Draft version September 20, 2021
Typeset using LaTeX **preprint** style in AASTeX63

The DESI PRObabilistic Value-Added Bright Galaxy Survey (PROVABGS) Mock Challenge

ChangHoon Hahn,[1] Gyubin Kwon, Rita Tojeiro, Malgorzata Siudek, and GQP WG

[1] *Department of Astrophysical Sciences, Princeton University, Peyton Hall, Princeton NJ 08544, USA*

ABSTRACT

The PRObabilistic Value-Added Bright Galaxy Survey (PROVABGS) will provide measurements of galaxy properties, such as stellar mass (M_*), star formation rate (SFR), stellar metallicity ($Z_{\rm MW}$), and stellar age ($t_{\rm age,MW}$), for >10 million galaxies of the DESI Bright Galaxy Survey. Full posterior distributions of the galaxy properties will be inferred using state-of-the-art Bayesian SED modeling of DESI spectroscopy and Legacy Surveys photometry. In this work, we present the SED model, the Bayesian inference framework, and all the methodology of PROVABGS. Furthermore, we construct realistic synthetic DESI spectra and photometry using galaxies in the L-Galaxies semi-analytic model, based on their star formation and chemical enrichment histories, and apply the PROVABGS SED modeling on the mock observations. Afterwards, we compare the galaxy properties that we infer to the true galaxy properties of the simulation using a hierarchical Bayesian framework to quantiy accuracy and precision. Overall, we accurately infer the true M_*, SFR, $Z_{\rm MW}$, and $t_{\rm age,MW}$ of the simulated galaxies. However, we find that priors on galaxy properties induced by the SED model have a significant impact on the posteriors. We characterize the priors in detail and quantify that they impose a SFR $> 10^{-1} M_\odot$/yr lower bound on SFR, a $\sim$0.3 dex bias on $\log Z_{\rm MW}$ for galaxies with low spectral signal-to-noise, and $t_{\rm age,MW} < 8\,$Gyr upper bound on stellar age. We also demonstrate that a joint analysis of spectrophotometry significantly improves the constraints on galaxy properties over photometry alone and is necessary to mitigate the impact of the priors. With the methodology presented and validated in this work, PROVABGS will maximize information extracted from DESI observation and provide a galaxy catalog that will extend current galaxy studies to new regimes and unlock new probabilistic analyses.

Keywords: cosmology: observations – galaxies: evolution – galaxies: statistics

1. INTRODUCTION

changhoon.hahn@princeton.edu

Large galaxy surveys have been transformational for our understanding of galaxy evolution. With surveys such as the Sloan Digital Sky Survey (SDSS; York et al. 2000), Galaxy and Mass Assembly survey (GAMA; Driver et al. 2011), and PRIsm MUlti-object Survey (PRIMUS; Coil et al. 2011), we have now established the global trends of galaxies in the local universe. For instance, population statistiscs, such as the stellar mass function (Li & White 2009; Marchesini et al. 2009; Moustakas et al. 2013) or quiescent fraction (Kauffmann et al. 2003; Blanton et al. 2003; Baldry et al. 2006; Taylor et al. 2009), and their evolution are now well understood. Many global scaling relations of galaxy propreties such as the mass-metallicity relation (Tremonti et al. 2004) or the "star formation sequence" (Noeske et al. 2007; Daddi et al. 2007; Salim et al. 2007) have also been firmly established by these observations. Despite their importance in building our current understanding, however, these empirical relations from existing observations are inadequate for shedding further light on how galaxies form and evolve.

More precise and accurate measurements of empirical relations can potentially reveal new trends among galaxies undetected by previous observations. New approaches that go beyond observed relations also show promise. Empirical prescriptions for physical proceses can be placed on top of N-body simulations that capture hierarchical structure formation in empirical models (*e.g.* UNIVERSEMACHINE Behroozi et al. 2019). The predictions of these models can be compared to the observed distributions of galaxy properties to derive insights into physical processes, such as the timescale of star formation quenching (Wetzel et al. 2013; Hahn et al. 2017; Tinker et al. 2017). Predicted distributions of galaxy properties of large-scale cosmological hydrodyanmical simulations can also be compared to observations (*e.g.* Genel et al. 2014; Somerville & Davé 2015; Davé et al. 2017; Trayford et al. 2017; Dickey et al. 2021; Donnari et al. 2021). Advances in machine learning techniques for accelerating and emulating hydrodynamical simulations (Villaescusa-Navarro et al. 2021) will enable such comparisons to explore a broad range of galaxy formation models. Soon we will be able to compare detailed galaxy formation models directly against observations and explore the parameter spaces of the models. All of these approaches require more statistically poweful galaxy samples with well controlled systematics and well understood selection function.

The Dark Energy Spectroscopic Instrument (DESI) marks the next stage in large galaxy surveys. Over the next five years, DESI will use its 5000 robotically-actuated fibers to provide redshifts of $\sim$30 million galaxies over $\sim$14,000 deg^2, a third of the sky (Collaboration et al. 2016a,b). The redshifts will be spectroscopically measured from optical spectra that spans the wavelength range $3600 < \lambda < 9800$Å with spectral resolutions $R = \lambda/\Delta\lambda = 2000 - 5000$. In addition, DESI targets will also have photometry from the Legacy Imaging Surveys Data Release 9 (LS; Dey et al. 2019), used for target selection. LS is a combination of three public projects (Dark Energy Camera Legacy Survey, Beijing-Arizona Sky Survey, and Mayall z-band Legacy Survey) that jointly imaged the DESI footprint in three optical bands (g, r, and z). It also includes photometry in the *Wide-field Infrared Survey Explorer* $W1$, $W2$, $W3$, and $W4$ infrared bands. The infrared photometry is derived from all imaging through year 4 of NEOWISE-Reactivation force-photometered in the unWISE maps at the locations of LS optical sources (Meisner et al. 2017b,a).

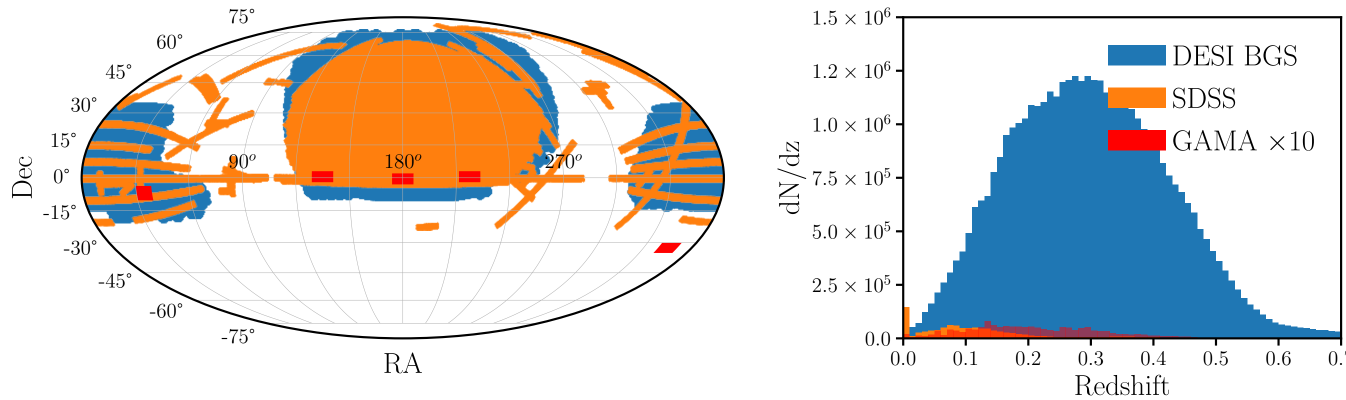

Figure 1. DESI will conduct the largest spectroscopic survey to date covering $\sim$14,000 deg^2. During dark time, DESI will measure >20 million spectra of luminous red galaxies, emission line galaxies, and quasars out to $z > 3$. During bright time, DESI will measure the spectra of $\sim$10 million galaxies out to $z\sim0.6$ with the Bright Galaxy Survey (BGS). *Left*: BGS (blue) will cover $\sim$2$\times$ the SDSS footprint (orange) and $\sim$45$\times$ the GAMA footprint (red). *Right*: We present the redshift distribution of BGS as predicted by the MXXL simulation (blue). We include the redshift distribution of SDSS (orange) and GAMA multiplied by 10$\times$ (red) for comparison. BGS will be roughly two orders of magnitude deeper than the SDSS main galaxy sample and deeper than GAMA. BGS will provide spectra for a magnitude limited sample of $\sim$10 million galaxies down to $r < 19.5$ (BGS Bright) and a deeper sample of $\sim$ 5 million galaxies as faint as $r < 20.175$ (BGS Faint).

During bright time, when the night sky is roughly $\sim$2.5$\times$ brighter than the nominal dark conditions, DESI will conduct the Bright Galaxy Survey (BGS). BGS will provide a $r < 19.5$ magnitude-limited sample of $\sim$10 million galaxies out to redshift $z < 0.6$ — the BGS Bright sample. It will also provide a surface brightness and color selected sample of $\sim$5 million faint galaxies with $19.5 < r < 20.175$. The completeness and effect of systematics of the BGS galaxy samples are characterized in detail in Hahn *et al.* (in prep.) using observations from DESI Survey Validation. Compared to the seminal SDSS main galaxy survey, BGS will provide optical spectra two magnitudes deeper, over twice the sky, and $\sim$4 billion years farther (Figure 1). It will observe a broader range of galaxies than previous surveys with unprecendented statistical power. Hence, BGS presents a unique opportunity to apply more sophisticated statistical analyses and new approaches to reveal the detailed connections among galaxy populations and advance galaxy evolution studies.

Observations alone, however, are not sufficient. New techniques and approaches require robust and consistent measuremnts of galaxy properties from the observations. The many advantages of the BGS observations would be squandered if they were analyzed inconsistently with a hodgepodge of methodologies. Value-added catalogs (VACs) that provide consistently measured galaxy properties for entire galaxy surveys are instrumental in this regard and have been used by hundreds of scientific analyses (see **?**, for a review). For SDSS galaxies, the NYU-VAGC (Blanton et al. 2005) provides photometric properties (*e.g.* absolute magnitudes) and the MPA-JHU catalog (Brinchmann et al. 2004)[1] provides spectral properties (*e.g.* emission line luminosities). Despite being released over a

[1] https://wwwmpa.mpa-garching.mpg.de/SDSS/DR7/

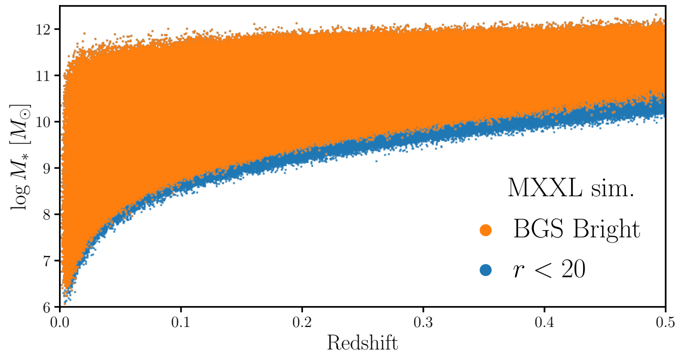

Figure 2. Stellar mass (M_*) distribution as a function of redshift of the $r < 19.5$ magnitude-limited BGS Bright sample (orange) as predicted by the MXXL simulation. We include the M_* distribution of MXXL galaxies with $r < 20$ (blue) for reference. Many such fainter galaxies will be included in the BGS Faint sample, which will observe galaxies as faint as $r < 20.175$. BGS will observe a wide range of galaxies with high completeness and provide galaxy samples with unprecedented statistical power. *We will apply state-of-the-art Bayesian SED modeling to all BGS galaxies to construct the PRObabilistic Value-Added BGS (PROVABGS) catalog, which will unlock new and more sophisticated approaches to galaxy evolution studies.*

decade ago, these VACs are still widely used today (*e.g.* Alpaslan & Tinker 2021; O'Donnell et al. 2021; Trevisan et al. 2021). Accompanying the BGS observations, the DESI Galaxy Quasar Physics (GQP) working group will produce the PRObabilistic Valued-Added Bright Galaxy Survey (PROVABGS) catalog.

For all >10 million BGS galaxies, PROVABGS will provide full posterior probability distributions of physical properties such as stellar mass (M_*), star formation rate (SFR), metallicity (Z), and stellar age ($t_{\rm age,MW}$). These properties will be inferred from both the LS photometry and DESI spectroscopy using a state-of-the-art Bayesian modeling of the galaxy spectral energy distribution (SED). PROVABGS will enable conventional analyses to be extended to a more statistically powerful spectroscopic galaxy sample. Population statistics such as the stellar mass function or the star formation sequence will be measured with higher precision than previously possible and for a much wider range of galaxies (Figure 2). The high completeness and simple selection function of the BGS Bright sample will also facilitate comparisons to empirical models or galaxy formation simulations with new approaches.

Moreover, PROVABGS will provide more accurate measurements of galaxy properties from full posterior distributions, rather than point estimates. Posteriors estimate the uncertainties on the galaxy property measurements and any degeneracies among them more accurately. This will enable a new level of statistical robustness in galaxy evolution studies. The PROVABGS posteriors will also open the door for Bayesian Hierarchical approaches. For instance, we can combine posteriors of galaxies to conduct principled population inference. Beyond mitigating biases, this will allow us to more fully exploit the observations since we can robustly include galaxies with less tightly constrained properties. Little explored low signal-to-noise regimes that can be probed with this approach may shed new light on galaxy evolution. Hierarchical inference can also improve the statistical power of

BGS through Bayesian shrinkage: the joint posterior of the galaxy sample can be used as the prior to shrink the uncertainties on the properties of individual galaxies. With these advantages, PROVABGS will provide a VAC that fully exploits the DESI observations and maximizes the scientific impact of BGS.

In this paper, we present the mock challenge for PROVABGS conducted by the GQP working group. We present the state-of-the-art SED modeling that will be used to infer the galaxy properties of BGS galaxies and construct the PROVABGS. We use an SED model with non-parametric prescriptions for galaxy star formation and metallicity histories and accelerate the parameter inference using neural emulators. Moreover, we validate our SED modeling on realistic mock BGS observations constructed using the L-Galaxies semi-analytic model (Henriques et al. 2015) and DESI survey simulations. By applying our SED model on mock observations, where we know the true galaxy properties, we demonstrate that we can accurately infer galaxy properties for PROVABGS and highlight the advantages of jointly analyzing photometric and spectroscopic observations. Furthermore, we characterize, in detail, the limitations of our SED modeling so that future studies using PROVABGS can use this work as a reference in interpreting their results.

In Section 2, we describe the L-Galaxies semi-analytic model and how we use them to construct mock BGS observations. We then present the SED model, our Bayesian parameter inference framework with neural emulators, and the mock challenge in Section 3. We present the results of the mock challenge in Section 4 and discuss their implications in Section 5.

2. SIMULATIONS

In this Section, we describe how we construct mock observations from simulated galaxies of the L-Galaxies semi-analytic galaxy formation model (SAM). We use a forward model that includes realistic noise, instrumental effects, and observational systematics and produces DESI-like photometry and spectra. Later, we apply Bayesian SED modeling to these mock observations and demonstrate that we can accurately infer the true galaxy properties.

2.1. *L-Galaxies*

L-Galaxies (hereafter LGal; Henriques et al. 2015) is a state-of-the-at semi-analytic galaxy formation model run on subhalo merger trees from the Millennium (Springel et al. 2005) and Millenium-II (Boylan-Kolchin et al. 2009) N-body simulations. Millenium-I and II provide a dynamic range of $10^{7.0} M_\odot < M_* < 10^{12} M_\odot$ and adopts a Collaboration et al. (2014) ΛCDM cosmology. LGal includes prescriptions for gas infall and cooling, star formation, disc and bulge formation, stellar and black hole feedback, and the environmental effects of tidal and ram-pressure stripping. AGN feedback, which prevents hot gas from cooling, is the major mechanism for quenching star formation in massive galaxies. LGal model parameters are calibrated against the observed stellar mass functions and passive (quiescent) fractions at four different redshifts from $z = 3$ to 0. For further detail on LGal, we refer readers to Henriques et al. (2015).

2.2. *Spectral Energy Distributions*

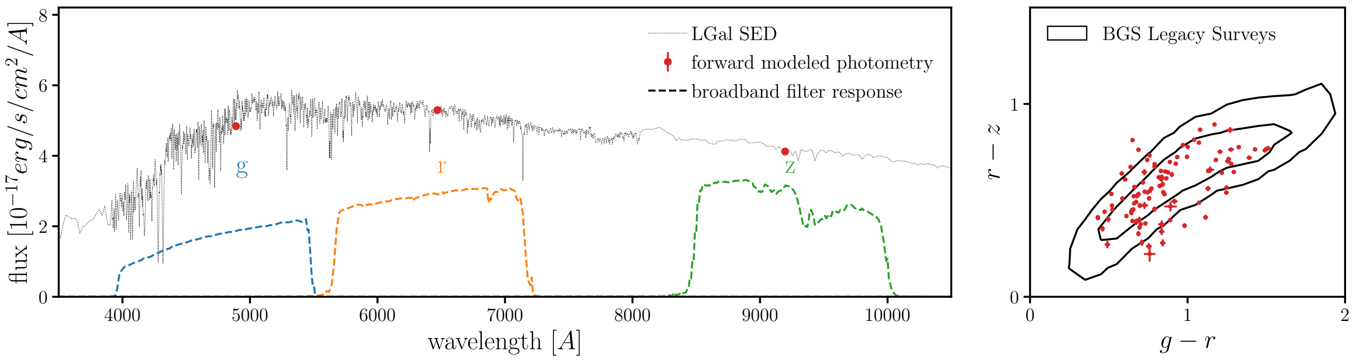

Figure 3. *Left*: We forward model DESI g, r, and z photometry (red) for our simulated galaxies (Section 2.1) by convolving their SEDs (dotted) with the broadband filters (dashed) and then applying an empirical noise model based on BGS objects in LS (Section 2.3). *Right*: The $g-r$ and $r-z$ color distribution of the forward modeled LGAL photometry is in good agreement with the color distribution of LS BGS objects (black contours).

For each simulated galaxy, LGAL provides the star formation histories (SFHs) and chemical enrichment histories (ZH) for its bulge and disk components in approximately log-spaced lookback time bins. We treat each lookback time bin, i, as a single stellar population (SSP) of age t_i. Then, we derive the luminosities of the bulge and disk components by summing up the luminosities of all of their SSPs:

$$L^{\text{comp.}}(\lambda) = \sum_i (\text{SFH}_i^{\text{comp.}} \Delta t_i)\ L_{\text{SSP}}(\lambda\,;\, t_i, Z_i^{\text{comp.}}). \tag{1}$$

$\text{SFH}_i^{\text{comp.}}$ and $Z_i^{\text{comp.}}$ are the star formation rate and metallicity of the bulge or disk component in lookback time bin i. Δt_i is the width of the bin. L_{SSP} corresponds to the luminosity of the SSP, which we calculate using the Flexible Stellar Population Synthesis (FSPS Conroy et al. 2009; Conroy & Gunn 2010) model. For FSPS, we use the MIST ischrones (Paxton et al. 2011, 2013, 2015; Choi et al. 2016; Dotter 2016), the MILES spectral library (Sánchez-Blázquez et al. 2006), and the Chabrier (2003) initial mass function (IMF).

Next, we apply velocity dispersions to $L^{\text{comp.}}(\lambda)$. For the disk, we apply a fixed 50 km/s velocity dispersion. For the bulge, we derive its velocity dispersion using the Zahid et al. (2016) empirical relation that depends on the total bulge mass. Afterwards, we apply dust attenuation to stellar emission in the disk component (L^{disk}) based on the cold gas content and orientation of the disk. We derive the attenuation curve using a mixed-screen model with the Mathis (1983) dust extinction curve. Stellar emission from stars younger than 30Myr are further attenuated with a uniform dust screen and a wavelength dependent optical depth. No dust attenuation is applied to the bulge component. We use the same dust attenuation as the dust prescription in Henriques et al. (2015) uses to construct galaxy colors.

Finally, we combine the attenuated disk component and the bulge component to construct the total luminosity of the simulated galaxy and then convert this rest-frame luminosity to observed-frame

SED flux using its redshift, z.

$$f_{\rm SED}(\lambda) = \frac{A(\lambda) L^{\rm disk}(\lambda) + L^{\rm bulge}(\lambda)}{4\pi d_L(z)^2 (1+z)}. \tag{2}$$

$A(\lambda)$ here is the dust attenuation for the disk component described above and $d_L(z)$ is the luminosity distance. In the left panel of Figure 3, we present an example of the SED flux constructed for an arbitrary LGal galaxy (black).

2.3. *Forward Modeling DESI Photometry*

In this section, we describe how we construct realistic LS-like photometry from the SEDs of simulated galaxies described in the last section. First, we convolve the SEDs with the broadband filters of the Legacy Survey to generate broadband photometric fluxes:

$$f_X = \int f_{\rm SED}(\lambda) R_X(\lambda) {\rm d}\lambda. \tag{3}$$

$f_{\rm SED}$ is the galaxy SED (Eq. 2) and R_X is the transmission for filter X. We generate photometry for the LS g, r, and z optical bands. Next, we apply realistic measurement uncertainties to the derived photometry by matching each simulated galaxy to a BGS target from LS DR9 with the nearest r-band magnitude and $g-r$ and $r-z$ color. The photometric uncertainties (σ_X) and r-band fiber flux ($f_r^{\rm fiber}$) of the BGS object are then assigned to the simulated galaxy. We apply photometric noise by sampling a Gaussian distribution with standard deviation σ_X:

$$\hat{f}_X = f_X + n_X \quad \text{where } n_X \sim \mathcal{N}(0, \sigma_X). \tag{4}$$

Finally, we impose the target selection criteria of BGS (Ruiz-Macias et al. 2021, Hahn *et al.* in prep.). In the left panel of Figure 3, we overplot the forward modeled photometry (red) on top of the SED flux (black) for an arbitrary LGalgalaxy. For reference, we also plot R_X for the g, r, and z bands of LS in blue, orange, and green, respectively. On the right panel, we compare the $g-r$ versus $r-z$ color distribution for the forward modeled LGal galaxies (red) to the color distribution of BGS objects in LS (black contour). The errorbars mark the photometric uncertainties. The forward modeled photometry show good agreement with LS BGS objects in color space.

2.4. *Forward Modeling DESI Spectra*

In this section, we describe how we construct realistic DESI-like spectroscopy from the SEDs of simulated galaxies. We forward model the fiber aperture effect and apply a noise model that accurately reproduces the bright time observations of BGS.

DESI uses fiber-fed spectrographs with fibers that have angular radii of 1”. Only the light from a galaxy within this fiber aperture is collected by the instrument. LS provides measurements of photometric fiber flux within a 1” radius aperture ($f_X^{\rm fiber}$), which estimates the flux that passes through to the fibers. When we assign photometric uncertainties to our simulated galaxies based on r, $g-r$, and $r-z$ in Section 2.3, we also assign r-band fiber flux. We model the SED flux that passes

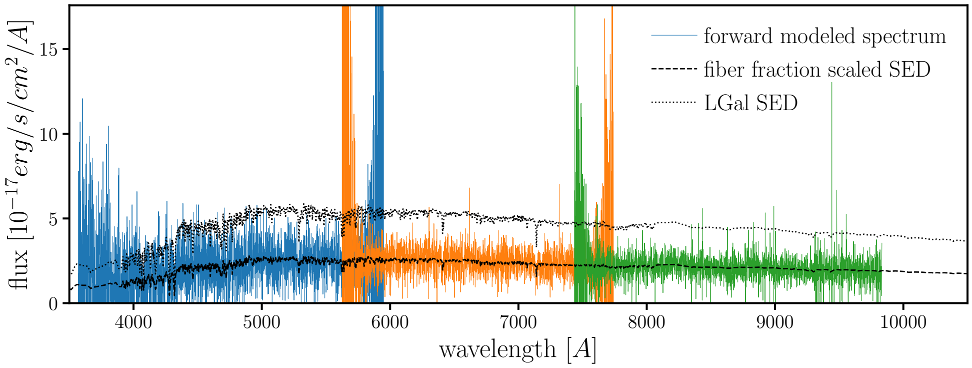

Figure 4. We construct simulated DESI spectra (solid) for LGAL simulated galaxies by applying a fiber aperture correction to the SED (dashed) and a realistic DESI noise model. We apply a fiber aperture correction by scaling down the full SED (dotted) by the r-band fiber fraction derived from the Legacy Surveys imaging. The noise model accounts for the DESI spectrograph response and the bright time observing conditions of BGS (**CH:** Hahn *et al.* in prep., Schlafly *et al.* in prep.). Our forward model produces DESI-like spectra for all three arms of the DESI spectrographs: b, r, and z (blue, orange, and green, respectively). For more details, we refer readers to Section 2.4.

through the fiber (fiber loss) by scaling the SED flux by the r band fiber fraction, the ratio of the r-band fiber flux over the total r band flux:

$$f^{\rm spec}(\lambda) = \left(\frac{f_r^{\rm fiber}}{f_r}\right) f_{\rm SED}(\lambda). \tag{5}$$

This fiber aperture correction assumes that there is no significant color dependence. We also assume that there are no significant biases in the fiber flux measurements in LS due to miscentering of objects. We discuss the implications of these assumptions later in Section 5. In addition to the aperture correction, we also use $f_r^{\rm fiber}$ to derive "measured" $\hat{f}_r^{\rm fiber}$:

$$\hat{f}_r^{\rm fiber} = f_r^{\rm fiber} + n_r^{\rm fiber} \quad \text{where } n_r^{\rm fiber} \sim \mathcal{N}\left(0, \frac{f_r^{\rm fiber}}{f_r}\sigma_r\right). \tag{6}$$

After all, when analyzing actual observations, we would not know the true fiber fraction. We later use $\hat{f}_r^{\rm fiber}$ to set the prior on the nuisance parameter of our SED modeling (Section 3).

Next, we apply a noise model that simulates the DESI instrument response and bright time observing conditions of BGS. We use the same noise model as the spectral simulations[2] used for the BGS survey design and validation (Hahn *et al.* in prep.). We refer readers to Schlafly *et al.* (in prep.) for details about the survey simulations. Specifically, we use nominal dark time observing conditions with 180s exposure time. These conditions accurately reproduce the spectral noise and redshift success rates of observed BGS spectra in DESI survey validation observations (Hahn *et al.* in prep.). In Figure 4, we present the forward modeled BGS spectrum of an arbitrary LGAL galaxy (solid). We mark the spectrum from each arm of the tree DESI spectrographs separately (blue, orange,

[2] https://specsim.readthedocs.io

green). For reference, we include the full SED (dotted) and fiber fraction scaled SED (dashed) of the galaxy.

3. JOINT SED MODELING OF PHOTOMETRY AND SPECTRA

3.1. *Stellar Population Synthesis Modeling*

PROVABGS will provide inferred galaxy properties derived from joint SED modeling of DESI photometry and spectra. For the SED modeling, we use a state-of-the-art stellar population synthesis (SPS) model that uses a non-parametric SFH with a star-burst, a non-parametric ZH that varies with time, and a flexible dust attenuation prescription.

The form of the SFH is one of the most important factors in the accuracy of an SPS model. In general, the form of the SFH requires balancing between being flexible enough to describe the wide range of SFHs in observations while not being too flexible that it can describe any SFH at the expense of constraining power. If the model SFH is not flexible enough to describe actual SFHs of galaxies, then unbiased galaxy properties cannot be inferred using the SPS model. For instance, most SPS models (*e.g.* CIGALE, Serra et al. 2011; BAGPIPES, Carnall et al. 2017) use parametric SFH such as the exponentially declining τ-model. Such functional forms, however, produce biased estimates of galaxy properties (*e.g.* M_* and SFR) when used to fit mock observations of simulated galaxies (Simha et al. 2014; Pacifici et al. 2015; Carnall et al. 2018). On the other hand, many non-parametric forms of the SFH are overly flexible and allow unphysical SFHs (Leja et al. 2019), which unncessarily increases parameter degeneracies and discards constraining power.

In our SPS model, we use a non-parametric SFH with two components: one based on non-negative matrix factorization (NMF) basis functions and a starburst component. For the first component, SFH is a linear combination of NMF SFH bases:

$$\mathrm{SFH}^{\mathrm{NMF}}(t, t_{\mathrm{age}}) = \sum_{i=1}^{4} \beta_i \frac{s_i^{\mathrm{SFH}}(t)}{\int\limits_0^{t_{\mathrm{age}}} s_i^{\mathrm{SFH}}(t)\,\mathrm{d}t}. \tag{7}$$

$\{s_i^{\mathrm{SFH}}\}$ are the NMF basis functions and $\{\beta_i\}$ are the coefficients. The integral in the denominator normalizes the NMF basis functions to unity. We constrain $\sum_i \beta_i = 1$, so the total SFH of the component over the age of the galaxy (t_{age}) is normalized to unity. $\{s_i^{\mathrm{SFH}}\}$ are derived from the IllustrisTNG cosmological hydrodynamic simulation (Nelson et al. 2018; Pillepich et al. 2018; Springel et al. 2018). The SFHs of simulated galaxies IllustrisTNG are compiled, rebinnined, and smoothed. Afterwards, we perform non-negative matrix factorization (Lee & Seung 1999; Cichocki & Phan 2009; Févotte & Idier 2011) on the smooth SFHs to derive $\{s_i^{\mathrm{SFH}}\}$. We find that 4 components is sufficient to accurately reconstruct the SFHs from IllustrisTNG. We present the NMF SFH bases as a function of lookback time in left panel of Figure 5. For further details on the derivation of the NMF bases, we refer readers to Appendix A. Assuming that the SFHs of IllustrisTNG galaxies resemble the SFHs of actual observed galaxies, our NMF form provides a compact and flexible representation of the SFHs.

The NMF basis functions are derived from smooth SFHs, which means that it does not include any stochasticity. However, observations and high resolution zoom-in hydrodyanmical simulations

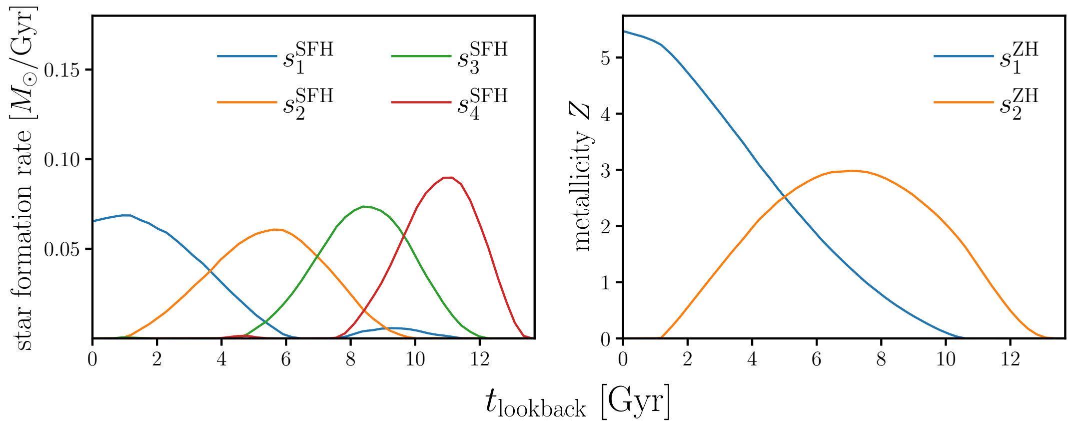

Figure 5. Non-negative matrix factorization basis functions for the SFH (left) and ZH (right) used in the non-parametric SFH and ZH prescriptions of our SPS model. These basis functions are derived from the SFHs and ZHs of simulated galaxies in the IllustrisTNG cosmological hydrodynamic simulations. With the NMF basis functions, we can reproduce the wide range of SFHs and ZHs of IllustrisTNG galaxies.

both find significant stochasticity in galaxy SFHs (Sparre et al. 2017; Caplar & Tacchella 2019; Hahn et al. 2019; Iyer et al. 2020). To include this stochasticity in our SPS model, we include a starburst component that consists of a SSP in the SFH. Thus, for the total SFH, we use

$$\mathrm{SFH}(t, t_{\mathrm{age}}) = (1 - f_{\mathrm{burst}})\ \mathrm{SFH}^{\mathrm{NMF}}(t, t_{\mathrm{age}}) + f_{\mathrm{burst}}\ \delta_{\mathrm{D}}(t - t_{\mathrm{burst}}). \tag{8}$$

f_{burst} is the fraction of total stellar mass formed during the starburst; t_{burst} is the time at which the starburst occurs; δ_{D} is the Dirac delta function. In total we use 6 free parameters in our SFH: 4 NMF basis coefficients (β_i), f_{burst}, and t_{burst}.

Another key part of an SPS model is the ZH, or chemical enrichment history. Current SPS models mostly assume a ZH that does not vary over time (Carnall et al. 2017; Leja et al. 2019). Since galaxies do not have constant metallicities throughout their history, this assumption can significantly bias the inferred galaxy properties. Instead, for our ZH, we take a similar approach to the SFH and use NMF basis functions:

$$\mathrm{ZH}(t) = \sum_{i=1}^{2} \gamma_i s_i^{\mathrm{ZH}}(t). \tag{9}$$

$\{s_i^{\mathrm{ZH}}(t)\}$ are the ZH NMF basis functions and $\{\gamma_i\}$ are the coefficients. $\{s_i^{\mathrm{ZH}}(t)\}$ are fit using the ZHs of simulated galaxies from IllustrisTNG in the same fashion as the SFH (Appendix A). In the right panel of Figure 5, we present the ZH NMF bases as a function of lookback time. We use two NMF components, so our ZH prescription has 2 free parameters.

We use the SFH and ZH above to model the unattenuated rest-frame luminosity as a linear combination of multiple SSPs, evaluated at logarithmically-spaced lookback time bins. We use a fixed log-binning with the bin egdes starting with $(0, 10^{6.05}\mathrm{yr})$, $(10^{6.05}, 10^{6.15}\mathrm{yr})$, and continuing on with bins of width 0.1 dex. The binning is truncated at the age of the model galaxy. For a $z = 0$ galaxy, we use 43 t_{lookback} bins. We use log-spaced t_{lookback} bins because it better reproduces galaxy luminosities evaluated with much higher resolution $t_{\rm lookback}$ binning than linearly-spacing, for the same number of bins. At every $t_{\rm lookback}$ bin i, we evaluate the luminosity of a SSP with ZH(t_i), where t_i is the center of $t_{\rm lookback}$ bin, and total stellar mass calculated by resampling the SFH in Eq. 8. We use FSPS to evaluate the SSP luminosities and use the MIST ischrones, MILES spectral library, and the Chabrier (2003) IMF (same as in Section 2.2). Since we use MIST isochrones, we impose a minimum and maximum limit to ZH based on its coverage: 4.49×10^{-5} and 4.49×10^{-2}, respectively. We note that our stellar metallicity range is significantly broader than previous studies (*e.g.* Carnall et al. 2017; Leja et al. 2017; Tacchella et al. 2021) for additional flexibility.

Before we combine the SSP luminosities, we apply dust attenuation. We use a two component Charlot & Fall (2000) dust attenuation model with birth cloud (BC) and diffuse-dust (ISM) components. The BC component represents the extra dust attenuation of young stars that are embedded in modecular clouds and HII regions. For SSPs younger than $t_i < 100$Myr, we apply the following BC dust attenuation:

$$L_i(\lambda) = L_i^{\rm unatten.}(\lambda) \exp\left[-\tau_{\rm BC}\left(\frac{\lambda}{5500\text{Å}}\right)^{-0.7}\right]. \tag{10}$$

$\tau_{\rm BC}$ is the BC optical depth that determines the strength of the BC attenuation. Afterwards, *all* SSPs are attenuated by the diffuse dust using the Kriek & Conroy (2013) attenuation curve parameterization:

$$L_i(\lambda) = L_i^{\rm unatten.}(\lambda) \exp\left[-\tau_{\rm ISM}\left(\frac{\lambda}{5500\text{Å}}\right)^{n_{\rm dust}}\left(k_{\rm Cal}(\lambda) + D(\lambda)\right)\right]. \tag{11}$$

$\tau_{\rm ISM}$ is the diffuse dust optical depth. $n_{\rm dust}$ is the Calzetti (2001) dust index, which determines the slope of the attenuation curve. $k_{\rm Cal}(\lambda)$ is the Calzetti (2001) attenuation curve and $D(\lambda)$ is the UV dust bump, parameterized using a Lorentzian-like Drude profile:

$$D(\lambda) = \frac{E_b(\lambda\ \Delta\lambda)^2}{(\lambda^2 - \lambda_0^2)^2 + (\lambda\ \Delta\lambda)^2} \tag{12}$$

where $\lambda_0 = 2175$Å, $\Delta\lambda = 350$Å, and $E_b = 0.85 - 1.9 n_{\rm dust}$ are the central wavelength, full width at half maximum, and strength of the bump, respectively. Once dust attenuation is applied to the SSPs, we sum them up to get the rest-frame luminosity of the galaxy. In total, our SPS model has 12 free parameters: M_*, 4 SFH basis coefficients, $f_{\rm burst}$, $t_{\rm burst}$, 2 ZH basis coefficients, $\tau_{\rm BC}$, $\tau_{\rm ISM}$, and $n_{\rm dust}$.

In practice, evaluating each SSP using FSPS requires X seconds. For each model evaluation, we evaluate ~ 43 SSPs in each of the log-spaced $t_{\rm lookback}$ bins. Though this is not a prohibitive computational cost on its own, sampling a high dimensional parameter space for inference requires $> 100,000$ evaluations — *i.e.* > 100 CPU hours *per galaxy*. For the >10 million BGS galaxies, this would require *a billion* CPU hours. Instead, we use an emulator for the model luminosity, which uses a Principal Component Analysis (PCA) neural network (NN) following the approach of Alsing et al. (2019). TODO TODO TODO

Our emulator consists of a NN and PCA basis functions. The NN provides a flexible and accurate mapping between the SPS model parameters and PCA coefficients — *i.e.* the NN predicts PCA coefficients for a given set of SPS parameters. Then the linear combination of the predicted coefficients

Table 1. Parameters of the PROVABGS SPS model used for joint SED modeling of DESI photometry and spectroscopy.

name	description	prior
$\log M_*$	log galaxy stellar mass	uniform over [7, 12.5]
$\beta_1, \beta_2, \beta_3, \beta_4$	NMF basis coefficients for SFH	Dirichlet prior
$f_{\rm burst}$	fraction of total stellar mass formed in starburst event	uniform over [0, 1]
$t_{\rm burst}$	time of starburst event	uniform over [10Myr, 13.2Gyr]
γ_1, γ_2	NMF basis coefficients for ZH	log uniform over $[4.5 \times 10^{-5}, 1.5 \times 10^{-2}]$
$\tau_{\rm BC}$	Birth cloud optical depth	uniform over $[0, 3]$
$\tau_{\rm ISM}$	diffuse-dust optical depth	uniform over $[0, 3]$
$n_{\rm dust}$	Calzetti (2001) dust index	unifrom over$[-2, 1]$
$f_{\rm fiber}$	spectrum fiber-aperture effect normalization	Gaussian $\mathcal{N}(\hat{f}_r^{\rm fiber}, \frac{f_r^{\rm fiber}}{f_r}\sigma_r)$

and PCA basis functions give us the emulated model luminosity. The PCA basis functions and NN are trained using $1,000,000$ SPS parameters and model luminosity pairs, $(\theta, L(\lambda; \theta))$. Throughout the wavelength range relevant for BGS, $3000 < \lambda < 9800$Å, we achieve $< 1\%$ accurate with the emulator. For details on the training, validation, and performance of our PCA NN emulator, we refer readers to Kwon *et al.* (in prep.).

From the rest-frame luminosity, we obtain the observed-frame, redshifted, flux in the same way as Eq. 2. In our case, redshift is not a free parameter since we will have high quality spectroscopic redshifts for every DESI BGS galaxy. BGS redshifts will have small redshift error, $\sigma_z < 0.0005(1+z)$ (150km/s), and <5% catastrophic failures, $\Delta z/(1 + z) < 0.003$ (<1000 km/s). To model DESI photometry, we convolve the model flux with the LS broadband filters as in Eq. 3. To model DESI spectra, we first apply Gaussian velocity dispersion. In this work, we keep velocity dispersion fixed at 0km/s but in practice velocity disperion can be set as a free parameter. Then the broadened flux is resampled into the wavelength binning of the observed DESI spectra, which has spectral resolution $R = 2000 - 5000$ over $3600 < \lambda < 9800$Å. Since the DESI spectra does not necessarily include all the light of a galaxy, we include a nuisance parameter $f_{\rm fiber}$, a normalization factor on the spectra to account for fiber aperture effects. Finally, the SPS model photometry and spectrum can be directly compared to observations.

3.2. *Bayesian Parameter Inference*

Using the SPS model above, we perform Bayesian parameter inference to derive posterior probability distributions of the SPS parameters from photometry and spectroscopy. From Bayes rule, we write down the posterior as

$$p(\theta \,|\, \mathbf{X}) \propto p(\theta)\; p(\mathbf{X} \,|\, \theta) \tag{13}$$

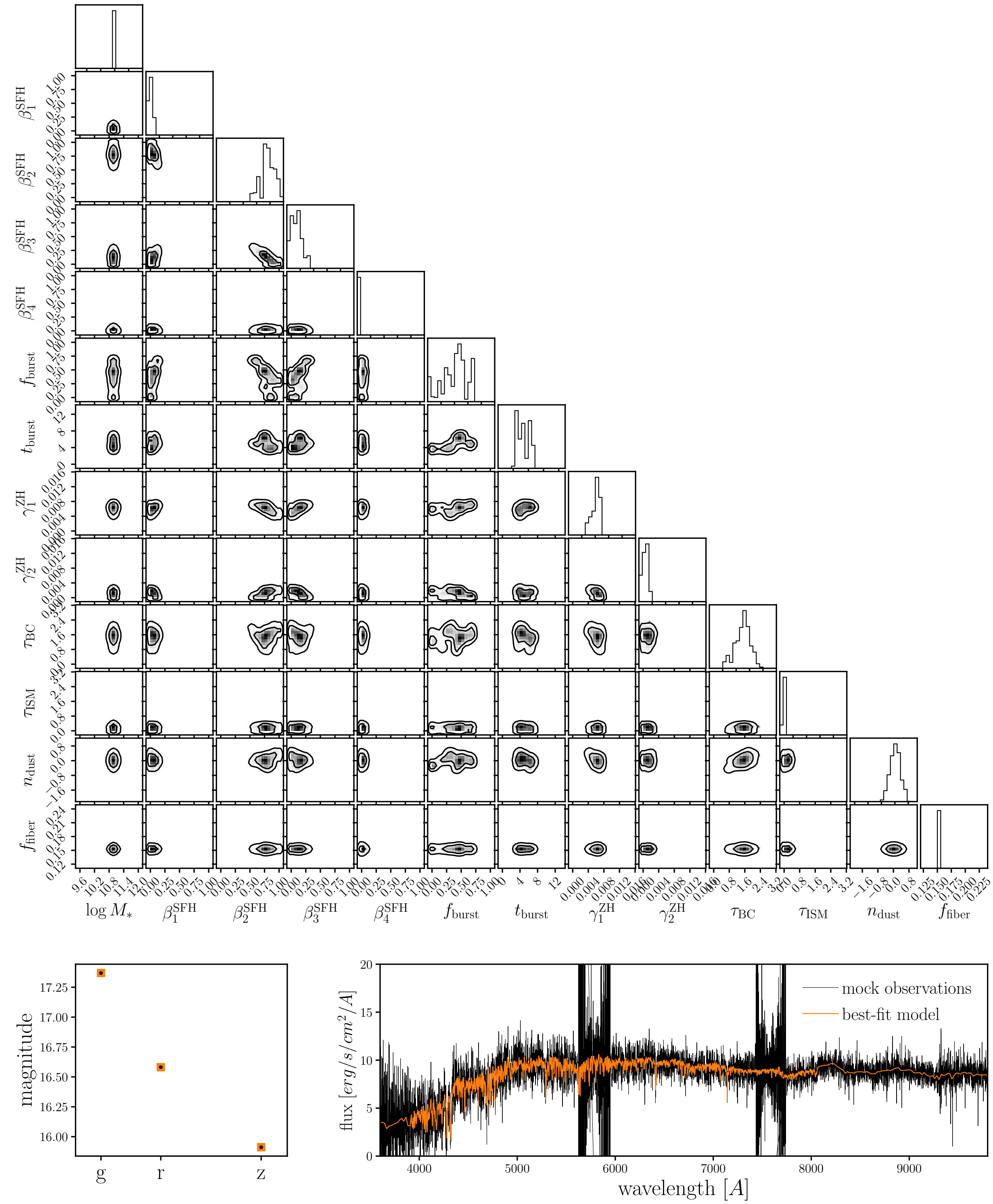

Figure 6. *Top*: Posterior probability distribution of our 12 SPS model parameters derived from joint SED modeling of the mock DESI photometry and spectrum. The contours mark the 68 and 95% percentiles. We use a Gaussian likelihood and the prior specified in Table 1 to evaluate the posterior and sample the distribution using ensemble slice MCMC. *With our Bayesian SED modeling approach, we capture the significiant parameter degeneracies and multimodality of the posterior distribution.*
Bottom: We compare the best-fit model observables (orange) to the mock observations (black). We find excellent agreement for both the LS photometry (left) and the DESI spectrum (right).

where $\mathbf{X}$ is the photometry or spectrum and θ is the set of SPS parameters. $p(\mathbf{X}\,|\,\theta)$ is the likelihood, which we calculate separately for the photometry

$$\mathcal{L}^{\rm photo} \propto \exp\left[-\frac{1}{2}\left(\frac{X^{\rm photo} - m^{\rm photo}(\theta)}{\sigma^{\rm photo}}\right)\right] \tag{14}$$

and for the spectrum

$$\mathcal{L}^{\rm spec} \propto \exp\left[-\frac{1}{2}\left(\frac{X^{\rm spec} - m^{\rm spec}(\theta)}{\sigma^{\rm spec}}\right)^2\right]. \tag{15}$$

$m^{\rm photo}$ and $m^{\rm spec}$ represent SPS model for photometry and spectroscopy. $\sigma^{\rm photo}$ and $\sigma^{\rm spec}$ respresent the uncertainties on the measured photometry and spectrum. We consider the photometry indepedent from the spectrum so we combine the likelihoods when jointly modeling the spectrophotometry:

$$\log \mathcal{L} \approx \log \mathcal{L}^{\rm photo} + \log \mathcal{L}^{\rm spec}. \tag{16}$$

$p(\theta)$ is the prior on the SPS parameters. For most of our parameters, we use uninformative uniform priors with conservatively chosen ranges that are listed in Table 1. However, for the priors of $\{\beta_1, \beta_2, \beta_3, \beta_4\}$, the NMF coefficients for the SFH, we use a Dirichlet distribution. With a Dirichlet distribution, β_i are within $0 < \beta_i < 1$ and satisfy the constraint $\sum_i \beta_i = 1$. This maintains the normalization of the SFH in Eq. 7.

Now that we can evaluate the posterior at given θ, we derive the posterior distributions using Markov Chain Monte Carlo (MCMC) sampling. We use the Karamanis & Beutler (2020) ensemble slice sampling MCMC algorithm with the ZEUS Python package[3]. Ensemble slice sampling is an extension of standard slice sampling that does not requires specifying the initial length scale or any further hand-tuning. It generally converges faster than other MCMC algorithms (*e.g.* Metropolis) and generates chains with significantly lower autocorrelation.

When we sample the posterior, we do not directly sample our 12 dimensional SPS parameter space because we impose a Dirichlet prior on the SFH NMF coefficients. Dirichlet distributions are difficult to directly sample so we instead use the Betancourt (2012) sampling method, which transforms an N dimensional Dirichlet distribution into an easier to sample $N-1$ dimensional space. Hence, we sample the posterior in the transformed 11 dimensional space. Given this dimensionality, we run our MCMC sampling with 30 walkers. Overall, we find that the sampling converges after 2,500 iterations with a 500 iteration burn in. Deriving the posterior distribution from a joint SED modeling of photometry and spectra, with the emulator, takes 0.5 CPU hours. In principle, since our emulator uses a PCA NN, we can further expedite our paremeter inference using more efficient sampling methods that exploit gradient information, such as Hamlitonian Monte Carlo. We will explore further speed ups to our SED modeling in future works. TODO

In Figure 6 we present the posterior distribution of our 12 SPS model parameters for an arbitrarily chosen LGAL mock observation. We mark the 68 and 95 percentiles of the distribution with the contours. The posterior distribution reveal there are significant degeneracies between SPS parameters

[3] https://zeus-mcmc.readthedocs.io/

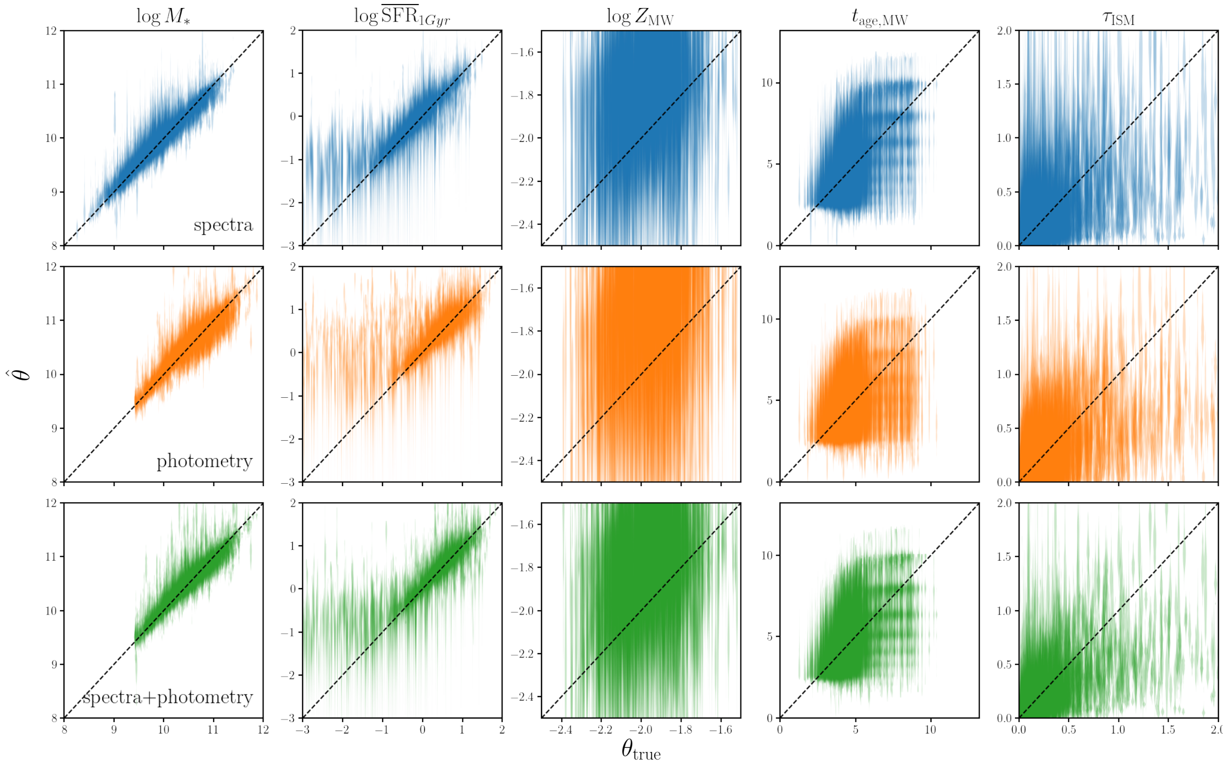

Figure 7. Comparison between the true galaxy properties, θ_{true}, and those inferred from SED modeling of mock observations, $\hat{\theta}$. From the left to right columns, we compare $\log M_*$, $\log \overline{\mathrm{SFR}}_{1\mathrm{Gyr}}$, $\log Z_{\mathrm{MW}}$, $t_{\mathrm{age,MW}}$ and τ_{ISM}. The inferred galaxy properties are derived from SED modeling of mock spectra (top), photometry (middle), and spectrophotometry (bottom). For each simulated galaxy, we represent the marginalized posterior of θ with a violin plot. *The posteriors demonstrate that, overall, we can derive accurate and precise constraints on certain galaxy properties from joint SED modeling of DESI photometry and spectra.*

— *e.g.* β_2^{SFH} and f_{burst}. Furthermore, the distribution is multimodal (see f_{burst} panels). With our Bayesian SED modeling, we are able to capture such complexities in the posterior that would be lost with point estimates or maximum likelihood approaches. In the bottom panels, we compare our SPS model evaluated at the best-fit parameters (orange) with the LGal mock observations (black). On the left, we compare the g, r, z band magnitudes; on the right, we compare the DESI spectroscopy. We find excellent agreement between the best-fit SPS model and mock observations.

4. RESULTS

The goal of this work is to demonstrate the precision and accuracy of inferred galaxy properties for PROVABGS. We apply our SED modeling to the mock observables of 2,123 LGal galaxies. From the posterior distributions of the SPS parameters, we derive the following physical galaxy properties: stellar mass (M_*), SFR averaged over 1 Gyr ($\overline{\mathrm{SFR}}_{1\mathrm{Gyr}}$), mass-weighted stellar metallicity (Z_{MW}), and diffuse-dust optical depth (τ_{ISM}). M_* and τ_{ISM} are both SPS model parameters, while $\overline{\mathrm{SFR}}_{1\mathrm{Gyr}}$ and

$Z_{\rm MW}$ are derived as

$$\overline{\rm SFR}_{\rm 1Gyr} = \frac{\int\limits_{t_{\rm age}-1{\rm Gyr}}^{t_{\rm age}} {\rm SFH}(t)\,{\rm d}t}{1{\rm Gyr}} \quad \text{and} \quad Z_{\rm MW} = \frac{\int\limits_{0}^{t_{\rm age}} {\rm SFH}(t)\,{\rm ZH}(t)\,{\rm d}t}{M_*}. \tag{17}$$

In Figure 7, we compare the galaxy properties inferred from SED modeling the mock observations, $\hat{\theta}$, to the true (input) galaxy properties, $\theta_{\rm true}$, of the simulated galaxies. In each column, we compare $\log M_*$, $\log \overline{\rm SFR}_{\rm 1Gyr}$, $\log Z_{\rm MW}$, $t_{\rm age,MW}$, and $\tau_{\rm ISM}$ from left to right. The inferred properties in the top, middle, and bottom rows are derived from SED modeling of spectra, photometry, and spectrophotometry, respectively. In each panel, we plot $\hat{\theta}$ using a violin plot, where the width of the marker represents the marginalized posterior distribution of θ. We note that in the comparison with SED modeling spectra only, we do not include $f_{\rm fiber}$ so the true stellar mass in this case corresponds to $f_{\rm fiber} \times M_*$. *Overall, the comparison demonstrates that we can robustly infer galaxy properties using the* PROVABGS *SED modeling.*

In more detail, we find that we infer unbiased and precise constraints on M_* throughout the entire M_* range. We also infer robust $\overline{\rm SFR}_{\rm 1Gyr}$ above $\log \overline{\rm SFR}_{\rm 1Gyr} > -1$ dex; below this limit, however, the inferred $\overline{\rm SFR}_{\rm 1Gyr}$ are significantly less precise and overestimate the true $\overline{\rm SFR}_{\rm 1Gyr}$. This bias at low $\overline{\rm SFR}_{\rm 1Gyr}$ is caused by model priors, which we discuss in further detail later in Section 5 and Appendix C. Both $Z_{\rm MW}$ and $t_{\rm age,MW}$ are not precisely constrained; however, we do not find clearn biases in the constraints. For $t_{\rm age,MW}$, the posteriors are less precise for galaxies with older stellar populations and they reveal the log-spaced $t_{\rm lookback}$ binning used in our SPS model for $t_{\rm age,MW} > 6$ Gyr. Lastly, $\tau_{\rm ISM}$ is overall accurately inferred for galaxies with low $\tau_{\rm ISM}$ but appears to be underestimated for for high $\tau_{\rm ISM}$.

The overall constraints on galaxy properties for the mock observations is especially encouraging due to the significant differences in the forward model used to generate the observations and the SPS model used in the SED modeling. First, the SFHs in the mock observations are taken directly from LGAL simulation outputs while the SFH parameterization in the SPS model is based on NMF bases fit to IllustrisTNG galaxy SFHs. Second, in the forward model, we construct the SED of the bulge and disk components of the simulated galaxies separately: the components have separate SFHs and ZHs. The SPS model treats all galaxies as having one component. Lastly, we use different dust prescriptions: Mathis (1983) dust attenuation curve in the forward model and Kriek & Conroy (2013) dust attenuation curve in the SPS model. Despite these significant differences, our constraints on certain galaxy properties are unbiased and precise.

Figure 7, also highlights the advantages of jointly modeling spectra and photometry. Comparing the constraints from spectrophotometry versus photometry alone, we find that including spectra significantly tightens the constraints for all properties. In addition, including spectra also appears to reduce biases of the constraints. For instance, with only photometry, we derive significantly more biased $\overline{\rm SFR}_{\rm 1Gyr}$ constraints. This is due to the limited constraining power of photometry, which allows the posteriors to be dominated by model priors. Adding spectra, significantly increases the contribution of the likelihood and ameliorates this effect.

Beyond qualitative comparisons of the posterior, we want to quantify the precision and accuracy of the inferred galaxy properties. Let $\Delta_{\theta,i}$ be the discrepancy between the inferred and true parameters for each galaxy: $\Delta_{\theta,i} = \hat{\theta}_i - \theta_i^{\rm true}$. Then, if we assume that $\Delta_{\theta,i}$ are sampled from a Gaussian distribution,

$$\Delta_{\theta,i} \sim \mathcal{N}(\mu_{\Delta_\theta}, \sigma_{\Delta_\theta}), \tag{18}$$

the mean (μ_{Δ_θ}) and standard deviation (σ_{Δ_θ}) of the distribution are population hyperparameters that represent the accuracy and precision of the inferred posteriors for the galaxy population. We can infer μ_{Δ_θ} and σ_{Δ_θ} using a hierarchical Bayesian framework (*e.g.* Hogg et al. 2010; Foreman-Mackey et al. 2014; Baronchelli et al. 2020).

Let $\{\boldsymbol{X}_i\}$ represent the photometry or spcetrum of a galaxy population and $\eta_\Delta = \{\mu_{\Delta_\theta}, \sigma_{\Delta_\theta}\}$ represent the population hyperparameters. Our goal is to constrain η_Δ from $\{\boldsymbol{X}_i\}$ — *i.e.* to infer $p(\eta_\Delta \,|\, \{\boldsymbol{X}_i\})$. We expand

$$p(\eta_\Delta \,|\, \{\boldsymbol{X}_i\}) = \frac{p(\eta_\Delta)\ p(\{\boldsymbol{X}_i\} \,|\, \eta_\Delta)}{p(\{\boldsymbol{X}_i\})} \tag{19}$$

$$= \frac{p(\eta_\Delta)}{p(\{\boldsymbol{X}_i\})} \int p(\{\boldsymbol{X}_i\} \,|\, \{\theta_i\})\ p(\{\theta_i\} \,|\, \eta_\Delta)\ {\rm d}\{\theta_i\}. \tag{20}$$

θ_i is the SPS parameters for galaxy i and $p(\{\boldsymbol{X}_i\} \,|\, \{\theta_i\})$ is likelihood of the set of observations $\{\boldsymbol{X}_i\}$ given the set of $\{\theta_i\}$. Since the likelihoods for each of the N galaxies, $p(\boldsymbol{X}_i \,|\, \theta_i)$, are not correlated, we can factorize and write the expression above as

$$= \frac{p(\eta_\Delta)}{p(\{\boldsymbol{X}_i\})} \prod_{i=1}^{N} \int p(\boldsymbol{X}_i \,|\, \theta_i)\ p(\theta_i \,|\, \eta_\Delta)\ {\rm d}\theta_i \tag{21}$$

$$= \frac{p(\eta_\Delta)}{p(\{\boldsymbol{X}_i\})} \prod_{i=1}^{N} \int \frac{p(\theta_i \,|\, \boldsymbol{X}_i)\ p(\boldsymbol{X}_i)}{p(\theta_i)}\ p(\theta_i \,|\, \eta_\Delta)\ {\rm d}\theta_i \tag{22}$$

$$= p(\eta_\Delta) \prod_{i=1}^{N} \int \frac{p(\theta_i \,|\, \boldsymbol{X}_i)\ p(\theta_i \,|\, \eta_\Delta)}{p(\theta_i)}\ {\rm d}\theta_i. \tag{23}$$

$p(\theta_i \,|\, \boldsymbol{X}_i)$ is the posterior for an individual galaxy, so the integral can be estimated using the Monte Carlo samples from the posterior:

$$\approx p(\eta_\Delta) \prod_{i=1}^{N} \frac{1}{S_i} \sum_{j=1}^{S_i} \frac{p(\theta_{i,j} \,|\, \eta_\Delta)}{p(\theta_{i,j})}. \tag{24}$$

S_i is the number of posterior samples and $\theta_{i,j}$ is the $j^{\rm th}$ sample of galaxy i. In practice, $p(\theta_{i,j} \,|\, \eta_\Delta) = p(\Delta_{\theta,i,j} \,|\, \eta_\Delta)$ is a Gaussian distribution and, hence, easy to evaluate. $p(\theta_{i,j}) = 1$ since we use uninformative and Dirichlet priors (Table 1). Finally, we derive the maximum a posteriori (MAP) value of η_Δ by maximizing the $p(\eta_\Delta \,|\, \{\boldsymbol{X}_i\})$ posterior distribution. This type of population inference is a major advantage of inferring full posteriors distributions of the galaxy properties. We discuss the derivation and interpretation of the hyperparameters in more detail in Appendix B.

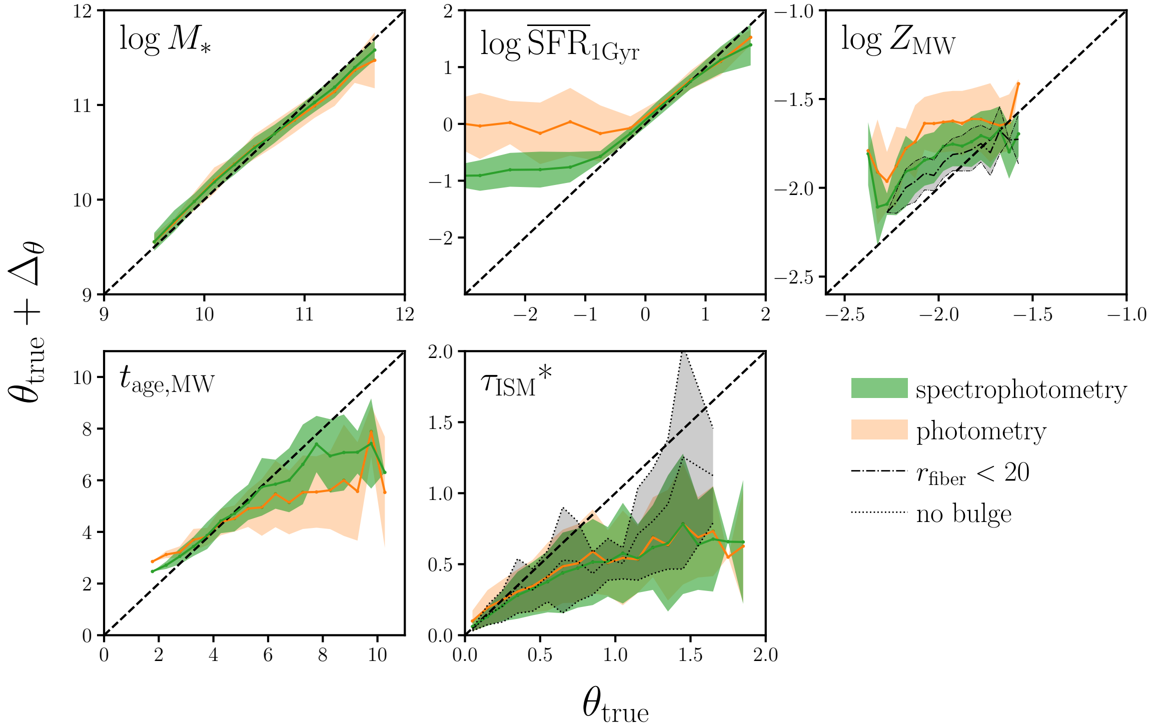

Figure 8. The accuracy and precision of galaxy property posteriors from our joint SED modeling of spectrophotometry, quantified using population hyperparameters $\eta_\Delta = \{\mu_{\Delta_\theta}, \sigma_{\Delta_\theta}\}$, as a function of true galaxy property (green). We plot $\theta_{\mathrm{true}} + \mu_{\Delta_\theta}$ in solid line and represent σ_{Δ_θ} with the shaded region. We include η_Δ for SED modeling of photometry alone (orange) for comparison. Including DESI spectra significantly improves both the accuracy and precision of the inferred galaxy properties. $\log \overline{\mathrm{SFR}}_{1\mathrm{Gyr}}$, $\log Z_{\mathrm{MW}}$, and $t_{\mathrm{age,MW}}$ constraints are significantly impacted by priors imposed by our SPS model (Appendix C). Discrepancies in the dust prescriptions between our SPS model and the mock observations drive the bias in τ_{ISM}. Nevertheless, *we accurately and precisely infer:* $\log M_*$ *for all* M_**,* $\log \overline{\mathrm{SFR}}_{1\mathrm{Gyr}}$ *above* $\log \overline{\mathrm{SFR}}_{1\mathrm{Gyr}} > -1\,\mathrm{dex}$*, and* $t_{\mathrm{age,MW}}$ *below* $8\,\mathrm{Gyr}$.

In Figure 8, we present the accuracy (μ_{Δ_θ}) and precision (σ_{Δ_θ}) of our joint SED modeling of spectra and photometry (green) as a function of true galaxy property. In each panel, we derive $p(\eta_\Delta \,|\, \{\boldsymbol{X}_i\})$ for $\log M_*$, $\overline{\mathrm{SFR}}_{1\mathrm{Gyr}}$, $\log Z_{\mathrm{MW}}$, $t_{\mathrm{age,MW}}$, and τ_{ISM} in bins of widths 0.2 dex, 0.5 dex, 0.05 dex, 0.5 Gyr, and 0.1, respectively. We only include bins with more than ten galaxies. μ_{Δ_θ} (solid) and σ_{Δ_θ} (shaded region) are the MAP values of $p(\mu_{\Delta_\theta}, \sigma_{\Delta_\theta} \,|\, \{\boldsymbol{X}_i\})$ posterior. For comparison, we include η_Δ for SED modeling of photometry alone (orange). We also include η_Δ for $\log Z_{\mathrm{MW}}$ of galaxies with $r_{\mathrm{fiber}} > 20$ (black dot-dashed) and η_Δ for τ_{ISM} of galaxies without bulges (black dotted), which we discuss later.

In Figure 9, we examine how the accuracy and precision of our galaxy parameter constraints are impacted by signal-to-noise ratio (SNR) or photometric color. We present η_Δ of our joint SED modeling of spectra and photometry as a function of r_{fiber}, r, $g-r$, and $r-z$. r_{fiber} and r magnitudes serve as proxies of the SNR for the spectra and photometry, respectively. In each row, we plot η_Δ for a different galaxy property: $\log M_*$, $\overline{\mathrm{SFR}}_{1\mathrm{Gyr}}$, $\log Z_{\mathrm{MW}}$, $t_{\mathrm{age,MW}}$ and τ_{ISM} (from top to bottom).

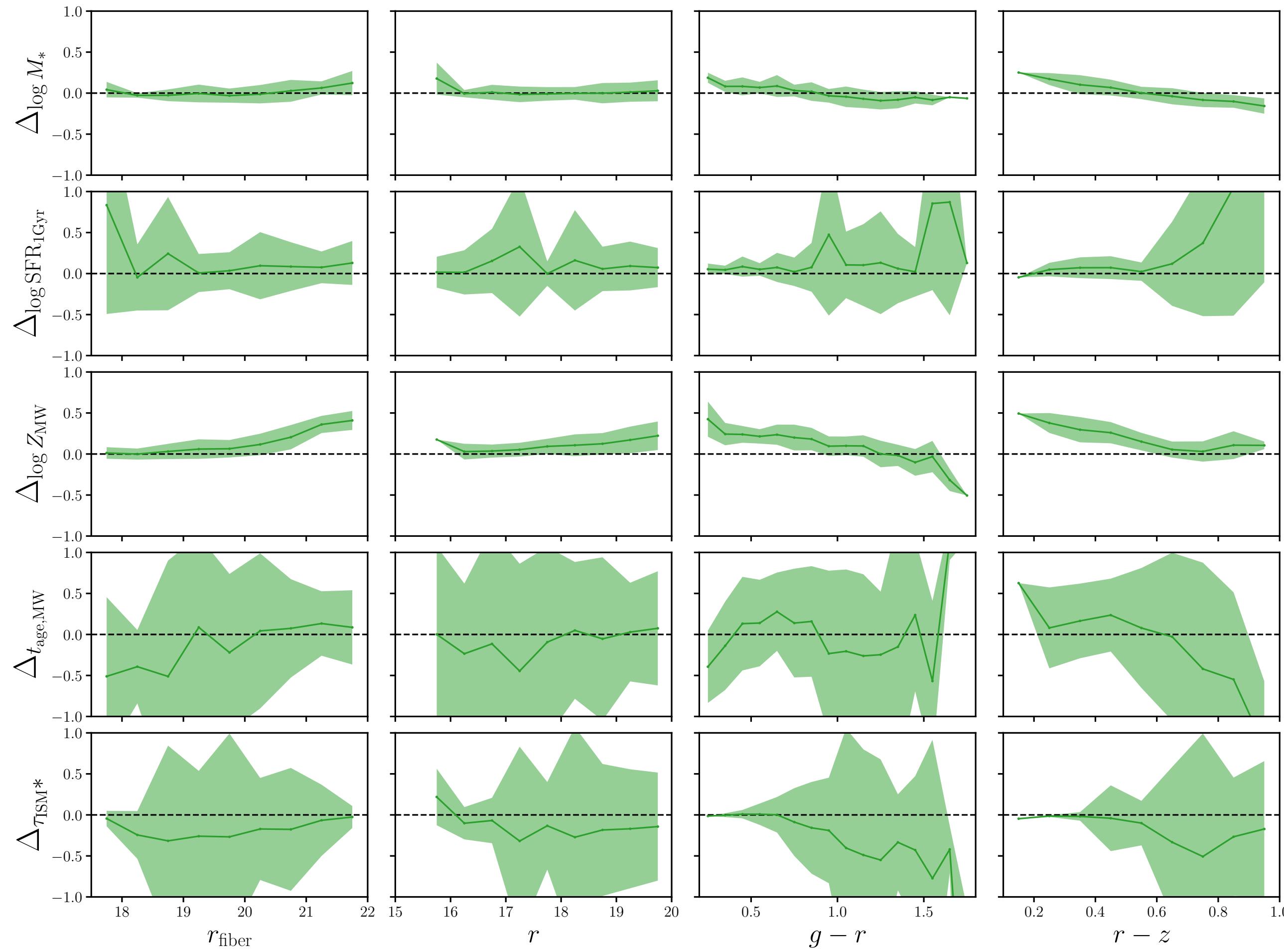

Figure 9. Accuracy and precision of the galaxy properties inferred from joint SED modeling of spectrophotometry as a function of r_{fiber}, r, $g-r$, and $r-z$. r_{fiber} and r magnitudes are proxies for spectral and photometric SNR. From the top to bottom rows, we present η_Δ for $\log M_*$, $\log \overline{\mathrm{SFR}}_{1\mathrm{Gyr}}$, $\log Z_{\mathrm{MW}}$, $t_{\mathrm{age,MW}}$ and τ_{ISM}. We find a significant dependence on spectral SNR in the inferred $\log Z_{\mathrm{MW}}$. When the spectral SNR is low ($r_{\mathrm{fiber}} > 20$), the prior on $\log Z_{\mathrm{MW}}$ imposed by our SPS model dominate the posterior and cause us to overestimate Z_{MW}. We find a significant color dependence on $\log \overline{\mathrm{SFR}}_{1\mathrm{Gyr}}$, $\log Z_{\mathrm{MW}}$, and $t_{\mathrm{age,MW}}$. For $\log Z_{\mathrm{MW}}$ and $t_{\mathrm{age,MW}}$, the dependence is driven by underlying correlations with spectral SNR and true $t_{\mathrm{age,MW}}$. Meanwhile, $\log \overline{\mathrm{SFR}}_{1\mathrm{Gyr}}$ is overestimated for the reddest galaxies with $r-z > 0.6$, which correspond to quiescent galaxies with $\log \overline{\mathrm{SFR}}_{1\mathrm{Gyr}} < -1$ dex. Otherwise we find no significant dependence on SNR or optical color.

Lastly, in Figure 10, we investigate whether there are any underlying dependence in the inferred galaxy properties on the M_*-SFR plane. In the top and bottom panels, we present μ_{Δ_θ} and σ_{Δ_θ} in $(\log M_*, \log \overline{\mathrm{SFR}}_{1\mathrm{Gyr}})$ bins for $\log M_*$, $\log \overline{\mathrm{SFR}}_{1\mathrm{Gyr}}$, $\log Z_{\mathrm{MW}}$, $t_{\mathrm{age,MW}}$ and τ_{ISM} (left to right). We use $\log M_*$ bins of width 0.225 dex and $\log \overline{\mathrm{SFR}}_{1\mathrm{Gyr}}$ bins of width 0.25 dex for $\log \overline{\mathrm{SFR}}_{1\mathrm{Gyr}} > 0$ dex and 0.5 dex for $\log \overline{\mathrm{SFR}}_{1\mathrm{Gyr}} < 0$ dex. We only present bins with more than 10 galaxies. On the $M_* - \mathrm{SFR}$ plane, we can examine whether the accuracy and precision of the inferred properties have significant

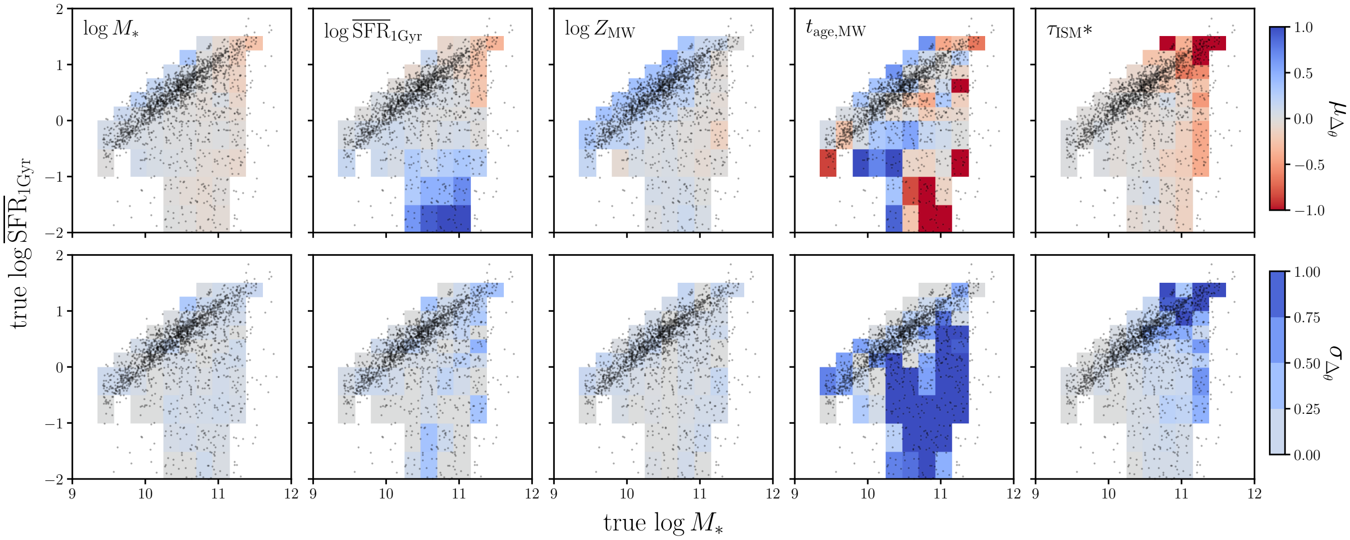

Figure 10. Accuracy and precision of the galaxy properties inferred from joint SED modeling of spectrophotometry as a function of the galaxies' true M_* and $\overline{\mathrm{SFR}}_{1\mathrm{Gyr}}$. We present μ_{Δ_θ} and σ_{Δ_θ} in (M_*, $\overline{\mathrm{SFR}}_{1\mathrm{Gyr}}$) bins for $\log M_*$, $\log \overline{\mathrm{SFR}}_{1\mathrm{Gyr}}$, $\log Z_{\mathrm{MW}}$, $t_{\mathrm{age,MW}}$ and τ_{ISM} in the top and bottom panels respectively. $\log M_*$ is accurately and precisely constrained for all types of galaxies. $\log \overline{\mathrm{SFR}}_{1\mathrm{Gyr}}$ is accurately and precisely constrained for all galaxies except for quiescent galaxies with $\log \overline{\mathrm{SFR}}_{1\mathrm{Gyr}} > -1$ dex. $\log Z_{\mathrm{MW}}$ is overestimated for star-forming galaxies, due to their overall lower spectral SNR. $t_{\mathrm{age,MW}}$ is accurately and precisely constrained for star-forming galaxies that have overall younger stellar populations. τ_{ISM} is accurately and precisely constrained for all galaxies except massive star-forming galaxies, which have high true τ_{ISM}.

dependencies for galaxy type.

Inferred $\log M_*$: Overall, we infer accurate and precise $\log M_*$ from the PROVABGS SED modeling. There is no significant dependence in μ_{Δ_θ} and σ_{Δ_θ} with true $\log M_*$ throughout the M_* range. This means that we accurately infer the true M_* throughout $\sim 10^9 - 10^{12} M_\odot$ with uniform precision of $\sigma_{\Delta_{\log M_*}} \sim 0.1$ dex. We also find no significant dependence on SNR dependence for M_* — neither r_{fiber} nor r magnitudes significantly affect $\mu_{\Delta_{\log M_*}}$ and $\sigma_{\Delta_{\log M_*}}$. There is a noticeable correlation with $g - r$ and $r - z$ color, which also appears in the $M_* - \mathrm{SFR}$ plane. However, this correlation is small compared to the precision of our inferred posterior on $\log M_*$. When we compare the η_Δ from spectrophotometry to η_Δ from photometry we find that including DESI spectra increases both the accuracy and precision of the constraints, especially at high $M_* > 10^{11} M_\odot$.

Inferred $\log \overline{\mathrm{SFR}}_{1\mathrm{Gyr}}$: We infer accurate $\log \overline{\mathrm{SFR}}_{1\mathrm{Gyr}}$ for galaxies with $\log \overline{\mathrm{SFR}}_{1\mathrm{Gyr}} > -1$ dex with $\sim$0.1 dex precision. In fact, we find a $\log \overline{\mathrm{SFR}}_{1\mathrm{Gyr}} \sim -1$ dex lower bound for the inferred $\log \overline{\mathrm{SFR}}_{1\mathrm{Gyr}}$. Below this limit, we significantly overestimate $\log \overline{\mathrm{SFR}}_{1\mathrm{Gyr}}$, consistent with the bias in Figure 7. At $\log \overline{\mathrm{SFR}}_{1\mathrm{Gyr}} < -1$ dex, the constraints are also significantly broader with $\sigma_{\Delta_{\log M_*}} \sim 0.25 - 0.3$ dex. Comparing μ_{Δ_θ} and σ_{Δ_θ} from spectrophotometry versus from only photometry, we confirm that including spectra significantly improves the accuracy and tightens the $\log \overline{\mathrm{SFR}}_{1\mathrm{Gyr}}$ constraints. For

$\overline{\mathrm{SFR}}_{\mathrm{1Gyr}}$ below $\log \overline{\mathrm{SFR}}_{\mathrm{1Gyr}} < -1$ dex, including spectra reduces the bias $\sim$1 dex — an order of magnitude.

We find no significant correlation between the accuracy and precision of $\overline{\mathrm{SFR}}_{\mathrm{1Gyr}}$ with spectral or photometric SNR. Meanwhile, there is a more significant color dependence where we overestimate $\log \overline{\mathrm{SFR}}_{\mathrm{1Gyr}}$ by $\mu_{\Delta_{\log \overline{\mathrm{SFR}}_{\mathrm{1Gyr}}}} > 0.5$ dex for the reddest galaxies ($g-r > 1.5$ and $r-z > 0.6$). The constraints for these galaxies are also significantly less precise: $\sigma_{\Delta_{\log \overline{\mathrm{SFR}}_{\mathrm{1Gyr}}}} \sim 0.5$ dex. The bias is also apparent in Figure 10: there is significant bias in the inferrd for quiescent galaxies where we overestimate $\overline{\mathrm{SFR}}_{\mathrm{1Gyr}}$. $\overline{\mathrm{SFR}}_{\mathrm{1Gyr}}$ is also slightly underestimated for the most massive ($M_* > 10^{11} M_\odot$) star-forming galaxies. These biases are consequences of our SPS model priors. $\overline{\mathrm{SFR}}_{\mathrm{1Gyr}}$ is a derived quantity; hence, the uniform priors we impose on SPS parameters induce non-uniform priors on them. Our SPS model imposes a prior on $\log \overline{\mathrm{SSFR}}_{\mathrm{1Gyr}}$ that is skewed and peaks at ~ -10.4 dex (Appendix C, Figure 12). Consequently, the posterior overestimates $\overline{\mathrm{SFR}}_{\mathrm{1Gyr}}$ at low $\overline{\mathrm{SFR}}_{\mathrm{1Gyr}}$ (red, quiescent galaxies) and underestimates $\overline{\mathrm{SFR}}_{\mathrm{1Gyr}}$ at the highest $\overline{\mathrm{SFR}}_{\mathrm{1Gyr}}$.

Inferred $\log Z_{\mathrm{MW}}$: Unlike in Figure 7, η_Δ in Figure 8 clearly reveal the accuracy and precision of the posteriors on $\log Z_{\mathrm{MW}}$. We find that μ_{Δ_θ} depends significantly on the true Z_{MW}: inferred $\log Z_{\mathrm{MW}}$ is overestimated by $\sim$0.2 dex below $\log Z_{\mathrm{MW}} < -2$ dex and slightly underestimated at the highest $\log Z_{\mathrm{MW}} > -1.6$ dex. $\sigma_{\Delta_\theta} \sim 0.15$ dex is uniform throughout the Z_{MW} range. Similar to $\overline{\mathrm{SFR}}_{\mathrm{1Gyr}}$, the bias in inferred Z_{MW} is a consequence of our SPS model priors. The prior skews $\log Z_{\mathrm{MW}}$ constraints towards the peak of the prior at $\log Z_{\mathrm{MW}} \sim -1.5$. Figure 8, also include η_Δ for posteriors derived from photometry alone (orange). This demonstrates that including DESI spectra substantially improves the accuracy of the $\log Z_{\mathrm{MW}}$ constraints — spectra reduces the overall bias on Z_{MW} by ~ 0.3 dex. The improvement comes from the likelihood contribution from DESI spectra reducing the relative contribution of the prior on the posterior.

In Figure 9, we find that the posteriors overestimate $\log Z_{\mathrm{MW}}$ at $r_{\mathrm{fiber}} > 20$. These correspond to mock observations with low spectral SNR and further underscore the constraining power of DESI spectra. Low spectral SNR means that the contribution of the likelihood from the spectra is reduced so the prior on $\log Z_{\mathrm{MW}}$ has a larger effect. The color dependence of μ_{Δ_θ} for Z_{MW} in Figure 9 is also a consequence of this spectral SNR dependence; so is the $M_* - \mathrm{SFR}$ dependence (Figure 10). If we exclude galaxies with low spectral SNR, both the color and $M_* - \mathrm{SFR}$ dependences are substantially reduced. More quantitatively, we present η_Δ for only $r_{\mathrm{fiber}} < 20$ galaxies in Figure 8 (black dot-dashed) where we infer $\log Z_{\mathrm{MW}}$ with $\mu_{\Delta_\theta} < 0.15$ dex and $\sigma_{\Delta_\theta} \sim 0.1$.

Inferred $t_{\mathrm{age,MW}}$: Figure 8 confirms that we derive unbiased and precise constraints on $t_{\mathrm{age,MW}}$ out to $t_{\mathrm{age,MW}} < 8$ Gyr. Below this limit, we infer $t_{\mathrm{age,MW}}$ with $\sigma_{\Delta_\theta}$$\sim$0.5 Gyr. For galaxies with older stellar populations above this limit, the log-spaced t_{lookback} binning in our SPS model (Section 3.1) expectedly underestimates $t_{\mathrm{age,MW}}$ constraints with larger uncertainties ($\sigma_{\Delta_{t_{\mathrm{age,MW}}}} \gtrsim 1$ Gyr). Meanwhile, we find no significant SNR or color dependence in Figure 9. At $r-z > 0.6$, $t_{\mathrm{age,MW}}$ is underestimated, but this is solely a consequence of the correlation between $r-z$ and true $t_{\mathrm{age,MW}}$. The simulated galaxies with $r-z > 0.6$ in our sample have overall older stellar populations. In Figure 10, we do not find a clear $M_* - \mathrm{SFR}$ dependence. However, $|\mu_{\Delta_{t_{\rm age,MW}}}|$ is larger and constraints are significantly less precise for galaxies with older stellar populations below the star-forming sequence.

Inferred $\tau_{\rm ISM}$: Lastly, we find that both the accuracy and precision of our $\tau_{\rm ISM}$ depend significant only the true $\tau_{\rm ISM}$ value. The inferred constraints increasingly underestimate $\tau_{\rm ISM}$ with lower precision for greater $\tau_{\rm ISM}$. The bias is due to discrepancies between the dust prescriptiosn of SPS model and the mock observations. First, we use a dust prescription with a different attenuation curve in the SPS model than in the forward model. This places a strict limit on how accurately we can derive $\tau_{\rm ISM}$. We intentially introduce this discrepancy since we do not know the "true" attenuation curve of observed galaxies in practice. Hence, with the discrepancy we test whether the PROVABGS SPS modeling can marginalize over the effect of dust.

Another reason for the biased $\tau_{\rm ISM}$ constraints is that we only attenuate the stellar emission in the disk component of the simulated galaxies and not the bulge component (Section 2.2). The true $\tau_{\rm ISM}$ is the optical depth for the disk componenet while our $\tau_{\rm ISM}$ constraints correspond to the optical depth of dust attenuation for the entire galaxies, a quantity that will be lower than the true $\tau_{\rm ISM}$ depending on how much the bulge contributes to the SED. Despite the discrepancies, we find no significant SNR or color dependence on the accuracy and precision of $\tau_{\rm ISM}$ constraints (Figure 9). Furthermore, we find unbiased and precise $\tau_{\rm ISM}$ constraints for all galaxies except star-forming galaxies above $M_* > 10^{11} M_\odot$ where we underestimate $\tau_{\rm ISM}$. Massive star-forming galaxies in this regime mainly have $\tau_{\rm ISM} > 1$. In Figure 8, we present a more apples-to-apples comparison of the $\tau_{\rm ISM}$ constraints, where we present η_Δ for only galaxies without bulge contributions (black dotted). For these galaxies, the bias in our $\tau_{\rm ISM}$ constraints is reduced and $\mu_{\Delta_\theta} < 0.5$ throughout the $\tau_{\rm ISM}$ range. Our constraints are still biased, however, due to the discrepant attenuation curves. We emphasize that the primary goal of dust prescription in our SPS model is to marginalize out the effect of dust. Based on the accuracy and precision of the constraints on other galaxy properties, the PROVABGS SPS model achieves this objective.

5. DISCUSSION

The most significant limitation of the PROVABGS SED modeling in inferring the true galaxy properties is the prior on galaxy properties imposed by the model. The effect of such priors is a major limitation for any SED modeling methods (*e.g.* Carnall et al. 2017; Leja et al. 2019). It is a consequence of the fact that galaxy properties are *not* parameters of the SPS model. For instance, $\overline{\rm SFR}_{\rm 1Gyr}$ and $Z_{\rm MW}$ are derived by integrating the SFH and ZH (Eq. 17), which are parameterized by $\beta_1, \beta_2, \beta_3, \beta_4$ and γ_1, γ_2. Uniform priors on βs and γs (Section 3.2 and Table 1) do not translate into uniform priors on $\overline{\rm SFR}_{\rm 1Gyr}$ and $Z_{\rm MW}$. Other galaxy properties (*e.g.* $t_{\rm age,MW}$, SFH, and ZH) likewise have non-uniform, and undesireable, priors.

One way to address this issue is to choose an SED model parameterization that does not impose extreme priors on galaxy properties and to characterize the priors in detail so that final posteriors can be appropriately interpreted. For the PROVABGS model, we explicitly chose our SFH prescription so that the prior on $\log \overline{\rm SSFR}_{\rm 1Gyr}$ extends the range -12 to -9 dex. Furthermore, we fully characterize the prior on $\overline{\rm SSFR}_{\rm 1Gyr}$, $Z_{\rm MW}$, $t_{\rm age,MW}$, SFH, and ZH in Appendix C (Figures 12). This way, we understand exactly how the model prior impacts the derived posteriors as we discuss in detail in Section 4. Beyond mitigating the effect of the priors, we can impose uniform prior (or any other desired prior distribution) on the derived galaxy properties by adjusting the priors on the SED model parameters. Handley & Millea (2019) recently demonstrated that maximum-entorpy priors can be used for this purpose to impose uniform priors on the inferred sum of neutrino masses in cosmological analyses. In an accompanying paper, Hahn (in prep.), I will demonstrate that maximum-entroy priors can also be used in Bayesian SED modeling to correct for the impact of priors on infer posteriors on derived galaxy properties.

In this work, we use forward modeled mock observations to demonstrate that we can infer accurate and precise posteriors on certain galaxy properties. The mock observations are constructed from LGal and include photometry and spectra. In generating the spectra, we model the fiber aperture effect — *i.e.* spectra only include light from a galaxy collected within its fiber diameter — by scaling the SED flux (Section 2.4). In our SED modeling, we account for this fiber aperture effect using a normalization factor, $f_{\rm fiber}$ (Section 3.1). Hence, our mock observations and SED modeling have a consistent treatment of the fiber aperture effect. In observations, however, aperture effects can be wavelength dependent (Gerssen et al. 2012; Richards et al. 2016), and if there is a strong dependence, an overall $f_{\rm fiber}$ factor would not be sufficient. In order to examine the wavelength dependence, we compare the ratio of the 1.5" aperture flux (fiber diamater) over total flux in g, r, and z bands of BGS targets from LS photometry. Overall, we find find no significant difference in the flux ratios of the different bands, which suggests that the fiber aperture effect in DESI does *not* have a strong wavelength dependence.

Flux calibration performed on DESI spectra can also induce wavelength dependent residuals. DESI spectra is measured using three-arm spectrographs that split the spectra into three b, r, and z channels with overlapping wavelength ranges: $3600-5930$, $5660-7720$, and $7470-9800$Å. After sky subtraction, flux calibration is performed on each channel of the spectra by matching physical stellar models to spectra of spectrohotometric standard stars observed in the same exposure. Since the calibration is performed for each channel seaparately, imperfections can imprint a wavelength dependent residual. **CH:** a statement on how well flux calibration works citing a KP1 paper. In a subsequent paper, Ramos *et al.* (in prep.), we examine the fiber aperture effect and wavelength dependent imprints on DESI spectra using BGS spectra from the DESI Survey Validataion data and MaNGA observations. Using galaxy properties derived using the PROVABGS pipeline for spectra from integrated field unit MaNGA observations, we will present aperture corrections that can be applied on derived BGS galaxy properties. We also note that the PROVABGS SED modeling pipline already includes flux calibration models beyond a single $f_{\rm fiber}$ and can easily be extended to include more sophisticated models (*e.g.* Chebyschev polynomial Carnall et al. 2017; Tacchella et al. 2021).

In both our PROVABGS SED model and mock observations, we use the MIST isochrones, MILES spectral library, and the Chabrier (2003) IMF. With the same set of choices, our analysis does not consider how different choices for stellar evolution or IMF can affect the inferred galaxy properties. Yet, it is well-established that there are major uncertainties in each of these choices (Conroy et al. 2009; Conroy 2013). For instance, recent observational works suggest that there may be significant variations in IMF (*e.g.* Treu et al. 2010; van Dokkum & Conroy 2010; Rosani et al. 2018; Sonnenfeld et al. 2019). Different SPS model choices can significantly impact the derived galaxy propeties. We reserve a detailed examination of this effect for future work. In the meantime, for the PROVABGS catalog we will release multiple catalogs each with different sets of choices for isochrone, spectral library, and IMF.

We demonstrate with the mock challenge that we can derive accurate and precise constraints on specific galaxy properties using the PROVABGS SED modeling. The PROVABGS catalog will have a number of key advantages over other value-added galaxy catalogs. First, PROVABGS will provide full Bayesian posteriors on galaxy properties instead of "best-fit" point estimates from maximizing the likelihood. Posterior distributions are essential for accurately estimating uncertainties on galaxy properties. As we find earlier, these uncertainties are significant, especially for properties such as $Z_{\rm MW}$ (Figure 7). Ignoring the uncertainties dramatically overestimates the statistical precision of the derived galaxy proprties and can significantly bias any galaxy study.

Furthermore, the PROVABGS posteriors will be derived from MCMC sampling rather than grid-based methods often used in the past (*e.g.* da Cunha et al. 2008; Moustakas et al. 2013; Boquien et al. 2019). As a result, they can accurately estimate posterior distributions with significant parameter degeneracies or multiple modes (peaks). For instance, in the posterior of Figure 6 we find degeneracies between $f_{\rm burst}$ and $\{\beta_1, \beta_2, \beta_3, \beta_4\}$ and between $\{\gamma_1, \gamma_2\}$ and $\{\beta_1, \beta_2, \beta_3, \beta_4\}$. The posterior is also multi-modal. Accurate estimates of the full posterior distribution are especially important, as they enable the maximum-entropy method, mentioned earlier, to correct for the significant impact of priors on derived galaxy properties. Grid-based methods also scale exponentially with the number of SPS parameters so they quickly become infeasible as the dimensionality of SPS models increase. Meanwhile, MCMC scales approximately linearly with the number of parameters.

In this work, we primarily focus on the following physical properties of galaxies: $\log M_*$, $\log \overline{\rm SFR}_{\rm 1Gyr}$, $\log Z_{\rm MW}$, $t_{\rm age,MW}$, and $\tau_{\rm ISM}$. The PROVABGS SPS model, however, can constrain galaxy properties beyond these properties. The SPS model employs nonparametric SFH and ZH prescriptions based on NMF bases and the model parameters include coefficients for these bases. Posteriors on the SPS model parameters can, thus, be used to derive constraints on the SFH and ZH. In Figure 11, we present the inferred SFH and ZH of two simulated galaxies from our LGAL sample: a star-forming galaxy (blue) and a quiescent galaxy (orange). We mark the 68 and 95% confidence intervals in the shaded regions. For comparison, we include the true SFH and ZH from LGAL (dashed). The inferred SFH and ZH is able to generally recover the true histories. We emphasize that current SPS models assume constant ZHs that does not vary over time (Carnall et al. 2017; Leja et al. 2019). Hence inferring ZH over time is a key advantage of the PROVABGS SPS model. Similar to the inferred $\overline{\rm SFR}_{\rm 1Gyr}$ and $Z_{\rm MW}$, the SFH and ZH constraints are also impacted by the priors imposed by our SPS model (Appendix C, Figure **??**).

Another key advantage of PROVABGS is that it will infer galaxy properties from joint SED modeling of photometry *and spectra*. Our results illustrate the advantages of including spectra in SED modeling. Galaxy spectra provide substantial statistical power for constraining galaxy properties. In addition to tightening constraints overall, the statistical power of galaxy spectra is essential

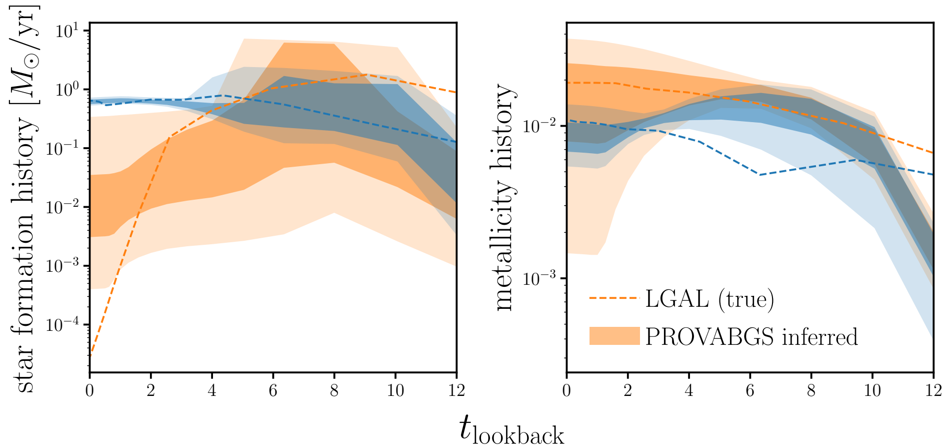

Figure 11. With the PROVABGS SPS model, we can infer posteriors on the full star formation and metallicity histories from observations. We present the inferred SFH and ZH for an arbitrarily chosen star-forming (blue) and quiescent galaxy (orange). The shaded region represent the 64 and 95% confidence intervals of the SFH and ZH posteriors. For comparison, we include the true SFH and ZH (dashed). The inferred SFH and ZH show good agreement with the true values; however, similar to the inferred $\overline{\rm SFR}_{\rm 1Gyr}$ and $Z_{\rm MW}$, the SFH and ZH are significantly impacted by priors imposed by the SPS model.

for mitigating the effect of the model priors. For instance, including spectra in the SED modeling significantly reduces the bias of our $Z_{\rm MW}$ and $t_{\rm age,MW}$ constraints (Figure 8). It also reduces the lower bound on the inferred $\overline{\rm SFR}_{\rm 1Gyr}$. In fact, without spectra, we are dominated by priors on $\overline{\rm SFR}_{\rm 1Gyr}$ and cannot robustly infer galaxy properties of quiescent galaxies with $\log \overline{\rm SFR}_{\rm 1Gyr} < 0$ dex.

PROVABGS will be a value-added galaxy catalog with unprecedented statistical power. With physical galaxy properties of over 10 million DESI BGS galaxies, PROVABGS will provide a transformational galaxy sample to extend previous statistical galaxy studies. For example, we will be able to make the most precise measurement of the stellar mass function (Li & White 2009; Moustakas et al. 2013, SMF), star-forming sequence (Noeske et al. 2007), or any other summary statistic of galaxy populations. Galaxy studies examining the galaxy-halo connection can also be extended to exploit the additional statistical power of PROVABGS (*e.g.* Tinker et al. 2011; Wetzel et al. 2013; Zu & Mandelbaum 2015; Hahn et al. 2017, 2019). With detailed galaxy properties, PROVABGS will also enable multiple-tracer galaxy clustering analyses that can circumvent cosmic variance in inferring cosmological parameters (Seljak 2009; McDonald & Seljak 2009; Wang & Zhao 2020). Analyses exploiting new forward modeling approaches, such as Hahn et al. (2021), will also greatly benefit from the statistical power of PROVABGS. insert more science applications here. TODO

In addition to the applications above, PROVABGS will also unlock applications that can exploit the full posteriors that the catalog will provide. In this work, we utilized the posteriors in order to quantify accuracy and precision of galaxy population constraints using population inference with a hierarchical Bayesian approach. This is only the *simplest* illustration of such an approach. We can also use population inference to robustly derive galaxy property distributions of galaxy subpopulations — *without stacking observations*. Stacking makes the strong assumption that galaxies that are grouped together in some *e.g.* color-space are from a subpopulation with the same properties. This assumption fails if, for instance, there are multiple disparate galaxy subpopulations that are degenerate in color-space and therefore are included in the stack. Furthermore, with the PROVABGS posteriors we can infer fully probabilistic galaxy population statistics, which will allow us to robustly probe even the lowest signal-to-noise regime. A probabilistic SMF of BGS, for example, will provide accurate constraints on the low mass end ($\sim 10^7 M_\odot$; Figure 2), which has important implications for both galaxy evolution and cosmology. With all of the applications listed above, PROVABGS will enable us to fully extract the statistical power of ∼10 million BGS galaxies.

6. SUMMARY

Over the next five years, DESI will measure spectra for >30 million galaxies, including >10 million galaxies in BGS during bright time. Each DESI galaxy will also have optical photometry from the Legacy Surveys. BGS, which will extend out to $z \sim 0.6$, will provide a $r < 19.5$ magnitude-limited sample of ∼10 million galaxies spanning a wide range of galaxy properties with high completeness and >95% redshift efficiency. It will also include a sample of fainter galaxies down to $r < 20.175$ selected based on a fiber magnitude and color. This upcoming dataset offers a unique opportunity to leverage its statistical power for galaxy evolution and maximize its scientific impact. Accurate galaxy properties for such a galaxy sample, for instance, would enable us to measure population statistics and empirical relations of galxaies with unprecedented precision. It would also enable more complete and precise comparisons between observations and galaxy formation models, which will shed light into the physical processes of galaxy evolution. To exploit this opportunity, we will construct the PRObabilistic Value-Added Bright Galaxy Survey (PROVABGS) catalog, where we will apply state-of-the-art Bayesian SED modeling to jointly analyze DESI spectroscopy and LS photometry. PROVBGS will provide full posterior distributions of galaxy properties, such as stellar mass (M_*), star formation rate (SFR), stellar metallicity ($Z_{\rm MW}$), and stellar age ($t_{\rm age,MW}$), for all >10 million BGS galaxies.

In this work, we present and validate the SED model, Bayesian inference framework, and other methodology that will be used to construct PROVABGS. We use 2,123 galaxies in the L-GALAXIES semi-analytic model to construct realistic synthetic DESI spectra and photometry. We build SEDs using SPS based on the star formation and chemical enrichment histories of the simulated galaxies. Then, we simulate the SEDs using the forward modeling pipeline used in the BGS survey design. Afterwards, we apply the PROVABGS SED modeling on the mock DESI observations to derive posteriors on M_*, $\overline{\rm SFR}_{\rm 1Gyr}$, $Z_{\rm MW}$, and $t_{\rm age,MW}$. From the posteriors and the population inference we conduct to quantify accuracy and precision, we find:

- Overall, we derive posteriors on galaxy properties that are in good agreement with the true properties of the simulated galaxies. Furthermore, with posteriors rather than point estimates we accurately estimate the uncertainties on the galaxy properties. We infer posteriors with the following levels of precision: $\sigma_{M_*} \sim 0.1$ dex, $\sigma_{\log \overline{\rm SFR}_{\rm 1Gyr}} \sim 0.1$ dex, $\sigma_{\log Z_{\rm MW}} \sim 0.15$ dex, and $\sigma_{t_{\rm age,MW}} \sim 0.5$ Gyr. Our results also demonstrate that we successfully marginalize over the effect of dust and other nuisance parameters.

- Like any SED model, the PROVABGS SED model also imposes significantly non-uniform priors on galaxy properties. We find that these priors impose a lower bound on $\overline{\rm SFR}_{\rm 1Gyr}$ of $\overline{\rm SFR}_{\rm 1Gyr} > 10^{-1} M_\odot/{\rm yr}$. It also biases $Z_{\rm MW}$ by ∼0.3 dex for observations with low spectral signal-to-noise and imposes an upper bound of $t_{\rm age,MW} < 8$ Gyr. We characterize the priors in detail so that constraints on galaxy properties can be interpreted in future studies that will use PROVABGS.
- We compare the posteriors derived from DESI spectrophotometry to those derived from photometry alone. Including DESI spectra substantially improves the constraints on galaxy properties. Moreover, jointly analyzing spectra is *essential* for mitigating the impact of the SED model priors. For example, with photometry alone, the priors impose a more restrictive $\overline{\rm SFR}_{\rm 1Gyr} > 1 M_\odot/{\rm yr}$ lower bound and bias $Z_{\rm MW}$ ∼0.5 dex.

We demonstrate with our mock challenge that we will derive accurate and precise constraints on specific galaxy properties in PROVABGS. Beyond M_*, $\overline{\rm SFR}_{\rm 1Gyr}$, $Z_{\rm MW}$, and $t_{\rm age,MW}$, which we focus on in this work, PROVABGS will also constrain star formation and metallicity histories. With galaxy properties of over >10 million BGS galaxies, current galaxy studies will be able to use the PROVABGS catalog to exploit the statistical power of BGS for the most precise measurements of various galaxy relations. Since the BGS samples span a wide range of galaxies, PROVABGS will also enable galaxy studies to investigate less explored regimes, such as the low mass galaxy populations. PROVABGS will be a fully probabilistic catalog with posteriors for all the properties that accurately capture their uncertainties. With these posteriors, we can also conduct more rigorous statistical analyses using new techinques such as population inference and Bayesian hierarchical inference. We demonstrate one such an approach in this work by using population inference, to estimate the overall accuracy and precision of our galaxy property constraints. These methods will not only improve the accuracy of our analyses but will allow us to fully exploit the statistical power of DESI observations.

Despite the overall success of the PROVABGS methodologies, as demonstrated in this mock challenge, there are some limitations. For instance, we only consider a simple model for the effect of the DESI fiber aperture and flux calibration. We reserve a more detailed investigation for Ramos *et al.* (in prep.). We also do not consider varying the isochrones, stellar library, or IMF. Instead, we will release multiple versions of PROVABGS with different sets of assumptions. Lastly, we find that the most significant limitation to deriving accurate galaxy properties comes from the prior imposed by the SED model. We will address this limitation and present a method to impose uniform priors on galaxy properties in Hahn (in prep.).

DESI has started its main 5 year operation. Already, as part of survey validation, DESI has collected over 150,000 spectra of BGS galaxies that will be released in the Survey Validation Data Assembly (SVDA). The SVDA release will also be accompanied by papers describing the data reduction pipeline, redshift fitting algorithm, fiber assignment, survey operation and simulations, visual inspection, and target selection for the various tracers. Finally, using BGS observations in the SVDA, we will construct and release the PROVABGS-SV catalog and present the stellar mass function measured from it in the subsequent paper.

ACKNOWLEDGEMENTS

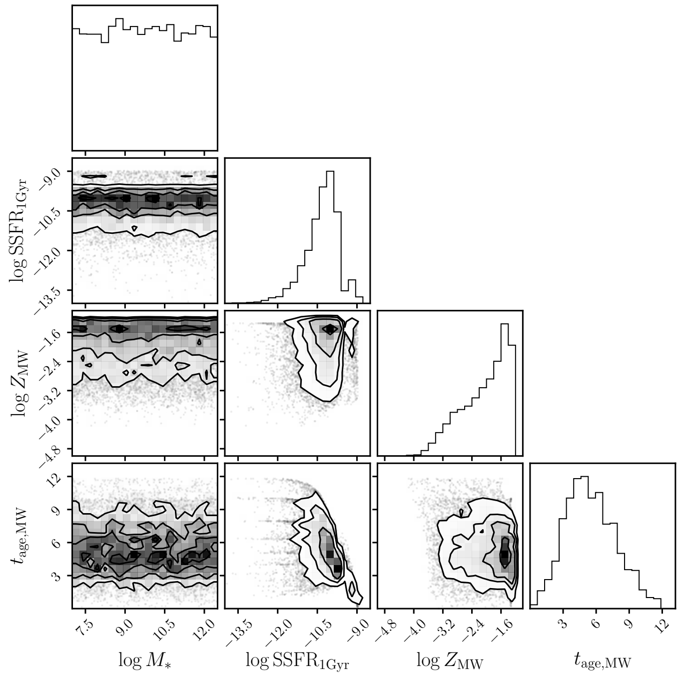

Figure 12. Priors imposed by our SPS model on $\log M_*$, $\log \overline{\mathrm{SSFR}}_{1\mathrm{Gyr}}$, and $\log Z_{\mathrm{MW}}$ at $z = 0.1$.

It's a pleasure to thank Adam Carnall, Kartheik Iyer, Joel Leja, Jenny Greene, and Peter Melchior. This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of High Energy Physics, under contract No. DE-AC02-05CH11231. This project used resources of the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. CH is supported by the AI Accelerator program of the Schmidt Futures Foundation.

APPENDIX

A. NON-NEGATIVE MATRIX FACTORIZATION BASES

B. POPULATION INFERENCE

C. SPS MODEL PRIORS

REFERENCES

Alpaslan M., Tinker J. L., 2021, Monthly Notices of the Royal Astronomical Society, 505, 5403

Alsing J., et al., 2019, arXiv:1911.11778 [astro-ph]

Baldry I. K., Balogh M. L., Bower R. G., Glazebrook K., Nichol R. C., Bamford S. P., Budavari T., 2006, Monthly Notices of the Royal Astronomical Society, 373, 469

Baronchelli L., Nandra K., Buchner J., 2020, Monthly Notices of the Royal Astronomical Society, 498, 5284

Behroozi P., Wechsler R. H., Hearin A. P., Conroy C., 2019, Monthly Notices of the Royal Astronomical Society, p. 1134

Betancourt M. J., 2012, arXiv:1010.3436 [physics 10.1063/1.3703631, pp 157–164

Blanton M. R., et al., 2003, The Astrophysical Journal, 594, 186

Blanton M. R., et al., 2005, The Astronomical Journal, 129, 2562

Boquien M., Burgarella D., Roehlly Y., Buat V., Ciesla L., Corre D., Inoue A. K., Salas H., 2019, Astronomy and Astrophysics, 622, A103

Boylan-Kolchin M., Springel V., White S. D. M., Jenkins A., Lemson G., 2009, Monthly Notices of the Royal Astronomical Society, 398, 1150

Brinchmann J., Charlot S., White S. D. M., Tremonti C., Kauffmann G., Heckman T., Brinkmann J., 2004, Monthly Notices of the Royal Astronomical Society, 351, 1151

Calzetti D., 2001, New Astronomy Reviews, 45, 601

Caplar N., Tacchella S., 2019, arXiv:1901.07556 [astro-ph]

Carnall A. C., McLure R. J., Dunlop J. S., Davé R., 2017, arXiv:1712.04452 [astro-ph]

Carnall A. C., Leja J., Johnson B. D., McLure R. J., Dunlop J. S., Conroy C., 2018, arXiv:1811.03635 [astro-ph]

Chabrier G., 2003, Publications of the Astronomical Society of the Pacific, 115, 763

Charlot S., Fall S. M., 2000, The Astrophysical Journal, 539, 718

Choi J., Dotter A., Conroy C., Cantiello M., Paxton B., Johnson B. D., 2016, The Astrophysical Journal, 823, 102

Cichocki A., Phan A.-H., 2009, IEICE Transactions on Fundamentals of Electronics Communications and Computer Sciences, 92, 708

Coil A. L., et al., 2011, The Astrophysical Journal, 741, 8

Collaboration P., et al., 2014, Astronomy & Astrophysics, Volume 571, id.A16, $<$NUMPAGES$>$66$<$/NUMPAGES$>$ pp., 571, A16

Collaboration D., et al., 2016a, arXiv:1611.00036 [astro-ph]

Collaboration D., et al., 2016b, arXiv:1611.00037 [astro-ph]

Conroy C., 2013, Annual Review of Astronomy and Astrophysics, 51, 393

Conroy C., Gunn J. E., 2010, The Astrophysical Journal, 712, 833

Conroy C., Gunn J. E., White M., 2009, The Astrophysical Journal, 699, 486

Daddi E., et al., 2007, The Astrophysical Journal, 670, 156

Davé R., Rafieferantsoa M. H., Thompson R. J., 2017, arXiv:1704.01135 [astro-ph]

Dey A., et al., 2019, AJ, 157, 168

Dickey C. M., et al., 2021, The Astrophysical Journal, 915, 53

Donnari M., Pillepich A., Nelson D., Marinacci F., Vogelsberger M., Hernquist L., 2021, Monthly Notices of the Royal Astronomical Society, 506, 4760

Dotter A., 2016, The Astrophysical Journal Supplement Series, 222, 8

Driver S. P., et al., 2011, Monthly Notices of the Royal Astronomical Society, 413, 971

Févotte C., Idier J., 2011, arXiv:1010.1763 [cs]

Foreman-Mackey D., Hogg D. W., Morton T. D., 2014, The Astrophysical Journal, 795, 64

Genel S., et al., 2014, Monthly Notices of the Royal Astronomical Society, 445, 175

Gerssen J., Wilman D. J., Christensen L., 2012, Monthly Notices of the Royal Astronomical Society, 420, 197

Hahn C., Tinker J. L., Wetzel A. R., 2017, The Astrophysical Journal, 841, 6

Hahn C., Tinker J. L., Wetzel A., 2019, arXiv:1910.01644 [astro-ph]

Hahn C., et al., 2021, IQ Collaboratory III: The Empirical Dust Attenuation Framework – Taking Hydrodynamical Simulations with a Grain of Dust

Handley W., Millea M., 2019, Entropy, 21, 272

Henriques B. M. B., White S. D. M., Thomas P. A., Angulo R., Guo Q., Lemson G., Springel V., Overzier R., 2015, Monthly Notices of the Royal Astronomical Society, 451, 2663

Hogg D. W., Myers A. D., Bovy J., 2010, The Astrophysical Journal, 725, 2166

Iyer K. G., et al., 2020, Monthly Notices of the Royal Astronomical Society, 498, 430

Karamanis M., Beutler F., 2020, arXiv e-prints, p. arXiv:2002.06212

Kauffmann G., et al., 2003, Monthly Notices of the Royal Astronomical Society, 341, 33

Kriek M., Conroy C., 2013, The Astrophysical Journal Letters, 775, L16

Lee D. D., Seung H. S., 1999, Nature, 401, 788

Leja J., Johnson B. D., Conroy C., van Dokkum P. G., Byler N., 2017, The Astrophysical Journal, 837, 170
Leja J., Carnall A. C., Johnson B. D., Conroy C., Speagle J. S., 2019, ApJ, 876, 3
Li C., White S. D. M., 2009, Monthly Notices of the Royal Astronomical Society, 398, 2177
Marchesini D., van Dokkum P. G., Förster Schreiber N. M., Franx M., Labbé I., Wuyts S., 2009, The Astrophysical Journal, 701, 1765
Mathis J. S., 1983, The Astrophysical Journal, 267, 119
McDonald P., Seljak U., 2009, Journal of Cosmology and Astroparticle Physics, 2009, 007
Meisner A. M., Lang D., Schlegel D. J., 2017a, The Astronomical Journal, 153, 38
Meisner A. M., Lang D., Schlegel D. J., 2017b, The Astronomical Journal, 154, 161
Moustakas J., et al., 2013, The Astrophysical Journal, 767, 50
Nelson D., et al., 2018, Monthly Notices of the Royal Astronomical Society, 475, 624
Noeske K. G., et al., 2007, The Astrophysical Journal Letters, 660, L43
O'Donnell C., Behroozi P., More S., 2021, Observing Correlations Between Dark Matter Accretion and Galaxy Growth: II. Testing the Impact of Galaxy Mass, Star Formation Indicator, and Neighbour Colours
Pacifici C., et al., 2015, Monthly Notices of the Royal Astronomical Society, 447, 786
Paxton B., Bildsten L., Dotter A., Herwig F., Lesaffre P., Timmes F., 2011, The Astrophysical Journal Supplement Series, 192, 3
Paxton B., et al., 2013, The Astrophysical Journal Supplement Series, 208, 4
Paxton B., et al., 2015, The Astrophysical Journal Supplement Series, 220, 15
Pillepich A., et al., 2018, Monthly Notices of the Royal Astronomical Society, 473, 4077
Richards S. N., et al., 2016, Monthly Notices of the Royal Astronomical Society, 455, 2826
Rosani G., Pasquali A., La Barbera F., Ferreras I., Vazdekis A., 2018, Monthly Notices of the Royal Astronomical Society, 476, 5233
Ruiz-Macias O., et al., 2021, Monthly Notices of the Royal Astronomical Society, 502, 4328
Salim S., et al., 2007, The Astrophysical Journal Supplement Series, 173, 267
Sánchez-Blázquez P., et al., 2006, Monthly Notices of the Royal Astronomical Society, 371, 703
Seljak U., 2009, Physical Review Letters, 102, 021302
Serra P., Amblard A., Temi P., Burgarella D., Giovannoli E., Buat V., Noll S., Im S., 2011, The Astrophysical Journal, 740, 22
Simha V., Weinberg D. H., Conroy C., Dave R., Fardal M., Katz N., Oppenheimer B. D., 2014, arXiv e-prints, p. arXiv:1404.0402
Somerville R. S., Davé R., 2015, Annual Review of Astronomy and Astrophysics, 53, 51
Sonnenfeld A., Jaelani A. T., Chan J., More A., Suyu S. H., Wong K. C., Oguri M., Lee C.-H., 2019, Astronomy & Astrophysics, Volume 630, id.A71, $<$NUMPAGES$>$19$<$/NUMPAGES$>$ pp., 630, A71
Sparre M., Hayward C. C., Feldmann R., Faucher-Giguère C.-A., Muratov A. L., Kereš D., Hopkins P. F., 2017, Monthly Notices of the Royal Astronomical Society, 466, 88
Springel V., et al., 2005, Nature, 435, 629
Springel V., et al., 2018, Monthly Notices of the Royal Astronomical Society, 475, 676
Tacchella S., et al., 2021, arXiv e-prints, 2102, arXiv:2102.12494
Taylor E. N., et al., 2009, The Astrophysical Journal, 694, 1171
Tinker J., Wetzel A., Conroy C., 2011, preprint, 1107, arXiv:1107.5046
Tinker J. L., Wetzel A. R., Conroy C., Mao Y.-Y., 2017, Monthly Notices of the Royal Astronomical Society, 472, 2504
Trayford J. W., et al., 2017, Monthly Notices of the Royal Astronomical Society, 470, 771
Tremonti C. A., et al., 2004, The Astrophysical Journal, 613, 898
Treu T., Auger M. W., Koopmans L. V. E., Gavazzi R., Marshall P. J., Bolton A. S., 2010, The Astrophysical Journal, 709, 1195
Trevisan M., Mamon G. A., Thuan T. X., Ferrari F., Pilyugin L. S., Ranjan A., 2021, Monthly Notices of the Royal Astronomical Society, 502, 4815
Villaescusa-Navarro F., et al., 2021, The Astrophysical Journal, 915, 71
Wang Y., Zhao G.-B., 2020, Research in Astronomy and Astrophysics, 20, 158

Wetzel A. R., Tinker J. L., Conroy C., van den Bosch F. C., 2013, Monthly Notices of the Royal Astronomical Society, 432, 336

York D. G., et al., 2000, The Astronomical Journal, 120, 1579

Zahid H. J., Geller M. J., Fabricant D. G., Hwang H. S., 2016, The Astrophysical Journal, 832, 203

Zu Y., Mandelbaum R., 2015, Monthly Notices of the Royal Astronomical Society, 454, 1161

da Cunha E., Charlot S., Elbaz D., 2008, Monthly Notices of the Royal Astronomical Society, 388, 1595

van Dokkum P. G., Conroy C., 2010, Nature, 468, 940