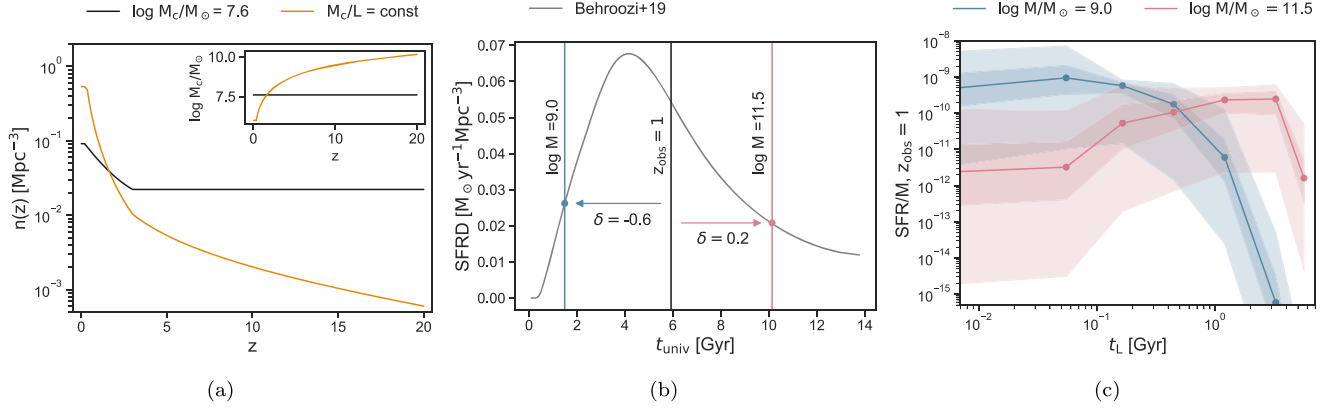
Figure 2. Here we show the technical details of the implementation of our priors from observations of galaxy evolution. (a) Number densities of galaxies, n, assuming different evolutions of the mass complete limit, M_c, are plotted as functions of redshift. Shown in black is n based on a constant M_c, while shown in dark orange is based on a constant mass-to-light ratio. The inserted panel shows M_c as a function of z. (b) A schematic diagram demonstrating our δ formalism, which introduces mass dependence to the SFH prior. The cosmic SFRD (Behroozi et al. 2019), which we take to be the expectation value, is plotted as a function of the age of the universe in gray. At an observed redshift, z_{obs}, the expected SFR values of a galaxy are shifted back/forward in time depending on its mass. The constructed SFH prior favors rising SFHs in the early universe, and falling SFHs in the late universe. (c) Specific SFRs of two galaxies of masses $10^8 M_\odot$ and $10^{10.5} M_\odot$, observed at $z = 1$, are plotted as functions of lookback time in blue and red, respectively. The x-axis is the time between the onset of star formation (birth of the galaxy) and $z = 1$. The shades indicate the 16–84th and the 2.5–97.5th percentiles of the prior distribution.

where

$$\delta_m = \begin{cases} -0.6, & \text{for } \log M/M_\odot < 9. \\ 1/3 \log M - 3.6, & \text{for } 9 \leqslant \log M/M_\odot \leqslant 12. \\ 0.4, & \text{for } \log M/M_\odot > 12. \end{cases} \quad (4)$$

$t_{univ}(z)$ is the age of the universe at an observed z. This does not, however, mean that the onset of star formation can be earlier than the Big Bang. The equation refers to the age of the universe at which the shape of the SFRD is taken as a prior for the expectation values of the time bins. In practice, we shift the start of the age bins along the SFRD curve, and then rescale the age bins to be within the age of the universe at a given z. This SFH prior effectively encodes an expectation that high-mass galaxies form earlier, and low-mass galaxies form later, than naive expectations from the cosmic SFRD. In other words, the physical motivation is the downsizing scenario in galaxy formation (Cowie et al. 1996; Thomas et al. 2005). We show four SFHs from our prior in Figure 1(c) as examples, and technical details of the δ formalism in Figures 2(b)–(c).

This prior does not, however, account for the quiescent/star-forming bimodality. We effectively approximate a double-peaked distribution of SFRs at a given mass and redshift seen in, e.g., Leja et al. (2022) as a wide single-peaked distribution. The quiescent fractions calculated from our prior match roughly with the observed trend at $z < 3$ (Leja et al. 2022), which justifies this simplification.

Having presented the new priors, we note that we include the stellar mass–stellar metallicity relationship measured from the Sloan Digital Sky Survey as a prior as well (Gallazzi et al. 2005). Following Leja et al. (2019), we take the conservative approach of widening the confidence intervals from this relationship by a factor of 2 to account for potential unknown systematics or redshift evolution.

4. Fitting Mock Galaxies

We fit SEDs with the Prospector-α model, as well as Prospector-β of this Letter, which essentially replaces the former's uniform priors with informative ones. We emphasize

that the physical model shared between Prospector-α and Prospector-β is more sophisticated than the one used for generating the mock galaxies. It consists of 16 free parameters describing the contribution of stars, active galactic nuclei (AGNs), gas, and dust (Table 1), although the dust emission and the infrared AGN parameters are only relevant at $z \lesssim 1$. The posteriors are sampled with the dynamic nested sampling code dynesty (Speagle 2020). We fit 5000 randomly drawn mock galaxies. This sample size is limited by the time required of FSPS/nested sampling, which takes >10 hr per fit. This problem will be circumvented via machine-learning-accelerated techniques in the near term (Alsing et al. 2020; Wang et al. 2022).

For comparison, we determine redshift using EAzY, which fits a nonnegative linear combination of a set of templates (Brammer et al. 2008). We use the vanilla version with the default set of tweak_fsps_QSF_12_v3.

5. Results

In this section, we report the parameter recoveries using the Prospector-β priors, and contrast them with the uniform priors. We also compare to the redshift inferred via the vanilla EAzY templates to put the redshift recovery in context of previous results in the literature. The medians of posterior distributions are quoted in all analyses.

We show histograms of residuals for redshift in Figure 3, and key physical parameters in Figure 4. The scatter in residuals is quantified using the normalized median absolute deviation (NMAD) given its advantage of being less sensitive to outliers than rms. It is defined as

$$\sigma_{NMAD} = 1.48 \times \text{median}|\Delta\theta|, \quad (5)$$

where $\Delta\theta = \theta_{med} - \theta_{true}$, with one exception: wherever we evaluate the accuracy of redshift inference, we replace $\Delta\theta$ with $\Delta z = (z_{med} - z_{true})/(1 + z_{true})$. We additionally quantify an outlier fraction, f_{out}, in which we define a catastrophic outlier as one with $\Delta\theta > 0.1$, and bias calculated using the mean bias

Table 1
Parameters and Priors for Fitting SEDs within Prospector

Parameter	Description	Prior
$\log(M/M_\odot)$	Total stellar mass formed	See Section 3.1
z	Redshift	See Section 3.2
SFH	Ratio of SFRs in adjacent log-spaced time bins	See Section 3.3
$\log(Z^\star/Z_\odot)$	Stellar metallicity	Gaussian approximating the M–$Z^\star$ relationship of Gallazzi et al. (2005)
n	Power-law index for a Calzetti et al. (2000) attenuation curve	Uniform: min $= -1.0$, max $= 0.4$
$\hat{\tau}_{dust,2}$	Optical depth of diffuse dust (Charlot & Fall 2000)	Truncated normal: min $= 0$, max $= 4$, $\mu = 0.3$, $\sigma = 1$
$\hat{\tau}_{dust,1}/\hat{\tau}_{dust,2}$	Ratio between the optical depths of birth cloud dust and diffuse dust (Charlot & Fall 2000)	Truncated normal: min $= 0$, max $= 2$, $\mu = 1$, $\sigma = 0.3$
$\log f_{AGN}$	Ratio between the object's AGN luminosity and its bolometric luminosity	Uniform: min $= -5$, max $= \log 3$
$\log \tau_{AGN}$	Optical depth of AGN torus dust	Uniform: min $= \log 5$, max $= \log 150$
$\log(Z_{gas}/Z_\odot)$	Gas-phase metallicity	Uniform: min $= -2.0$, max $= 0.5$
q_{PAH}	Fraction of grain mass in PAHs (Draine & Li 2007)	Truncated normal: min $= 0$, max $= 7$, $\mu = 2$, $\sigma = 2$

error (MBE) as

$$\delta_{MBE} = \frac{1}{n}\sum_{i=1}^{n} \Delta\theta. \qquad (6)$$

5.1. Assessing Parameter Recovery

We start by evaluating the accuracy of redshift inference, then proceed to key stellar populations metrics. Results on redshifts inferred from the mock catalog assuming uniform priors, and the proposed priors of this Letter are shown in the upper panels of Figure 3. Redshifts estimated from EAzY are overplotted as a reference point. The widths of the central distributions in parameter inference of Prospector-β and EAzY exhibit no significant differences, as evident from their NMAD values. It is reassuring to see that Prospector-β and EAzY share similar performance, although one discrepancy may be noticed. There exists a small cluster around $\Delta z/(1+z) \sim 0.7$ in the EAzY redshifts. This is probably a manifestation of the survey geometry of the mock catalog, which is mostly comprised of young galaxies at $z > 1$ with a rising ultraviolet/optical slope due to nebular emission and young stars. EAzY tends to mistake these for a rising near-infrared slope of old/quiescent galaxies at $z = 0.1$–0.3. Prospector recovers the correct redshift by correctly putting star-forming galaxies at high z, since its fiducial model already installs a hard limit that no stars can be older than the age of the universe at each time step. On the other hand, EAzY just fits linear combinations of templates. Its default template set is optimized for wider/shallower surveys in which more old quiescent galaxies are present. Adapting a new set of z-dependent templates in EAzY is likely to alleviate this problem (e.g., the recent set sfhz, which will be described in G. Brammer et al. 2023, in preparation). Prospector-β performs understandably better than EAzY in this special survey geometry because it is informed by priors based on galaxy evolution that low-mass, high-z galaxies are more common and highly star forming.

Similar trends are observed in the lower panels of Figure 3, where we test our model on data from a JWST DD program. This perhaps serves as an even more convincing piece of evidence that our approach can perform as least as well as EAzY in the context of JWST surveys, with the added bonus of joint constraints on key stellar populations metrics.

Figure 4 demonstrates that the priors proposed in this Letter result in a more accurate recovery of mass, mass-weighted age, and SFR. This is straightforward to conclude from the smaller absolute MBEs because our model is informed by our expectation from galaxy evolution. Reducing the scattering requires more informative data; it is thus natural that we see no significant changes in NMADs. The recovery of the rest of the parameters in Table 1 shows no material difference between difference cases, which is reasonable given that we place no additional prior on them. Discussions on how well these parameters can be constrained within the Prospector framework can be found in, e.g., Lower et al. (2020), Tacchella et al. (2022b), and Lower et al. (2022).

Losing the HST coverage spanning 0.4–1 μm in observed frame does not deteriorate our model performance, beside a 0.1 dex increase in MBE of the redshift recovery. This is because the HST data tend to be associated with larger uncertainties, which means that the likelihood is mostly affected by the JWST bands.

5.2. Age–Mass–Redshift Degeneracy

What is surprising is that the different choices of Prospector-α and Prospector-β priors lead to comparable redshift recovery despite the aforementioned improvements in recovered age, mass, and SFR. Upon examining cases where the uniform priors recover true redshifts, one example of which is shown as Figure 5(a), we believe that the flat priors work by exploiting an age–mass–redshift degeneracy: an overestimation in age leads to a galaxy having a higher mass-to-light ratio; then mass must be overestimated to produce the same luminosity; however, redshift stays the same since color is mostly unaffected. This argument is supported by Figure 5(b). The degeneracy is so strong that the flat priors can mostly get the redshift right despite the wrong mass or age. We further discuss the significance of this finding in the following section.

6. Discussion and Conclusions

We aim to improve the SED fitting performed on less-informative data by including well-motivated priors based on observations of galaxy evolution. This Letter presents the construction, behavior, and influence of three joint priors on mass, galaxy number density, and dynamic SFH. The priors are incorporated into the Prospector inference framework. As a
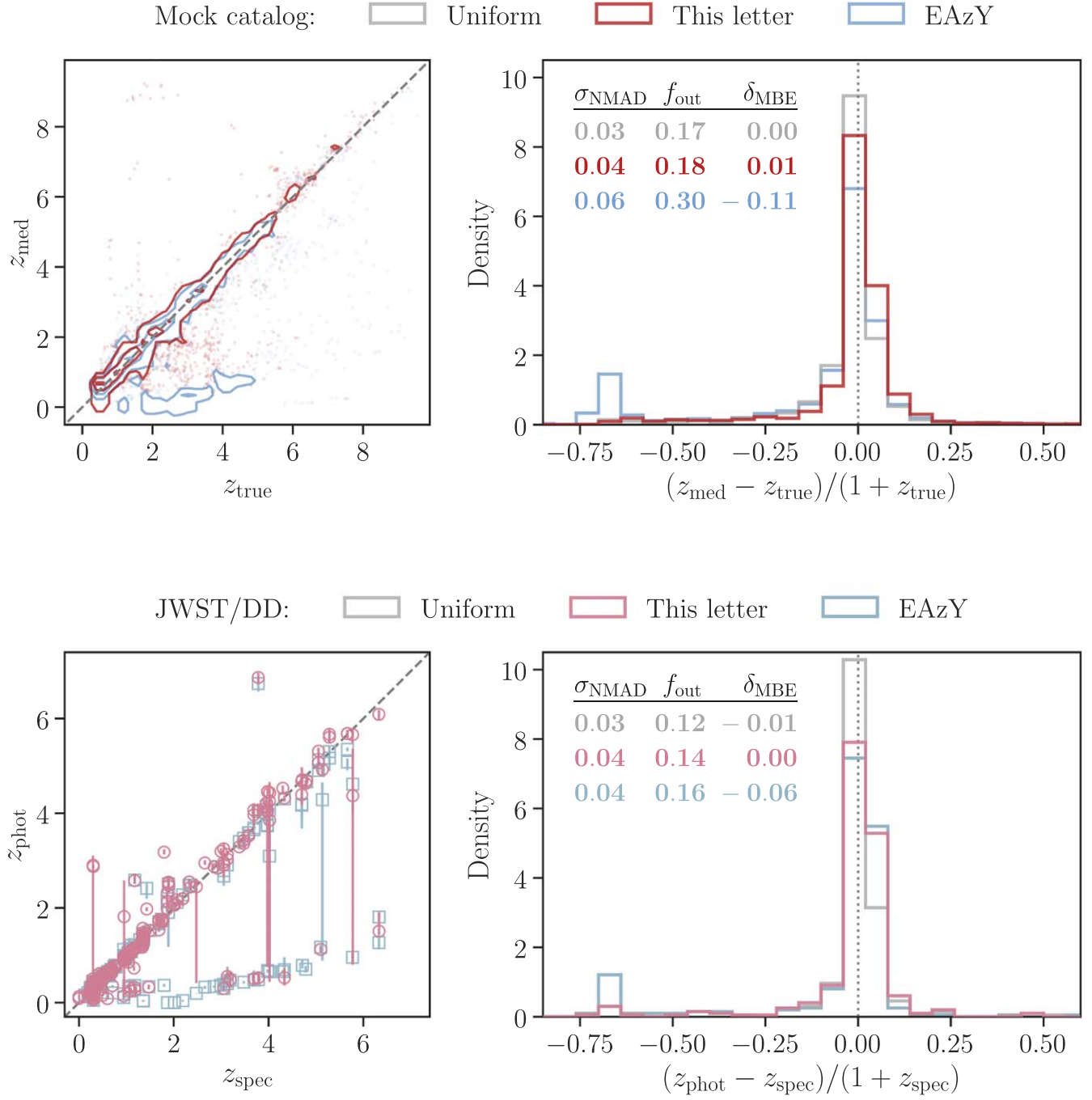
Figure 3. We compare the photo-z accuracy between `Prospector-β` and `EAzY` using both simulated and observed data. (Upper panel) On the left, we show true z vs. posterior median z_{med} from a mock catalog. Red contours show results from `Prospector-β`, while blue contours show these from `EAzY`. On the right, we show residuals as histograms. We also include results based on uniform priors, plotted as gray histograms. NMAD, outlier fraction, and MBE are reported in the same colors as the corresponding histogram. (Lower panel) Same as the upper panel, but for redshifts inferred from a spec-z sample with newly acquired data from a JWST DD program. In both cases, we find that `Prospector-β` performs at least as accurately as `EAzY`.

result, we are able to better recover stellar mass, age, and SFR. Below we discuss the key findings in this work.

First, we identify and characterize the age–mass–redshift degeneracy. This is likely a general feature of galaxy SEDs (e.g., Nagaraj et al. 2021), and in this analysis, it manifests as a deceivingly good recovery of redshift using the fiducial `Prospector` model that assumes uniform priors.

Second, our `Prospector-β` model breaks the above age–mass–redshift degeneracy. This has important ramifications.

Estimating masses and SFHs at high z is crucial to many studies such as those establishing the onset of galaxy formation and the reionization history. Breaking this degeneracy means that we now have a framework in which redshift, mass, SFH, and other contributions coming from stars, AGNs, gas, and dust can be accurately and self-consistently derived. An immediate gain is that we can use joint distributions on all parameters from `Prospector` to fit populations models that include full redshift uncertainties. For instance, it becomes
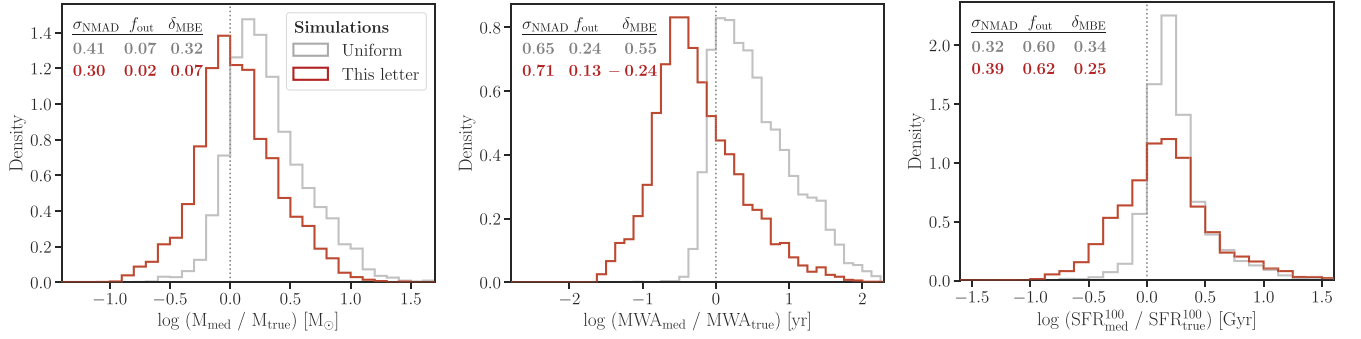
Figure 4. Here we demonstrate the performance of Prospector-β in recovering physical parameters, and contrast it with the fiducial Prospector-α model that assumes uniform priors. Residuals in mass, mass-weighted age, and SFR(0–100Myr) are shown as histograms. Results based on uniform priors are shown in gray, while those based on Prospector-β are shown in red. NMAD, outlier fraction, and MBE are reported in the same colors as the corresponding histogram. To avoid counting the error in redshift twice, we only include residuals for the subset where $(z_{\mathrm{med}} - z_{\mathrm{true}})/(1 + z_{\mathrm{true}}) \leqslant 0.1$. We find that Prospector-β results in less bias consistently in mass, age, and SFR as compared to uniform prior cases.
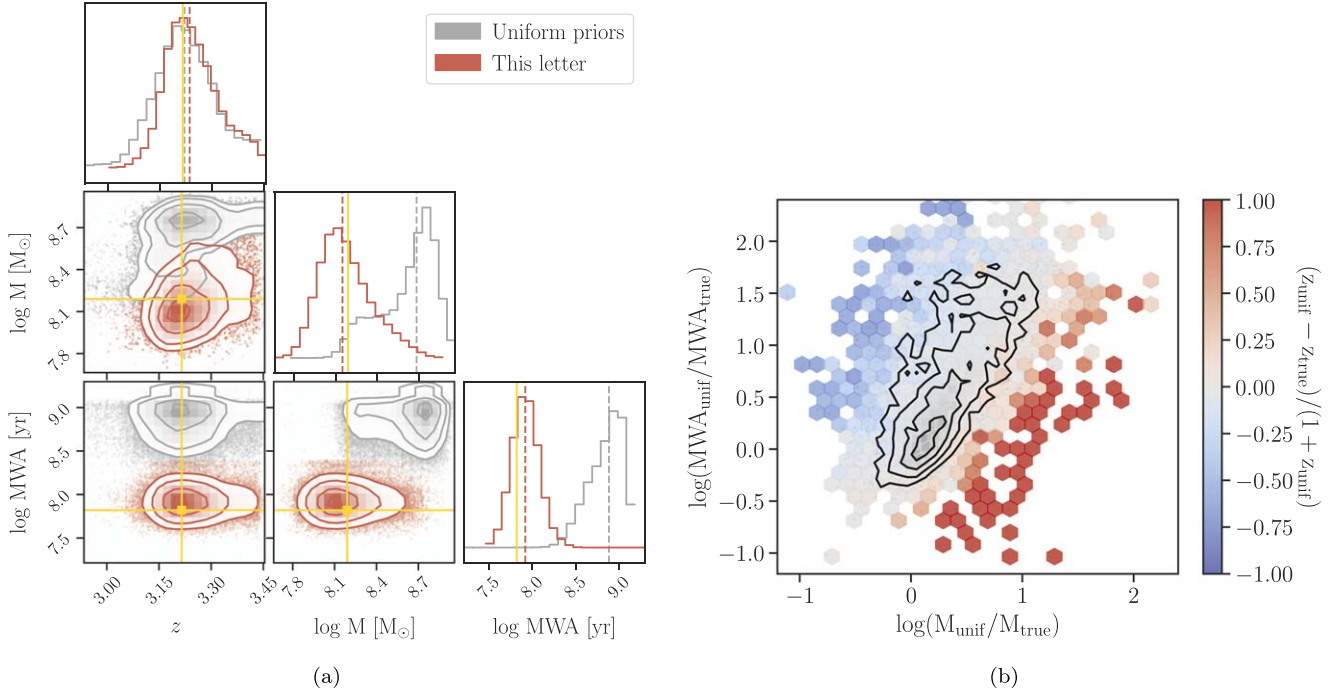
Figure 5. (a) Posteriors of a sample fit illustrating the degeneracy. The red contours show results assuming the Prospector-β priors, while the gray contours show those assuming uniform priors. Truths are overplotted on the corner plot in yellow. (b) The 2D histogram shows the average redshift offset when assuming uniform priors. Overplotted are contours showing the density of points. The fact that the contour lines are right along the line of the age–mass–redshift degeneracy suggests that uniform priors get the correct redshifts by exploiting this degeneracy.

possible to improve the mass function methodology from Leja et al. (2020) via the propagation of non-Gaussian redshift uncertainties.

Third, we compare Prospector to EAzY redshifts. Leistedt et al. (2023) also infers photo-z using Prospector-α and compares to EAzY; in this work, they find that a Bayesian hierarchical model is required to achieve similar performance in estimating redshift as EAzY. The fact that the fiducial Prospector model already shows comparable results in Figure 4(a) unambiguously demonstrates the influence of the age–mass–redshift degeneracy. Furthermore, the advantage of encoding galaxy evolution into the inference process is also made visible by the more accurate redshifts from Prospector-β. The less-satisfying performance of EAzY is likely due to the fact that we deploy the vanilla EAzY templates on simulations from a deep but small field, in which massive galaxies are rare. The default EAzY setting may become

insufficient since it is optimized for a more traditional setup where old quiescent galaxies constitute a significant proportion. Including redshift dependence into the templates will likely alleviate this problem (G. Brammer et al. 2023, in preparation). We do not carry out a comprehensive comparison to EAzY in further detail as the intention here is merely to have a well-tested code to act as a point of reference. We additionally emphasize that the better redshift recovery of Prospector-β demonstrates another advantage of our approach: modifying one's model based on a particular survey renders unnecessary once galaxy evolution is encoded in the inference process.

Fourth, we would like to highlight the simplicity of our SFH prior, which is motivated by downsizing in galaxy formation, and creates rising histories early in the universe and falling histories late in the universe. Even though the known behavior of SFH(M, z) is fairly complex, we deliberately choose a simple parameterization. There is no particular reason to tune this in

detail to the JAGUAR mock universe, which itself may not be a faithful representation of the true universe. Empirical constraints on SFH at high z remain weak, so we think that such simple approach is preferred, allowing data to inform the inference process to the greatest extent. It is feasible to develop a full population model for $0 < z < 1$ galaxies (e.g., Alsing et al. 2023), and constrain this hyperparameter with data. We leave this improvement for future studies.

Finally, a word of caution: our tests are performed on a mock catalog built for a deep JWST survey with a small area, which means that our mock does not test massive galaxies well since those are rare in this survey. We hope to calibrate the proposed model soon on real data, and also in broader surveys.

To conclude, high-z surveys in the era of JWST increasingly require us to infer more physical information from less-informative data. Applying `Prospector-`β to early JWST observations will downweight solutions that are disfavored by current observational constraints on galaxy formation, and hence increase our confidence in the results of our fitting (e.g., Labbe et al. 2022; Nelson et al. 2022).

This work is thus a timely addition to address the many challenges being put forth by JWST surveys. In the near future, we plan to apply `Prospector-`β to UNCOVER—an ultradeep Cycle 1 JWST survey targeting first-light galaxies (Bezanson et al. 2022). The data will be an even more stringent test on our model, and will provide a path for future improvements.

Appendix

We supply further details on the properties of our mock catalog in this Appendix. The simulated photometry consists of seven HST and seven JWST bands spanning ∼0.4–5 μm in the observed frame listed in Table 2, in which we also summarize the assumed depths in each filter. We also show the magnitude in F444W as a function of redshift in Figure 6.

The noise for the mock observations in the JWST bands are generated using the JWST Exposure Time Calculator as follows. First, the uncertainty within a $0\rlap{.}{''}16$ radius aperture is calculated given the planned total exposure time in each filter. A flat, faint f_ν source and the point-spread function (PSF) enclosed energy fraction curve is used to convert the flux from $e^- s^{-1}$ to nJy. This per-filter aperture uncertainty is then

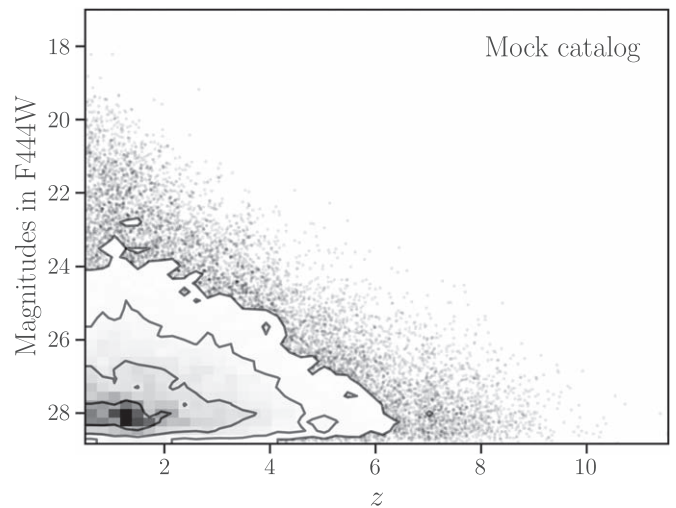**Figure 6.** Magnitudes in F444W are plotted as a function of redshifts for the the mock galaxies that we test our model on.

Table 2
Depth per Filter in the Mock Catalog

Filter	5σ Depth [AB mag]	Instrument
F435W	26.49	HST ACS (WFC)
F606W	26.62	HST ACS (WFC)
F814W	27.16	HST ACS (WFC)
F105W	25.51	HST WFC3 (IR)
F115W	26.48	JWST NIRCam
F125W	26.16	HST WFC3 (IR)
F140W	25.77	HST WFC3 (IR)
F150W	28.35	JWST NIRCam
F160W	29.02	HST WFC3 (IR)
F200W	28.79	JWST NIRCam
F277W	28.62	JWST NIRCam
F356W	28.62	JWST NIRCam
F410M	28.23	JWST NIRCam
F444W	28.65	JWST NIRCam

converted to a total flux uncertainty for each object using a per-object filter aperture correction that is determined using 2D Sérsic profiles convolved with the PSF. This total flux uncertainty for each band is then used directly as the mock filter flux uncertainty, δf_X. Noise is randomly drawn assuming a normal distribution with a standard deviation equal to δf_X, and added to the intrinsic fluxes to yield the mock flux f_X.

As for the noise in the HST bands, we similarly assume Gaussian distributions. The standard deviations are estimated from available HST observations overlapping the JWST DD program.

ORCID iDs

Bingjie Wang (王冰洁) ⓘ https://orcid.org/0000-0001-9269-5046
Joel Leja ⓘ https://orcid.org/0000-0001-6755-1315
Rachel Bezanson ⓘ https://orcid.org/0000-0001-5063-8254
Benjamin D. Johnson ⓘ https://orcid.org/0000-0002-9280-7594
Gourav Khullar ⓘ https://orcid.org/0000-0002-3475-7648
Ivo Labbé ⓘ https://orcid.org/0000-0002-2057-5376

Sedona H. Price ● https://orcid.org/0000-0002-0108-4176
John R. Weaver ● https://orcid.org/0000-0003-1614-196X
Katherine E. Whitaker ● https://orcid.org/0000-0001-7160-3632

References

Alsing, J., Peiris, H., Leja, J., et al. 2020, ApJS, 249, 5
Alsing, J., Peiris, H., Mortlock, D., Leja, J., & Leistedt, B. 2023, ApJS, 264, 29
Arnouts, S., Cristiani, S., Moscardini, L., et al. 1999, MNRAS, 310, 540
Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, AJ, 156, 123
Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33
Behroozi, P., Wechsler, R. H., Hearin, A. P., & Conroy, C. 2019, MNRAS, 488, 3143
Bell, E. F., & de Jong, R. S. 2001, ApJ, 550, 212
Benítez, N. 2000, ApJ, 536, 571
Bennett, C. L., Larson, D., Weiland, J. L., et al. 2013, ApJS, 208, 20
Bezanson, R., Labbe, I., Whitaker, K. E., et al. 2022, arXiv:2212.04026
Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2015, ApJ, 803, 34
Brammer, G. B., van Dokkum, P. G., & Coppi, P. 2008, ApJ, 686, 1503
Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000
Byler, N., Dalcanton, J. J., Conroy, C., & Johnson, B. D. 2017, ApJ, 840, 44
Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, ApJ, 533, 682
Carnall, A. C., Leja, J., Johnson, B. D., et al. 2019, ApJ, 873, 44
Carnall, A. C., McLure, R. J., Dunlop, J. S., & Dave, R. 2018, MNRAS, 480, 4379
Chabrier, G. 2003, PASP, 115, 763
Charlot, S., & Fall, S. M. 2000, ApJ, 539, 718
Chevallard, J., & Charlot, S. 2016, MNRAS, 462, 1415
Choi, J., Dotter, A., Conroy, C., et al. 2016, ApJ, 823, 102
Conroy, C., Dutton, A. A., Graves, G. J., Mendel, J. T., & van Dokkum, P. G. 2013, ApJL, 776, L26
Conroy, C., & Gunn, J. E. 2010, ApJ, 712, 833
Cowie, L. L., Songaila, A., Hu, E. M., & Cohen, J. G. 1996, AJ, 112, 839
Dotter, A. 2016, ApJS, 222, 8
Draine, B. T., & Li, A. 2007, ApJ, 657, 810
Feldmann, R., Carollo, C. M., Porciani, C., et al. 2006, MNRAS, 372, 565
Finkelstein, S. L., Bagley, M. B., Ferguson, H. C., et al. 2022a, arXiv:2211.05792
Finkelstein, S. L., Bagley, M. B., Haro, P. A., et al. 2022b, ApJL, 940, L55
Foreman-Mackey, D. 2016, JOSS, 1, 24
Furlanetto, S. R., & Mirocha, J. 2022, arXiv:2208.12828

Gallazzi, A., Charlot, S., Brinchmann, J., White, S. D. M., & Tremonti, C. A. 2005, MNRAS, 362, 41
Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, Natur, 585, 357
Hunter, J. D. 2007, CSE, 9, 90
Ilbert, O., Arnouts, S., McCracken, H. J., et al. 2006, A&A, 457, 841
Johnson, B. D., Leja, J., Conroy, C., & Speagle, J. S. 2021, ApJS, 254, 22
Kodra, D., Andrews, B. H., Newman, J. A., et al. 2023, ApJ, 942, 36
Labbe, I., van Dokkum, P., Nelson, E., et al. 2022, arXiv:2207.12446
Leistedt, B., Alsing, J., Peiris, H., Mortlock, D., & Leja, J. 2023, ApJS, 264, 23
Leitner, S. N. 2012, ApJ, 745, 149
Leja, J., Carnall, A. C., Johnson, B. D., Conroy, C., & Speagle, J. S. 2019, ApJ, 876, 3
Leja, J., Johnson, B. D., Conroy, C., Dokkum, P. G. v., & Byler, N. 2017, ApJ, 837, 170
Leja, J., Speagle, J. S., Johnson, B. D., et al. 2020, ApJ, 893, 111
Leja, J., Speagle, J. S., Ting, Y.-S., et al. 2022, ApJ, 936, 165
Li, Y., & Leja, J. 2022, ApJ, 940, 88
Lower, S., Narayanan, D., Leja, J., et al. 2020, ApJ, 904, 33
Lower, S., Narayanan, D., Leja, J., et al. 2022, ApJ, 931, 14
Mobasher, B., Dahlen, T., Ferguson, H. C., et al. 2015, ApJ, 808, 101
Nagaraj, G., Ciardullo, R., Lawson, A., et al. 2021, ApJ, 912, 145
Nelson, E. J., Suess, K. A., Bezanson, R., et al. 2022, arXiv:2208.01630
Newman, J. A., & Gruen, D. 2022, ARA&A, 60, 363
Oesch, P. A., Bouwens, R. J., Illingworth, G. D., Labbe, I., & Stefanon, M. 2018, ApJ, 855, 105
Pacifici, C., Kassin, S. A., Weiner, B. J., et al. 2016, ApJ, 832, 79
Park, M., Belli, S., Conroy, C., et al. 2022, arXiv:2210.03747
Pforr, J., Maraston, C., & Tonini, C. 2012, MNRAS, 422, 3285
Pozzetti, L., Bolzonella, M., Zucca, E., et al. 2010, A&A, 523, A13
Richard, J., Claeyssens, A., Lagattuta, D., et al. 2021, A&A, 646, A83
Salvato, M., Ilbert, O., & Hoyle, B. 2019, NatAs, 3, 212
Speagle, J. S. 2020, MNRAS, 493, 3132
Tacchella, S., Bose, S., Conroy, C., Eisenstein, D. J., & Johnson, B. D. 2018, ApJ, 868, 92
Tacchella, S., Conroy, C., Faber, S. M., et al. 2022b, ApJ, 926, 134
Tacchella, S., Finkelstein, S. L., Bagley, M., et al. 2022a, ApJ, 927, 170
Thomas, D., Maraston, C., Bender, R., & de Oliveira, C. M. 2005, ApJ, 621, 673
Treu, T., Roberts-Borsani, G., Bradac, M., et al. 2022, ApJ, 935, 110
Treu, T., Schmidt, K. B., Brammer, G. B., et al. 2015, ApJ, 812, 114
Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, NatMe, 17, 261
Wang, B., Leja, J., Villar, A., & Speagle, J. S. 2022, arXiv:2211.03747
Weaver, J. R., Cutler, S. E., Pan, R., et al. 2023, arXiv:2301.02671
Whitler, L., Stark, D. P., Endsley, R., et al. 2023, MNRAS, 519, 5859
Williams, C. C., Curtis-Lake, E., Hainline, K. N., et al. 2018, ApJS, 236, 33
Windhorst, R. A., Cohen, S. H., Jansen, R. A., et al. 2023, AJ, 165, 13