Draft version March 28, 2023
Typeset using LaTeX **preprint** style in AASTeX63

PROVABGS Probabilistic Stellar Mass Function of the BGS One-Percent Survey

ChangHoon Hahn[1,*] and DESI collaboration

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

(Dated: `DRAFT -- 6d00025 -- 2021-01-14 -- NOT READY FOR DISTRIBUTION`)

ABSTRACT

We present the probabilistic stellar mass function (pSMF) of galaxies in the DESI Bright Galaxy Survey (BGS), observed during the One-Percent Survey. The One-Percent Survey was one of DESI's survey validation programs that was conducted from April to May 2021, before the start of the main DESI survey. It observed with the same target selection and observing strategy as the main survey and successfully observed the spectra of 143,017 galaxies in the $r < 19.5$ magnitude-limited BGS Bright sample and 95,499 galaxies in the fainter surface brightness and color selected BGS Faint sample over $0 < z < 0.6$. We derive pSMFs from posteriors of stellar mass, M_*, inferred from DESI photometry and spectroscopy using the Hahn et al. (2022a) PROVABGS Bayesian SED modeling framework. We use a hierarchical population inference framework that statistically rigorously propagates the M_* uncertainties. Furthermore, we include correction weights that account for the selection effects and incompleteness of the BGS observations. We present the redshift evolution of the pSMF in BGS as well as the pSMFs of star-forming and quiescent galaxies classified using average specific star formation rates from PROVABGS. Overall, the pSMFs show good agreement with previous SMF measurements in the literature. Our pSMFs showcase the potential and statistical power of BGS, which in its main survey will observe >100× more galaxies. Moreover, we present the population inference framework for subsequent population statistics measurements using BGS, which will characterize the global galaxy population and galaxy scaling relations at low redshifts with unprecendent precision.

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

1. INTRODUCTION

With large galaxy 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), the galaxy population can now largely be characterize with a small number of scaling relations and population statistics (see Blanton & Moustakas 2009, for a review). The stellar mass function (SMF), for instance, precisely characterizes the overall stellar mass, M_*, distribution of

* changhoon.hahn@princeton.edu

galaxies and its evolution across cosmic history (Li & White 2009; Marchesini et al. 2009; Moustakas et al. 2013; Muzzin et al. 2013; Leja et al. 2019; Driver et al. 2022).

The relationship between the stellar masses and star formation rates of galaxies reveal a bimodality in the galaxy population with star-forming galaxies lying on a tightly correlated "star formation sequence" (SFS; Noeske et al. 2007; Daddi et al. 2007; Salim et al. 2007; Speagle et al. 2014; Hahn et al. 2019). From $z \sim 2$, there is an overall decline in star formation rates of galaxies in the SFS. This is accompanied by an increase in the fraction of quiescent galaxies caused by the quenching of star formation in some galaxies (**??**Baldry et al. 2006; **?**). Additional scaling relations among galaxy propreties such as the mass-metallicity relation (**?**) or add more scaling relations here have also been firmly established. They connect some physical insight More precise and accurate measurements of the statistical distributions of the properties for galaxy populations at different cosmic epochs have the potential to shed further light on galaxy formation and evolution.

TODO
TODO

For one, they have the potential to reveal new trends among galaxies undetected by previous observations and open new discovery space. They can also be used to test galaxy formation models spanning empirical models (*e.g.* UniverseMachine; Behroozi et al. 2019), semi-analytic models (*e.g.* Benson 2012; Henriques et al. 2015; Somerville & Davé 2015), and hydrodyanmical simulations (see Somerville & Davé 2015, for a review). Empirical models, for example, have been used to measure the timescale of timescale of star formation quenching (Wetzel et al. 2013; Hahn et al. 2017; Tinker et al. 2017) or the dust content of galaxies (Hahn et al. 2021).

Furthermore, observations have already been used to infer parameters that dictate the physical processes in semi-analytic models (*e.g.* Henriques et al. 2009; Lu et al. 2014; Henriques et al. 2015) Although full parameter exploration is currently computationally prohibitively for hydrodynamical simulations, they have been extensively compared to observations: *e.g.* Genel et al. (2014); Davé et al. (2017); Trayford et al. (2017); Dickey et al. (2021); Donnari et al. (2021). Soon machine learning techniques for accelerating and emulating simulations will enable us to go beyond such comparisons and broadly explore parameter space and galaxy formation models (*e.g.* Villaescusa-Navarro et al. 2022; Jamieson et al. 2022). While many different approaches are available for expanding our understanding of galaxies, they all require statistically powerful galaxy samples with well controlled systematics and well understood selection functions.

One survey that will provide galaxy samples with unprecedented statistical power is the Dark Energy Spectrscopic Instrument (DESI; Collaboration et al. 2016a,b; Abareshi et al. 2022). Over its 5 year operation, DESI will observe galaxy spectra using the 4-meter Mayll telescope at Kitt Peak National Observatory with a focal plane filled with 5000 robotically-actuated fibers that direct the light to ten optical spectrographs. It will observe ∼40 million galaxy spectra over $360 < \lambda < 980$ nm with spectral resolution of $2000 < \lambda/\Delta\lambda < 5500$ over $\sim 14,000$ deg^2, a third of the sky. In addition, DESI galaxies will also have photometry from the Legacy Imaging Surveys Data Release 9 (LS; **?**). 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). DESI began observing its main survey in May 14, 2021.

As part of the survey, DESI is conducting the Bright Galaxy Survey (BGS; Hahn et al. 2022b). BGS spans the same 14,000deg^2 footprint and will include low redshift $z < 0.6$ galaxies that can be observed during bright time, when the night sky is $\sim$2.5$\times$ brighter than nominal dark conditions, BGS will provide two galaxy samples: the BGS Bright sample, a $r < 19.5$ magnitude-limited sample of $\sim$10 million galaxies, and the BGS Faint sample, a fainter $19.5 < r < 20.175$ sample of ~ 5 million galaxies selected using a surface brightness and color. The selection and completeness of the BGS samples are characterized in detail in Hahn et al. (2022b). Compared to the seminal SDSS main galaxy survey, BGS will provide a galaxy sample two magnitudes deeper, over twice the sky, and double the median redshift $z$$\sim$0.2. It will observe a broader range of galaxies than previous surveys and provide an opportunity to measure galaxy population statistics with unprecendented precision.

BGS will also be accompanied by a value-added catalog: the Probabilistic Value-Added BGS (PROVABGS; Hahn et al. 2022a; Kwon et al. 2022). For every BGS galaxy, PROVABGS will provide physical properties such as stellar mass (M_*), average star formation rate ($\overline{\mathrm{SFR}}$), stellar metallicity (Z_*), stellar age ($t_{\rm age}$), and dust content. These galaxy properties will be from state-of-the-art Spectral Energy Distribution (SED) modeling of both DESI photometry and spectroscopy in a full Bayesian inference framework. The SED model is designed to minimize model misspecification by using a highly flexible non-parameteric star formation and metallicity histories as well as a flexible dust attenuation model. Furthermore, the properties will be inferred using a fully Bayesian inference framework and provide statistically rigorous estimates of uncertainties and degeneracies among the properties. Ultimately, PROVABGS will provide consistently measured galaxy properties that will enable analyses to fully take advantage of the statistical power of BGS with new techniques and approaches.

A key application for PROVABGS will be measuring population statistics using statistically rigorous methodology that correctly propagates the uncertainties in galaxy property measurements. Current population statistics are by and large derived from simply binning best-fit point estimates of galaxy properties. Malz & Hogg (2020) demonstrated, in the context of inferring redshif distributions from individual photometric redshift measurements, that using point estimates is statistically incorrect and can lead to biased redshift distributions. Simiarly, the standard approach can also lead to biased population statistics.

Instead, we can estimate population statistics from combining individual PROVABGS posteriors of galaxy properties using population inference in a hierarchical Bayesian framework (*e.g.* Hogg et al. 2010; Foreman-Mackey et al. 2014; Baronchelli et al. 2020). This approach correctly propagates the uncertainties in the galaxy properties from the individual posteriors of galaxies. As a result, they significantly improve the accuracy of population statistics measurements and will enable more accurate measurements of key galaxy scaling relations. In this work, we present the first of such population statistic measurement for BGS: the probabilistic stellar mass function (pSMF).

In particular, we present the pSMF of BGS galaxies observed during the DESI One-Percent Survey, a survey validation program conducted before the main survey operations. We also present the statistical methodology for the population inference as well as the our methods for accounting for observational incompleteness. We begin in Section 2 with an overview of the BGS galaxies observed during the DESI One-Percent Survey. Then, in Section 3, we briefly summarize the PROVABGS SED modeling framework used to infer the physical properties of the BGS galaxies. Afterwards, we present the pSMF inferred from the BGS observations in Section 4. We summarize and discuss our results in Section 5. Throughout this work, we assume AB magnitudes and a flat ΛCDM cosmology described by the final Planck results (Collaboration et al. 2014): . $\Omega_m = 0.307$, $\Omega_b = 0.0483$, $H_0 = 67.8\,\mathrm{km\,s^{-1}Mpc^{-1}}$, $A_s = 2.19 \times 10^{-9}$, $n_s = 0.9635$

2. THE DESI BRIGHT GALAXY SURVEY: ONE-PERCENT SURVEY

DESI began its five years of operations in May 14, 2021. Before its start, DESI conducted the Survey Validation (SV) campaign to verify that the survey will meets its scientific and performance requirements. The SV campaign was divided into two main programs: the first, SV1, characterized the survey's performance for different observing conditions and was used to optimize sample selection. The second, the One-Percent Survey (or SV3), observed a dataset that can be used for representative clustering measurements and deliver a 'truth' sample with high completeness over an area at least 1% of the expected main survey footprint. We refer readers to **?** for details on the DESI SV programs. In this work, we focus on BGS galaxies observed during the One-Percent Survey.

The One-Percent Survey observed on 38 nights from April 2021 to the end of May 2021. During this time DESI observed 288 bright time exposures that cover 214 BGS 'tiles', planned DESI pointings. The tiles were arranged so that a set of 11 overlapping tiles has their centers arranged around a 0.12 deg circle, forming a 'rosette' completeness pattern. In total, the One-Percent Survey observed 20 rosettes covering 180 deg^2 spanning the northern galactic cap (see Figure 1 in Hahn et al. 2022a).

All BGS spectra observed during the One-Percent Survey are reduced using the 'Fuji' version of the DESI spectroscopic data reduction pipeline (**?**). First, spectra are extracted from the spectrograph CCDs using the *Spectro-Perfectionsim* algorithm of **?**. Then, fiber-to-fiber variations are corrected by flat-fielding and a sky model, empirically derived from sky fibers, is subtracted from each spectrum. Afterwards, the fluxes in the spectra are calibrated using stellar model fits to standard stars. The final processed spectra is then derived by co-adding the calibrated spectra across expoures of the same tile. In total, DESI observed spectra of 155,022 BGS Bright and 109,418 BGS Faint targets during the One-Percent Survey.

For each spectrum, redshift is measured using Redrock[1] (**?**), a redshift fitting algorithm that uses χ^2 minimization computed from a linear combination of Principal Component Analysis (PCA) basis spectral templates in three template classes ("stellar", "galaxy", and "quasar"). Redrock also provides measures of redshift uncertainty, `ZERR` and redshift confidence, $\Delta\chi^2$, which corresponds to the difference between the χ^2 values of the best-fit model and the next best-fit model. We restrict our sample to galaxy targets with reliable redshift measurements. We only keep targets with spectra classified as galaxy spectra by Redrock, no Redrock warning flags, $\Delta\chi^2 > 40$, and Redrock redshift uncertainty `ZERR` $< 0.0005(1+z)$. We also exclude any targets observed using malfunctioning fiber positioners. Lastly, we impose a redshift range of $0 < z < 0.6$. After these cuts, our One-Percent Survey BGS sample includes 143,074 BGS Bright galaxies and 96,771 BGS Faint galaxies.

[1] https://redrock.readthedocs.io

3. PROVABGS SED MODELING

For each BGS galaxy, we derive its M_* and other physical properties, $\overline{\rm SFR}$, $Z_{\rm MW}$, and $t_{\rm age,MW}$ from DESI photometry and spectroscopy using the PROVABGS SED modeling framework (Hahn et al. 2022a). PROVABGS models galaxy SEDs using stellar population synthesis with non-parametric star-formation history (SFH) with a starburst, a non-parametric metallicity history (ZH) that evolves with time, and a flexible dust attenuation prescription. The non-parameteric SFH and ZH prescriptions are derived from SFHs and ZHs of simulated galaxies in the Illustris hydrodynamic simulation (Vogelsberger et al. 2014; Genel et al. 2014; Nelson et al. 2015) and provide compact and flexibly representations of SFHs and ZHs. For the stellar population synthesis, PROVABGS uses the Flexible Stellar Population Synthesis (FSPS; Conroy et al. 2009, 2010) model with MIST isochrones (Paxton et al. 2011, 2013, 2015; Choi et al. 2016; Dotter 2016), Chabrier (2003) initial mass function (IMF), and a combination of MILES (Sánchez-Blázquez et al. 2006) and BaSeL (Lejeune et al. 1997, 1998; Westera et al. 2002) spectral libraries.

Furthermore, PROVABGS provides a Bayesian inference framework that infers full posterior probability distributions of the SED model parameter: $p(\theta\,|\,\mathbf{X}^{\rm photo}, \mathbf{X}^{\rm spec})$, where $\mathbf{X}^{\rm photo}$ represents the photometry and $\mathbf{X}^{\rm spec}$ represents the spectroscopy. In total, θ has 13 parameters: M_*, 6 parameters specifying the SFH ($\beta_1, \beta_2, \beta_3, \beta_4, f_{\rm burst}, t_{\rm burst}$), 2 parameters specifying ZH (γ_1, γ_2), 3 parameters specifying dust attenuation ($\tau_{\rm BC}, \tau_{\rm ISM}, n_{\rm dust}$), and a nuisance parameter for the fiber aperture effect. Posteriors accurately estimate uncertainties and degeneracies among galaxy properties. Furthermore, as we later demonstrate, they are essential for statistically rigorous hierarchical population inference.

In practice, accurately estimating a 13 dimensional posterior requires a large number ($\gtrsim$100,000) SED model evaluations, which requires prohibitive computational resources — ~ 10 CPU hours per galaxy. To address this challenge, PROVABGS samples the posterior using the Karamanis & Beutler (2020) ensemble slice Markov Chain Monte Carlo (MCMC) sampling with the ZEUS Python package[2]. PROVABGS further accelerates the inference by using neural emulators for the SED models. The emulators are accurate to subpercent level and $>$100$\times$ faster than the original SED model based on FSPS (Kwon et al. 2022). With ZEUS and neural emulation, deriving a posterior takes $\sim$5 min per galaxy with PROVABGS. Moreover, Hahn et al. (2022a) demonstrated PROVABGS can accurately infer M_* overall the full expected M_* range of BGS, using forward modeled synthetic DESI BGS observations.

In Figure 1, we demonstrate the PROVABGS SED modeling framework for a randomly selected BGS Bright galaxy with $z = 0.2242$ (target ID: 39627757520424630). In the top panels, we present the posteriors of galaxy properties, M_*, $\overline{\rm SFR}$, $Z_{\rm MW}$, and $t_{\rm age,MW}$, inferred from DESI photometry and spectroscopy. We mark the 12, 40, 68, and 86 percentiles of posterior with the contours. The posteriors illustrate that we can precisely measure the properties of BGS galaxies from DESI photometry and spectroscopy. Furthermore, with the full posterior, we accurately estimate the uncertainties on the galaxy properties and the degeneracies among them (*e.g.* M_* and $\overline{\rm SFR}$). In the bottom panels, we compare the PROVABGS SED model prediction using the best-fit parameter values (black) to DESI

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

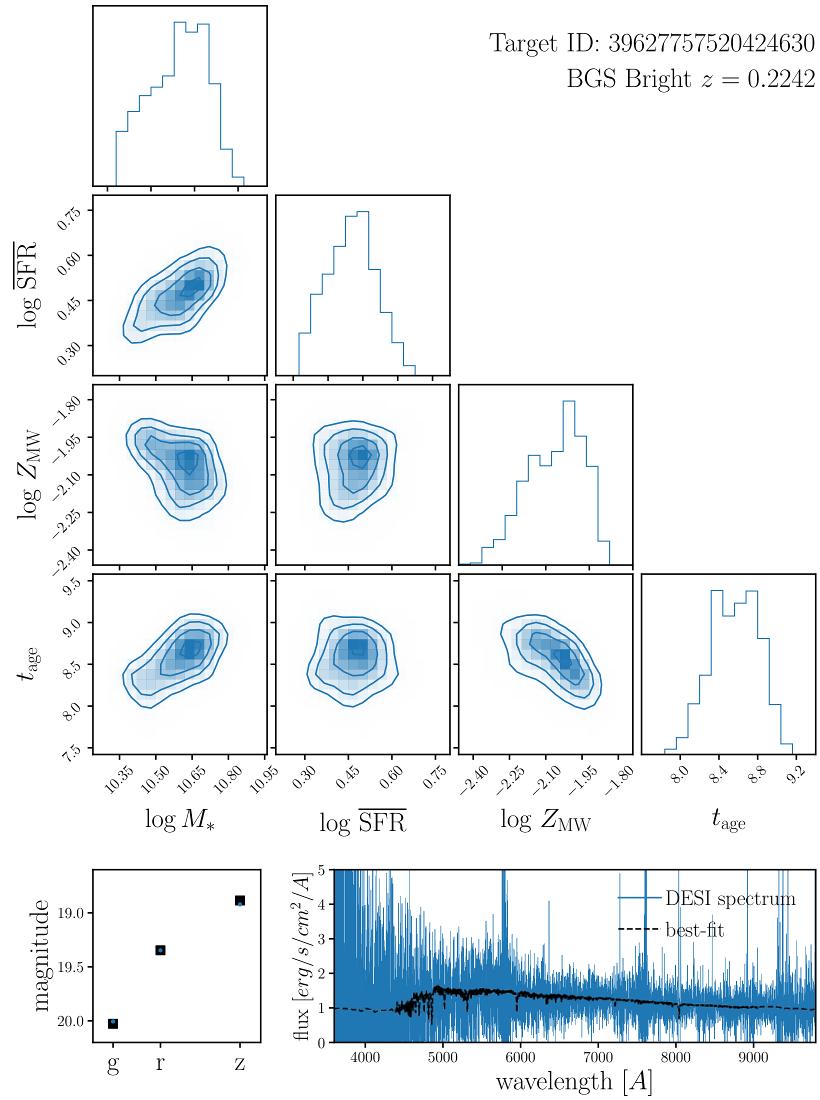

Figure 1. *Top panels*: Posteriors of galaxy properties, M_*, $\overline{\mathrm{SFR}}$, Z_{MW}, and $t_{\mathrm{age,MW}}$, for a randomly selected BGS Bright galaxy with $z = 0.2242$ (target ID: 39627757520424630) inferred using the PROVABGS SED modeling framework from DESI photometry and spectroscopy. The contours mark the 12, 40, 68, and 86 percentiles of the posterior. With the PROVABGS posteriors, we accurately estimate the galaxy properties, their uncertainties, and any degeneracies among them. *Bottom panels*: Comparison of the best-fit PROVABGS SED model prediction (black) to observations (blue). We compare the g, r, and z band photometry in the left panel and spectra in the right panel. We infer the posterior of galaxy properties for every BGS galaxies in the DESI One-Percent Survey.

observations (blue). The left panel compares the optical g, r, and z band photometry while the right panel compares the spectra. The comparison shows good agreement between the best-fit model and the observations.

We derive a PROVABGS posterior (*e.g.* Figure 1) for every galaxy in the DESI One-Percent Survey. In Figure 2, we present the best-fit M_* measurements as a function of z for the BGS galaxies

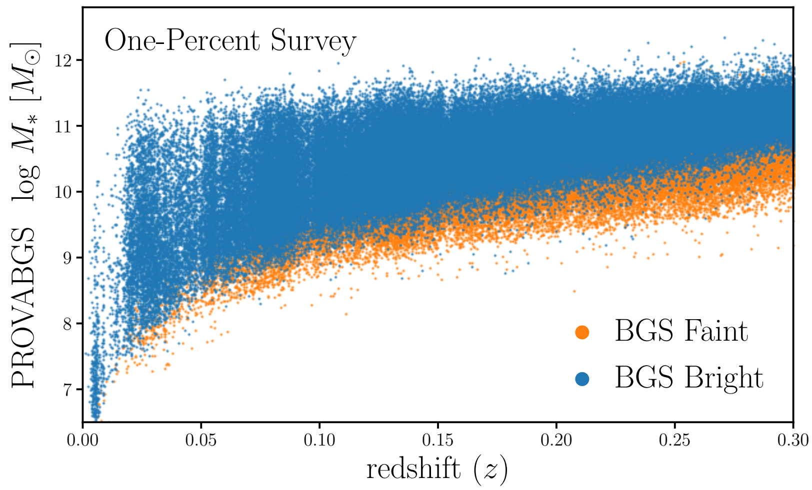

Figure 2. M_* as a function of z of BGS Bright (blue) and Faint (orange) galaxies in the DESI One-Percent Survey. For M_*, we use the best-fit values derived using PROVABGS. BGS Bright is a magnitude-limited sample to $r < 19.5$ while BGS Faint includes fainter galaxies $19.5 < r < 20.175$ selected using $r_{\rm fib}$ and color (Hahn et al. 2022b). In total, we infer the posteriors of 143,017 BGS Bright and 95,499 BGS Faint galaxies in the DESI One-Percent Survey spanning $0 < z < 0.6$.

in DESI One-Percent Survey. We mark the galaxies in the BGS Bright sample in blue and the ones in the BGS Faint sample in orange.

4. RESULTS

From the posteriors of galaxy properties inferred using PROVABGS (Section 3), we derive the marginalized posteriors: $p(M_* \,|\, \boldsymbol{X}_i)$, the marginalized 1D posterior of M_* from observed spectrophotometry $\boldsymbol{X}_i$ of galaxy i. Using these posteriors, we can estimate the probabilistic SMF (pSMF) of BGS galaxies using population inference in a hierarchical Bayesian framework (*e.g.* Hogg et al. 2010; Foreman-Mackey et al. 2014; Baronchelli et al. 2020). In other words, we can infer $p(\phi \,|\, \{\boldsymbol{X}_i\})$, the probability distribution of ϕ given the full DESI BGS observations, $\{\boldsymbol{X}_i\}$. ϕ is the set of population hyperparameters that describe the pSMF, $\Phi(M_*;\phi)$. We again emphasize that this approach is statistically rigorous and correctly propagates the uncertainties in our M_* measurements to the pSMF.

In this work, we estimate the pSMF using a Gaussian Mixture Model (GMM; Press et al. 1992; McLachlan & Peel 2000), which provides a highly flexible description of the M_* distribution:

$$\Phi(M_*;\phi) = \sum_{j=1}^{k} \mathcal{N}(M_*;\phi_j). \tag{1}$$

k represents the number of Gaussian components. ϕ_j represent the mean and standard deviation of the $j^{\rm th}$ Gaussian component of the GMM. Previous works have used parametric functions (*e.g.* Schechter function) to describe the pSMF (Leja et al. 2019). We opt for GMMs in order to produce a non-parametric measuerment of the pSMF. In a subsequent work, Speranza *et al.* (in prep.), we will present the BGS pSMF measured using a parametric model with continuous redshift evolution.

To infer $p(\phi \,|\, \{\boldsymbol{X}_i\})$, we follow the same approach described in Hahn et al. (2022a):

$$
\begin{aligned}
p(\phi \,|\, \{\boldsymbol{X}_i\}) &= \frac{p(\phi)\; p(\{\boldsymbol{X}_i\} \,|\, \phi)}{p(\{\boldsymbol{X}_i\})} && (2)\\
&= \frac{p(\phi)}{p(\{\boldsymbol{X}_i\})} \int p(\{\boldsymbol{X}_i\} \,|\, \{\theta_i\})\; p(\{\theta_i\} \,|\, \phi)\; \mathrm{d}\{\theta_i\}. && (3)\\
&= \frac{p(\phi)}{p(\{\boldsymbol{X}_i\})} \prod_{i=1}^{N} \int p(\boldsymbol{X}_i \,|\, \theta_i)\; p(\theta_i \,|\, \phi)\; \mathrm{d}\theta_i && (4)\\
&= \frac{p(\phi)}{p(\{\boldsymbol{X}_i\})} \prod_{i=1}^{N} \int \frac{p(\theta_i \,|\, \boldsymbol{X}_i)\; p(\boldsymbol{X}_i)}{p(\theta_i)}\; p(\theta_i \,|\, \phi)\; \mathrm{d}\theta_i && (5)\\
&= p(\phi) \prod_{i=1}^{N} \int \frac{p(\theta_i \,|\, \boldsymbol{X}_i)\; p(\theta_i \,|\, \phi)}{p(\theta_i)}\; \mathrm{d}\theta_i. && (6)
\end{aligned}
$$

We can estimate the integral using S_i Monte Carlo samples from the individual posteriors $p(\theta_i \,|\, \boldsymbol{X}_i)$:

$$
p(\phi \,|\, \{\boldsymbol{X}_i\}) \approx \; p(\phi) \prod_{i=1}^{N} \frac{1}{S_i} \sum_{j=1}^{S_i} \frac{p(\theta_{i,j} \,|\, \phi)}{p(\theta_{i,j})}. \tag{7}
$$

Since the sample of BGS galaxies is not volume-limited and complete as a function of M_*, we must account for the selection effect and incompleteness when estimating the pSMF. To account for the selection effects of the BGS samples, we include weights derived from $z_i^{\rm max}$, the maximum redshift that galaxy i could have and still be included in the BGS samples. We derive $z_i^{\rm max}$ for every galaxy using by redshifting the SED predicted by the best-fit parameters. We then derive $V_i^{\rm max}$, the comoving volume out to $z_i^{\rm max}$, and include a factor of $1/V_i^{\rm max}$ in the galaxy weight w_i.

Next, we include correction weights for spectroscopic incompleteness driven by fiber assignment and redshift failures. Incompletenss from fiber assignment is due to the fact that DESI is not able to assign fibers to all galaxies included in the BGS target selection. Furthermore, due to the clustering of galaxies there is significant variation in the assignment probability. Meanwhile, incompleteness from redshift failure is caused by the fact that we do not successfully measure the redshift for every spectra and the redshift failure rate depends significantly on the surface brightnesses of the galaxies and the signal-to-noise ratio of the spectra. We describe how we derive the incompleteness correction weights for fiber assignment and redshift failures, $w_{i,\rm FA}$ and $w_{i,\rm ZF}$, in Appendix A. Each BGS galaxy is assigned a weight of $w_i = (w_{i,\rm FA} \times w_{i,\rm ZF})/V_i^{\rm max}$.

We modify Eq. 2 to include galaxy weights, w_i:

$$p(\phi \,|\, \{\boldsymbol{X}_i\}) \approx \frac{p(\phi)}{\prod_{i=1}^{N} p(\boldsymbol{X}_i)^{w_i}} \prod_{i=1}^{N} \left(\int p(\boldsymbol{X}_i \,|\, \theta_i)\; p(\theta_i \,|\, \phi)\; \mathrm{d}\theta_i \right)^{w_i} \tag{8}$$

$$\approx \frac{p(\phi)}{\prod_{i=1}^{N} p(\boldsymbol{X}_i)^{w_i}} \prod_{i=1}^{N} \left(\sum_{j=1}^{S_i} \frac{p(\theta_{i,j} \,|\, \phi)}{p(\theta_{i,j})} \right)^{w_i} \tag{9}$$

$$\approx \frac{p(\phi)}{\prod_{i=1}^{N} p(\boldsymbol{X}_i)^{w_i}} \prod_{i=1}^{N} \left(\sum_{j=1}^{S_i} \frac{q_\phi(\theta_{i,j})}{p(\theta_{i,j})} \right)^{w_i}. \tag{10}$$

In practice, we do not derive the full posterior $p(\phi \,|\, \{\boldsymbol{X}_i\})$. Instead we derive the maximum a posteriori (MAP) hyperparameter $\phi_{\rm MAP}$ that maximizes $p(\phi \,|\, \{\boldsymbol{X}_i\})$ or $\log p(\phi \,|\, \{\boldsymbol{X}_i\})$. We expand,

$$\log p(\phi \,|\, \{\boldsymbol{X}_i\}) \approx \log p(\phi) + \sum_{i=1}^{N} w_i \log \left(\sum_{j=1}^{S_i} \frac{q_\phi(\theta_{i,j})}{p(\theta_{i,j})} \right). \tag{11}$$

Since the first two terms are constant, we derive $\phi_{\rm MAP}$ by maximizing

$$\max_{\phi} \quad \sum_{i=1}^{N} w_i \log \left(\sum_{j=1}^{S_i} \frac{q_\phi(\theta_{i,j})}{p(\theta_{i,j})} \right). \tag{12}$$

using the Adam optimizer (Kingma & Ba 2017). We derive $\phi_{\rm MAP}$ for BGS galaxies in redshift bins of width $\Delta z = 0.04$ starting from $z = 0.01$ in order to examine the redshift evolution of the SMF within BGS.

4.1. *The Probabilistic Stellar Mass Function*

We present the probabilistic SMF (pSMF) of $0.01 < z < 0.05$ BGS galaxies in the One-Percent Survey in Figure 3 (black line). The shaded regions represent the uncertainties of the pSMF from sample variance, which we derive using a standard jackknife technique (Appendix B) and are conservative estimates (Norberg et al. 2009). In the left panel, we also present the pSMFs of the BGS Bright (blue) and Faint (orange) galaxies. BGS Bright galaxies are selected using a $r > 19.5$ magnitude limit. As a result, the BGS Bright sample is M_* complete above $M_{\rm lim} > 10^{8.975} M_\odot$. We derive $M_{\rm lim}$ in Appendix C and mark the pSMF above the completeness limit in solid and below the limit in dashed. Meanwhile, the BGS Faint sample is selected using a surface brightness and color selection. It includes fainter galaxies, $19.5 < r < 20.175$, with overall lower M_* than the BGS Bright sample.

We also include the SMF estimated using the standard approach (black dotted). This SMF is derived using the best-fit M_* point estimates with the same galaxy weights, w_i. At intermediate M_* range, $10^9 < M_* < 10^{11} M_\odot$, we find good agrement with the pSMF. However, the standard approach significantly underestimates the SMF outside this M_* range. These discrepancies is due to the fact that point estimates of M_* ignore the uncertainties which contribute significantly at the most and

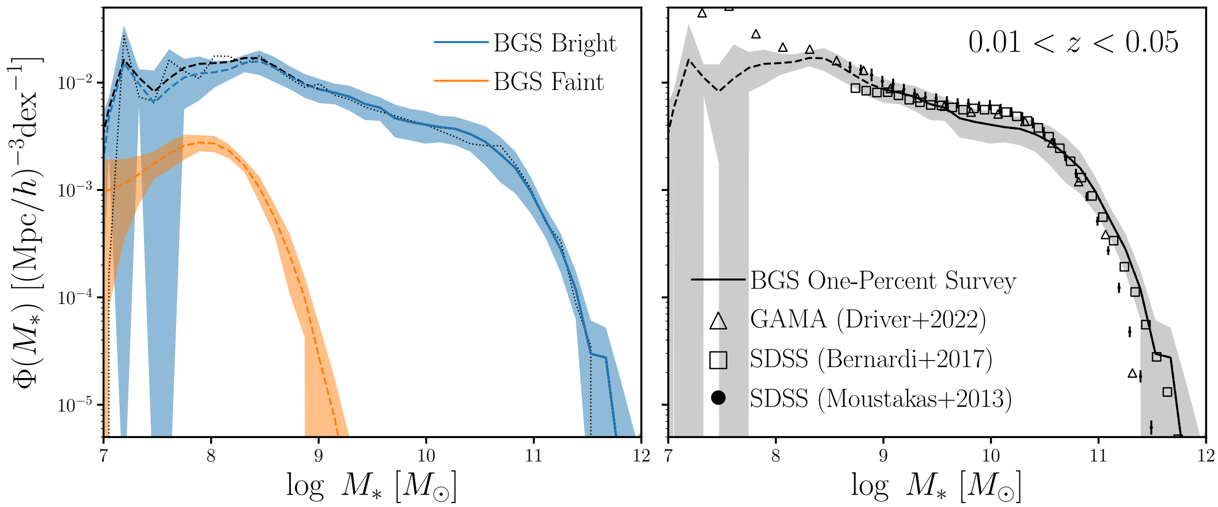

Figure 3. The probabilistic SMF (pSMF) of BGS galaxies in the One-Percent Survey at $0.01 < z < 0.05$ (black line). We represent uncertainties on the pSMF, estimated using a standard jackknife technique (Appendix B), in the shaded regions. The solid line represents the pSMF above the completeness limit $M_* > M_{\rm lim} = 10^{8.975} M_\odot$ (Appendix C). In the left panel, we present the pSMFs of BGS Bright (blue) and Faint (orange) galaxies. In the right panel, we include SMF measurements from previous spectroscopic surveys for comparison: SDSS (Moustakas et al. 2013; Bernardi et al. 2017) and GAMA (Driver et al. 2022). Overall, the pSMF of BGS are in good agreement with SMF measurements from previous surveys.

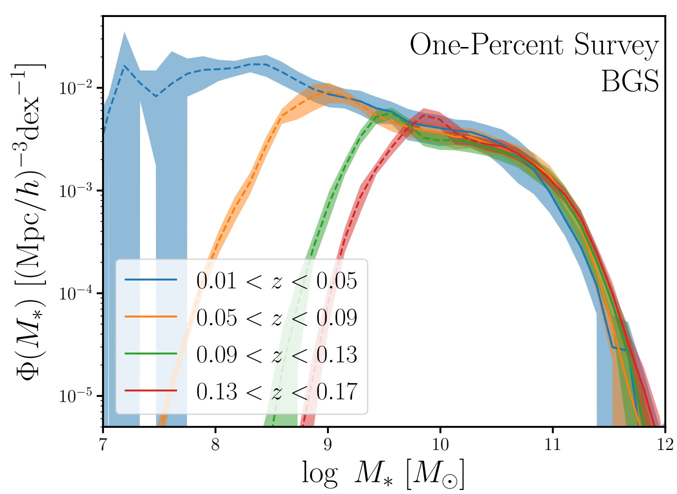

Figure 4. The BGS pSMF over the redshift range $0.01 < z < 0.17$ in bins of $\Delta z = 0.04$. The shaded regions represent the uncertainties on the pSMF, estimated using a standard jackknife technique. The solid lines represent the pSMF above the completeness limit $M_* > M_{\rm lim}$ while the dashed lines represent the pSMF below the limit. There is no significant redshift evolution of the pSMF given the statistical uncertainties. The main BGS survey will observe $> 100\times$ more galaxies than the One-Precent Survey.

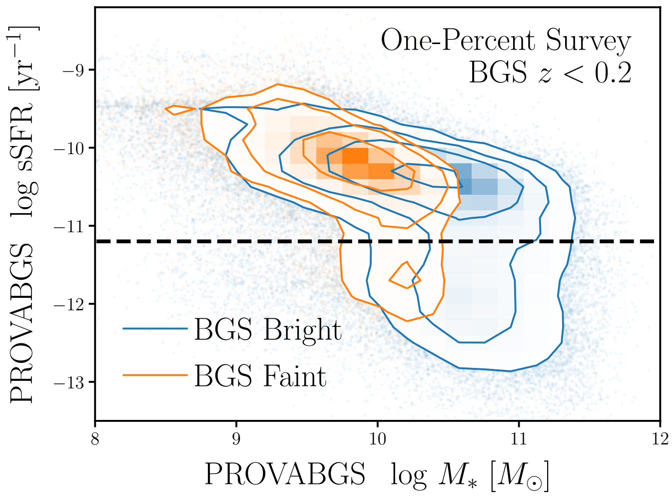

Figure 5. The $M_* -$sSFR distribution of BGS galaxies at $z < 0.2$. sSFR is derived using $\overline{\rm SFR}$, average SFR over the last 1 Gyr, inferred using PROVABGS. The $M_* -$ sSFR distribution is bimodal with star-forming galaxies lying on the star-forming sequence. We classify galaxies with sSFR $> 10^{-11.2}\,{\rm yr}^{-1}$ as star-forming galaxies and sSFR $< 10^{-11.2}\,{\rm yr}^{-1}$ as quiescent.

least massive ends of the SMF. The discpreancies are present in all other redshift bins and underscore the importance of correctly propagating the M_* uncertainties.

In the right panel, we compare the BGS pSMF to SMF measurements from previous spectroscopic surveys: SDSS (Moustakas et al. 2013; Bernardi et al. 2017) (black circle and square) and GAMA (Driver et al. 2022) (black triangle). We note that there is significant variance in SMF measurements in the literature, especially at the high M_* end. This is partly due to the different modeling methodologies used to derive M_*, which can contribute >0.1 dex discrepancies (Pacifici et al. 2023). Furthermore, there are also discrepancies due to photometric corrections applied to SDSS photometry, assumptions on the stellar populations, and dust (Bernardi et al. 2017). In a subsequent work, we will present a detailed comparison of BGS M_* measurements using different methods. Overall, we find good agreement with previous SMF measurements, especially in the intermediate M_* range where we precisely infer the pSMF.

In Figure 4, we present the redshift evolution of the pSMF over $0.01 < z < 0.17$ in redshift bins of $\Delta z = 0.04$. The shaded region represent the jackknife uncertainties for the pSMF. The solid line represents the pSMF above $M_{\rm lim}$ while the dashed lines represent the pSMF below the limit. We only include 4 redshift bins, since $M_{\rm lim} > 10^{10.5} M_\odot$ for $z > 0.17$ (Table 2). The pSMFs in Figure 4 do not reveal a significant redshift dependence given their uncertainties. We note that the large uncertainties for the $0.01 < z < 0.05$ pSMF is driven by large-scale structure at RA ~ 195 deg, Dec ~ 28 deg, and $z \sim 0.244$. The main BGS survey will observe $> 100\times$ more BGS galaxies than the One-Percent Survey and enable pSMF measurements with unprecedented precision.

4.2. *Star-Forming and Quiescent Galaxies in the BGS*

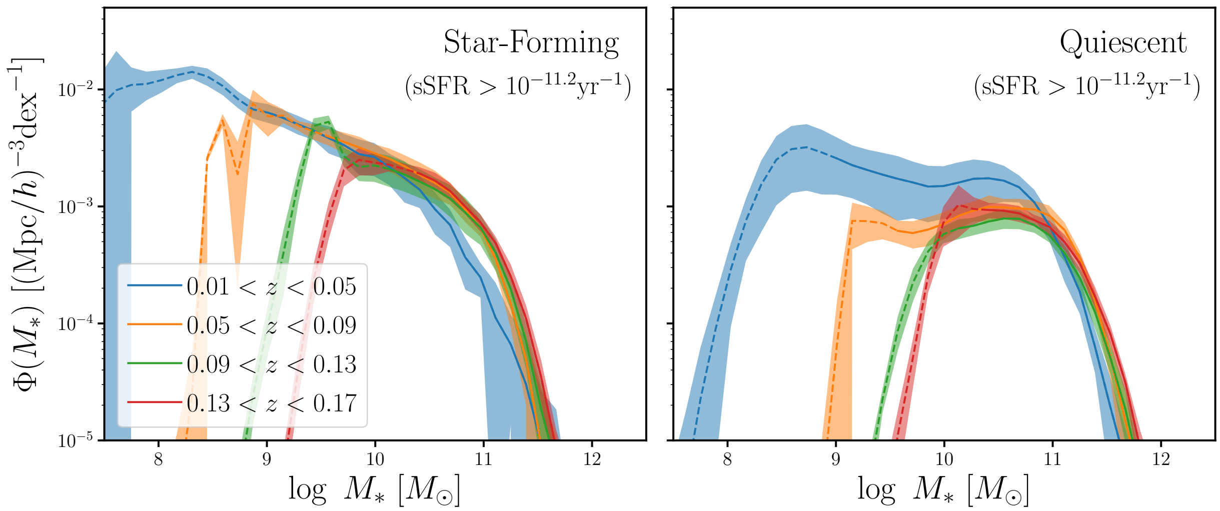

Figure 6. The pSMF of star-forming (left) and quiescent (right) BGS Bright galaxies over $0.01 < z < 0.17$ in bins of $\Delta z = 0.04$. Star-forming and quiescent galaxies are classified using an empirically determined $\mathrm{sSFR} = 10^{-11.2}\mathrm{yr}^{-1}$ cut. We represent the uncertainties for the pSMF in the shaded regions and the pSMFs above/below the M_* completeness limits in solid/dashed lines. The pSMFs suggest a decline in massive, $M_* > 10^{11}M_\odot$ star-forming galaxies and an increase in the quiescent galaxy population at $M_* < 10^{11}M_\odot$ at lower redshifts.

In addition to the pSMF of the full galaxy population, we can also examine the pSMF of the star-forming and quiescent subpopulations using $\overline{\mathrm{SFR}}$, average SFR over the last 1 Gyr, inferred using PROVABGS. In Figure 5, we present the distribution of M_* versus average specific SFR, $\overline{\mathrm{sSFR}} = \overline{\mathrm{SFR}}/M_*$, for BGS Bright (blue) and Faint (orange) galaxies at $z < 0.2$. The $M_* - \mathrm{sSFR}$ distribution of the BGS galaxies reveal a clear bimodality with star-forming galaxies lying on the SFS and quiescent galaxies lying $\gtrsim$1 dex below the sequence. Figure 5 also confirms that BGS Faint galaxies have overall lower M_* than BGS Bright galaxies and are primarily star-forming galaxies. This is due to the fact that the $(z - W1) - 1.2(g - r) + 1.2$ color used to select BGS Faint galaxies is a proxy for Hα and Hβ emission lines.

To further examine the star-forming and quiescent galaxy populations, we classify BGS Bright galaxies as star-forming or quiescent using a $\overline{\mathrm{sSFR}} = 10^{-11.2}\,\mathrm{yr}^{-1}$ cut. We determine this cut empirically based roughly on the sSFR of the "green valley" between the SFS and the quiescent mode. We opt for a $\overline{\mathrm{sSFR}}$ cut rather than more sophisticated methods in the literature (*e.g.* Hahn et al. 2019; Donnari et al. 2019) for simplicity. In Figure 6, we present the pSMF of star-forming and quiescent BGS Bright galaxies at $0.01 < z < 0.17$ in bins of $\Delta z = 0.04$. The shaded regions represent the jackknife uncertainties for the pSMF. The solid lines represent the pSMFs above the completeness limit while the dashed lines represent the pSMFs below the limit. The pSMF of quiescent galaxies suggest an increase in the number of galaxies below $M_* < 10^{11}M_\odot$. Meanwhile, the pSMF of star-forming galaxies shows a possible decline at the massive end over $0.01 < z < 0.17$.

Next, we present the fraction of quiescent galaxies in BGS Bright as a function of M_* over $0.01 < z < 0.17$ in Figure 7. The quiescent fraction is derived by taking the ratio of the pSMFs of quiescent

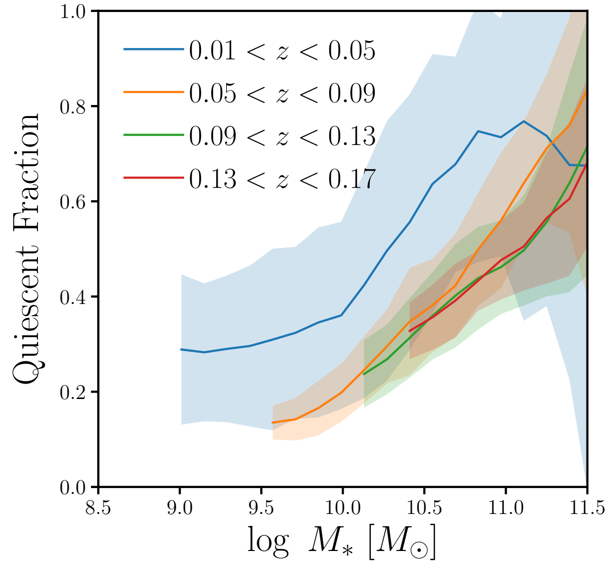

Figure 7. The quiescent fraction of BGS Bright galaxies over $0.01 < z < 0.17$ in bins of $\Delta z = 0.04$. We present the uncertainties in the shaded region and only include the quiescent fraction above the M_* completeness limit. The quiescent fractions increase with M_* at all redshifts. Furthermore, the quiescent fractions suggest an overall increase in the quiescent population with lower redshift.

galaxies over all galaxies and measured for each $\Delta z = 0.04$ bin. The shaded region represent the uncertainties derived from propagating the jackknife uncertainties of the pSMFs. We focus on the quiescent fraction of BGS Bright galaxies above the M_* completeness limit: $M_* > M_{\rm lim}$. At each redshift bin, the quiescent fraction increases with M_* to ∼1 at $M_* \sim 10^{11.5} M_\odot$. The quiescent fraction also suggests an increase in the quiescent population with redshift. Although the significant statistical uncertainties obfuscate a clear trend, the quiescent fraction evolution is in good qualitative agreement with previous works (*e.g.* Baldry et al. 2006; Iovino et al. 2010; Peng et al. 2010; Hahn et al. 2015). Upcoming observations from the DESI main survey will increase the number of BGS galaxies by >100× and enable precise comparisons of the quiescent fraction measurements.

5. SUMMARY AND DISCUSSION

Over its five year operation, starting on May 2021, the DESI Bright Galaxy Survey (BGS) will observe the spectra of ∼15 million galaxies out to $z < 0.6$ over 14,000 deg^2. BGS will produce two main galaxy samples: a $r < 19.5$ magnitude-limited BGS Bright sample and a fainter $19.5 < r < 20.175$ surface brightness and color selected BGS Faint sample. Compared to the SDSS main galaxy survey, the BGS galaxy samples will be over two magnitudes deeper, twice the sky, and double the median redshift $z \sim 0.2$. They will include diverse galaxy subpopulations that have the potential to reveal new trends among galaxies that were preivously undetectable and open new discovery space.

In addition, each galaxy in BGS will have measurements of its detailed physical properties (*e.g.* M_*, SFR, Z_*, $t_{\rm age}$) from PROVABGS. These properties will be inferred from DESI spectrophotometry using state-of-the-art SED modeling in a fully Bayesian inference framework. PROVABGS will provide statistically rigorous estimates of uncertainties and degeneracies among the properties. With these measurements, BGS will be a consistent and statistical powerful galaxy sample to measure scaling relations and population statistics to characterize the global galaxy population and test galaxy formation models with unprecedented precision.

In this work, we showcase the potential of BGS by presenting the probabilistic stellar mass function (pSMF) using ∼250,000 BGS galaxies observed solely during one of DESI's survey validation program. The pSMF are derived using a hierarhical population inference framework that statistically rigorously propgates uncertainties on M_* and provide improve estimates of the SMF at the lowest and highest M_* regimes. We also describe how we account for selection effect and incompleteness in the BGS observations (Appendix A). Overall, we find good agreement between our pSMF and previous SMF measurements in the literature. We also examine the pSMF of the star-forming and quiescent galaxy population classified using a simple $\overline{\rm sSFR} = 10^{-11.2}{\rm yr}^{-1}$ cut and find qualitative agreement with previous works.

This work is first of a series of papers that will present population statistics for BGS galaxies using PROVABGS. For the pSMF in this work, we used a flexible GMM to provide a non-parametric measurement of the SMF. In a subsequent work, Speranza *et al.* (in prep.), we will present the pSMF of BGS measured using a parametric model with continuous redshift evolution. In another work, we will present in depth comparison M_* measured using different methodologies and assumptions. Lastly, the hierarchical population inference framework presented in this work can be extended to population statistics beyond the SMF. We will extend the framework to the SFR-M_* distribution and present the probabilistic SFR-M_* distribution and quiescent fraction in future work.

All of the pSMFs presented in this work are measured from BGS galaxies observed from April to May of 2021 during the DESI One-Percent Survey. Since then, DESI has already completed nearly two years of observations. As of writing (March 2023), DESI has observed nearly 20 million galaxy spectra in total and nearly 9 million BGS galaxy spectra. With three out of the five years of operation remaining, BGS has completed completed ∼60% of its observations and is ahead of schedule. DESI observations will be publicly released periodically, starting with the Early Data Release (EDR) later this year. The EDR will include observations from the One-Percent Survey used in this this work. An accompanying PROVABGS catalog will be released with each data release.

ACKNOWLEDGEMENTS

It's a pleasure to thank . This work was supported by the AI Accelerator program of the Schmidt Futures Foundation.

This research is supported by the Director, Office of Science, Office of High Energy Physics of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231, and by the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility under the same contract; additional support for DESI is provided by the U.S. National Science Foundation, Division of Astronomical Sciences under Contract No. AST-0950945 to the NSF's National Optical-Infrared Astronomy Research Laboratory; the Science and Technologies Facilities Council of the United Kingdom; the Gordon and Betty Moore Foundation; the Heising-Simons Foundation; the French Alternative Energies and Atomic Energy Commission (CEA); the National Council of Science and Technology of Mexico (CONACYT); the Ministry of Science and Innovation of Spain (MICINN), and by the DESI Member Institutions: https://www.desi.lbl.gov/collaborating-institutions.

The DESI Legacy Imaging Surveys consist of three individual and complementary projects: the Dark Energy Camera Legacy Survey (DECaLS), the Beijing-Arizona Sky Survey (BASS), and the Mayall z-band Legacy Survey (MzLS). DECaLS, BASS and MzLS together include data obtained, respectively, at the Blanco telescope, Cerro Tololo Inter-American Observatory, NSF's NOIRLab; the Bok telescope, Steward Observatory, University of Arizona; and the Mayall telescope, Kitt Peak National Observatory, NOIRLab. NOIRLab is operated by the Association of Universities for Research in Astronomy (AURA) under a cooperative agreement with the National Science Foundation. Pipeline processing and analyses of the data were supported by NOIRLab and the Lawrence Berkeley National Laboratory. Legacy Surveys also uses data products from the Near-Earth Object Wide-field Infrared Survey Explorer (NEOWISE), a project of the Jet Propulsion Laboratory/California Institute of Technology, funded by the National Aeronautics and Space Administration. Legacy Surveys was supported by: the Director, Office of Science, Office of High Energy Physics of the U.S. Department of Energy; the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility; the U.S. National Science Foundation, Division of Astronomical Sciences; the National Astronomical Observatories of China, the Chinese Academy of Sciences and the Chinese National Natural Science Foundation. LBNL is managed by the Regents of the University of California under contract to the U.S. Department of Energy. The complete acknowledgments can be found at https://www.legacysurvey.org/.

The authors are honored to be permitted to conduct scientific research on Iolkam Du'ag (Kitt Peak), a mountain with particular significance to the Tohono O'odham Nation.

APPENDIX

A. SPECTROSCOPIC COMPLETENESS

Spectroscopic galaxy surveys, such as BGS, do not successfully measure the redshift for all of the galaxies they target. As a result, this spectroscopic incompleteness must be accounted for when measuring galaxy population statistics such as the SMF. In this appendix, we present how we estimate the spectroscopic incompleteness for BGS and derive the weights we use to correct for its impact on the SMF.

For BGS, spectroscopic incompleteness is primarily driven by fiber assignment and redshift failures. DESI uses 10 fiber-fed spectrographs with 5000 fibers but targets more galaxies than available fibers. For instance, the BGS Bright and Faint samples have ~ 860 and $530\,\mathrm{targets/deg^2}$, respectively. For the 8 $\deg^2$ field-of-view of DESI, this roughly correspond to 11,000 targets, significantly more than the 5000 available fibers. DESI only measures the spectra of targets that are assigned fibers. In fact, of the 5000, a minimum of 400 'sky' fibers are dedicated to measuring the sky background for accurate sky subtraction and an additional 100 fibers are assigned to standard stars for flux calibration **?**.

Furthermore, each fiber is controlled by a robotic fiber positioner on the focal plane. These positioners can rotate on two arms and be positioned within a circular patrol region of radius 1.48 arcmin (**????**). Although the patrol regions of adjacent positioners slightly overlap, the geometry of the positioners cause higher incompleteness in regions with high target density (**?**). To mitigate the incompleteness from the fiber assignment, BGS will observe its footprint with four passes. With this strategy, BGS achieves ∼80% fiber assignment completeness (**?**).

To estimate fiber assignment completeness, we run the fiber assignment algorithm (**?**) on BGS targets 128 separate times. For each BGS galaxy, i, we count the total number of times out of 128 that the galaxy is assigned a fiber: $N_{i,\mathrm{FA}}$. Then to correct for the fiber assignment incompleteness, we assign correction weights

$$w_{i,\mathrm{FA}} = \frac{128}{N_{i,\mathrm{FA}}} \tag{A1}$$

to each BGS galaxy. explain what this means

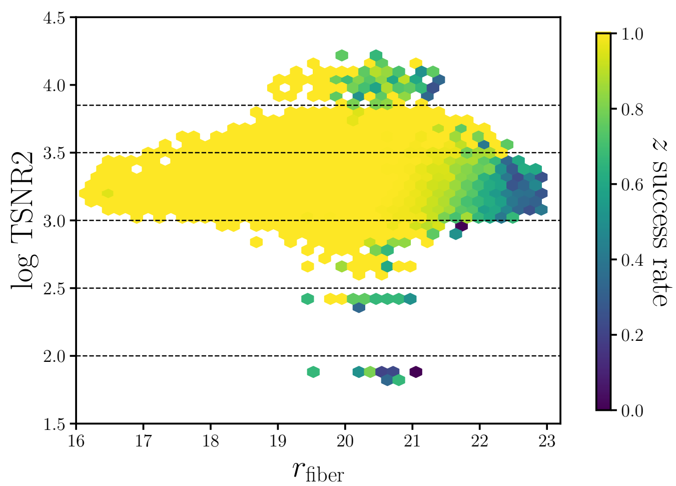

Figure 8. Redsift success rate of BGS Bright galaxies as a function of r_{fiber} and TSNR2. TSNR2 is a statistic that quantifies the signal-to-noise ratio of the observed spectrum. The color map represents the mean redshift success rate in each hexbin. We mark the TSNR2 bins (black dashed) that we use to separately fit the redshift success rate as a function of r_{fiber} using Eq. A3. In each TSNR2 bin, redshift success decreases as r_{fiber} increases.

Although we measure a spectrum for each galaxy assigned a fiber, we do not accurately measure redshifts for every spectra. This redshift measurement failure significantly contributes to spectroscopic incompleteness. For BGS, redshift failure of an observed galaxy spectrum depends mainly on fiber magnitude and a statistic, TSNR2. Fiber magnitude is the predicted flux of the BGS object within a 1.5″diameter fiber; we use r-band fiber magnitude, r_{fiber}. TSNR2 roughly corresponds to the signal-to-noise ratio of the spectrum and is the statistic used to calibrate the effective exposure times in DESI observations (CITE). TODO

In Figure 8, we present the redshift, z, success rate of BGS Bright galaxies as a function of r_{fiber} and TSNR2. In each hexbin, the color map represents the mean z-success rate. We include all hexbins

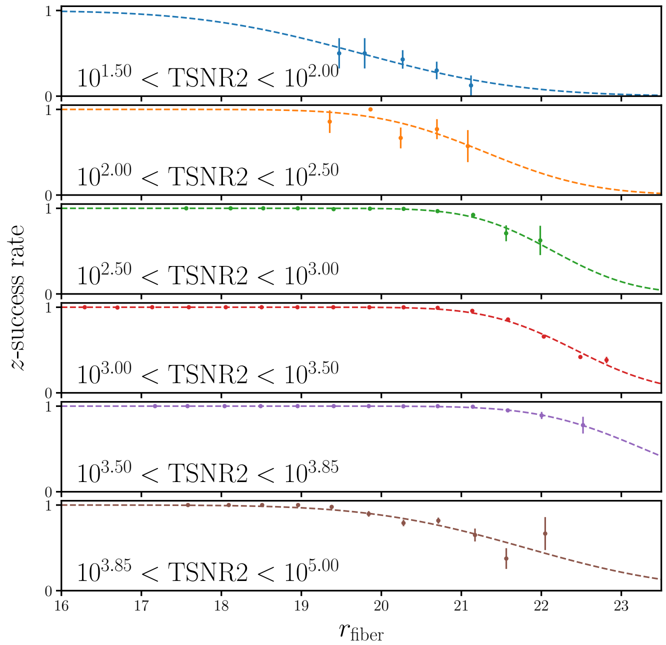

Figure 9. Redshift success rates of BGS Bright galaxies as a function of r_{fiber} in 6 TSNR2 bins. The error bars represent the poisson uncertainties. In each panel, we include the best-fit analytic (Eq. A3) approximation of the redshift success rate (dashed) derived from χ^2 minimization. We use this analytic approximation to calculate the galaxy weights to correct for spectroscopic incompleteness caused by failures to accurately measure redshifts from observed spectra.

with more than 2 galaxies. Overall, the z-success rate depends significantly on r_{fiber}: galaxies with fainter r_{fiber} have lower z-success rates. However, the r_{fiber} dependence itself varies in bins of TSNR2. We mark the edges of the bins in black dashed: $\log \mathrm{TSNR2} = 2.0, 2.5, 3.0, 3.5, 3.85$. Within each of the TSNR2 bins, the r_{fiber} dependence of the z-success rate does not vary significantly. In Figure 9, we present the z-success rate of BGS Bright galaxies as a function of r_{fiber} for each of the 6 TSNR2 bins. We mark the range of TSNR2 in the bottom left of each panel. The errorbars represent the Poisson uncertainties of the z-success rate.

To correct for the effect of redshift failures, we include an additional correction weight for each BGS galaxy:

$$w_{i,\mathrm{ZF}} = \frac{1}{f_{z-\mathrm{sucess}}(r_{\mathrm{fiber},i}, \mathrm{TSNR2}_i)}. \tag{A2}$$

$f_{z-\mathrm{sucess}}(r_{\mathrm{fiber},i}, \mathrm{TSNR2}_i)$ is the z-success rate as a function of r_{fiber} and TSNR2 of the galaxy. Galaxies with $f_{z-\mathrm{sucess}} = 1$ (100% z-success) will have $w_{i,\mathrm{ZF}} = 1.0$ while galaxies with $f_{z-\mathrm{success}} = 0.1$ (10% z-success) will have $w_{i,\mathrm{ZF}} = 10$. For $f_{z-\mathrm{sucess}}(r_{\mathrm{fiber},i}, \mathrm{TSNR2}_i)$, we fit the following functional form for each TSNR2 bin:

$$f_{z-\mathrm{success}}(r_{\mathrm{fiber}}) = \frac{1}{2}\bigg(1 - \mathrm{erf}(c_0(r_{\mathrm{fiber}} - c_1))\bigg). \tag{A3}$$

In Figure 9, we present the best-fit $f_{z-\mathrm{success}}(r_{\mathrm{fiber}})$ for each of the TSNR2 bins in dashed. The best-fit coefficients, c_0, c_1, are derived from χ^2 minimization. We repeat this procedure indepedently for BGS

Table 1. Best-fit coefficients of the z-success rate as a function of $r_{\rm fiber}$ for different TSNR2 bins for BGS Bright and Faint samples.

TSNR2 range	c_0	c_1
BGS Bright		
$10^{1.5} - 10^{2}$	0.443	19.7
$10^{2} - 10^{2.5}$	0.668	21.3
$10^{2.5} - 10^{3}$	0.888	22.1
$10^{3} - 10^{3.5}$	0.822	22.4
$10^{3.5} - 10^{3.85}$	0.698	23.3
$10^{3.85} - 10^{5}$	0.465	21.8
BGS Faint $(z - W1) - 1.2(g - r) + 1.2 \geq 0$		
$10^{1.5} - 10^{2.5}$	1.67	21.1
$10^{2.5} - 10^{3}$	1.65	21.8
$10^{3} - 10^{3.1}$	1.49	22.1
$10^{3.1} - 10^{3.2}$	1.32	22.3
$10^{3.2} - 10^{3.3}$	1.33	22.4
$10^{3.3} - 10^{3.5}$	0.907	23.1
$10^{3.5} - 10^{3.85}$	1.03	23.0
$10^{3.85} - 10^{5}$	0.924	21.6
BGS Faint $(z - W1) - 1.2(g - r) + 1.2 < 0$		
$10^{2.5} - 10^{3}$	1.48	20.9
$10^{3} - 10^{3.1}$	2.40	21.2
$10^{3.1} - 10^{3.2}$	1.30	21.8
$10^{3.2} - 10^{3.3}$	1.27	22.0
$10^{3.3} - 10^{3.5}$	1.83	21.6
$10^{3.5} - 10^{3.85}$	0.798	22.9
$10^{3.85} - 10^{5}$	1.29	20.6

Bright galaxies as well as the BGS Faint galaxies with $(z - W1) - 1.2(g - r) + 1.2 \geq 0$, and BGS Faint galaxies with $(z - W1) - 1.2(g - r) + 1.2 < 0$. We list the best-fit values in bins of TSNR2 for each of the samples in Table 1.

B. UNCERTAINTIES ON THE SMF

We estimate the uncertainties of the SMF from sample variance using the standard jackknife technique. This involves splitting our BGS sample into subsamples and then estimating uncertainties

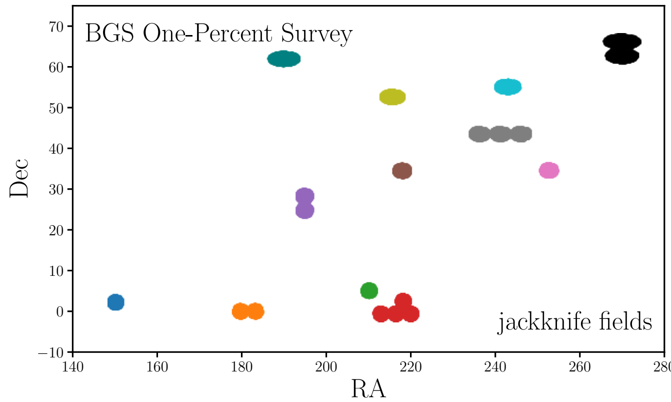

Figure 10. The RA and Dec of the 12 jackknife fields of the BGS One-Percent Survey used to estimate the uncertainties on the SMF from sample variance. We mark each field with a distinct color.

using the subsample-to-subsample variations:

$$\sigma_\Phi = \left(\frac{N_{\rm jack} - 1}{N_{\rm jack}} \sum_{k=1}^{N_{\rm jack}} (\Phi_k - \Phi)^2 \right). \tag{B4}$$

$N_{\rm jack}$ is the number of jackknife subsamples and Φ_k represents the SMF estimated from the BGS galaxies excluding the jackknife subample k. In this work, we split the BGS sample into 12 jackknife fields based on the angular positions of galaxies. We present the jackknife fields in Figure 10 with distinct colors.

C. STELLAR MASS COMPLETENESS

In this appendix, we describe how we derive $M_{\rm lim}$, the M_* limit above which our BGS Bright sample is complete. Although there are various methods for estimating $M_{\rm lim}$ in the literature, *e.g.* based on estimating the mass-to-light ratio (**?**Moustakas et al. 2013), we adopt a simple approach that takes advantage of the fact that BGS Bright is a magnitude-limited sample.

To derive $M_{\rm lim}$ in redshift bins of width $\Delta z = 0.04$, we first split the galaxy sample into narrower bins of $\Delta z/2$. For each narrower redshift bin, $i\Delta z/2 < z < (i+1)\Delta z/2$, we take all the best-fit PROVABGS SEDs from all galaxies in the bin and artificially redshift it to $z' = z + \Delta z/2$:

$$f'_\lambda = f_\lambda \frac{d_L(z)^2}{d_L(z')^2}. \tag{C5}$$

$d_L(z)$ reprents the luminosity distance at redshift z. Afterward, we calculate the r-band magnitudes, r', for f'_λ and impose the $r' < 19.5$ magnitude limit of the BGS Bright. We then compare the M_* distribution of all the galaxies in $i\Delta z/2 < z < (i+1)\Delta z/2$ to the galaxies in $i\Delta z/2 < z < (i+1)\Delta z/2$ with $r' < 19.5$. For instance, we present the M_* distributions of all BGS Bright galaxies in $0.01 < z < 0.03$ (blue) and the BGS Bright galaxies in $0.01 < z < 0.03$ with $r' < 19.5$ (orange) in Figure 11.

Since galaxies become fainter when they are placed at higher redshifts, *i.e.* $r' > r$, the $r' < 19.5$ sample has fewer low M_* galaxies. We determine the M_* at which, more than 10% of galaxies are

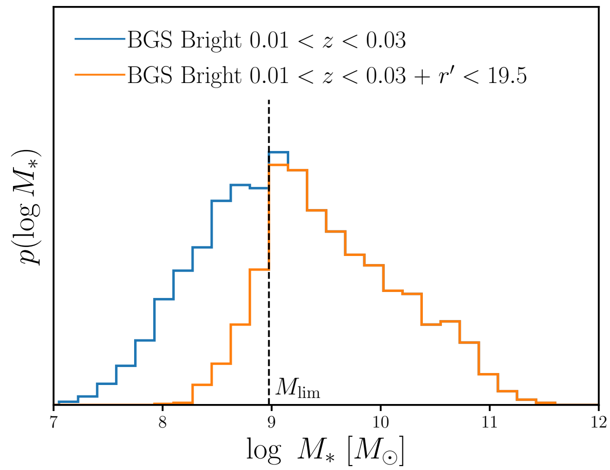

Figure 11. The M_* distribution of BGS Bright galaxies with $0.01 < z < 0.03$ (blue) and the M_* distribution of same set of galaxies that would remain in the BGS Bright magnitude limit if they were redshifted to $z' = z + 0.02$: $r' < 19.5$. We set the stellar mass completeness limit, $M_{\rm lim}$, for $0.01 < z < 0.05$ to the M_* where more than 10% of galaxies are excluded in the latter distribution.

Table 2. Stellar mass completeness limit, $M_{\rm lim}$ for redshift bins of width $\Delta z = 0.04$.

z range	$\log_{10} M_{\rm lim}$
$0.01 - 0.05$	8.975
$0.05 - 0.09$	9.500
$0.09 - 0.13$	10.20
$0.13 - 0.17$	10.38
$0.17 - 0.21$	10.72

excluded in the $r' < 19.5$ sample (black dashed) and set this limit as $M_{\rm lim}$ for the redshift bins: $0.01 < z < 0.05$. Our procedure for deriving $M_{\rm lim}$ takes advantage of the fact that galaxy samples at lower redshifts are complete down to lower M_* than at higher redshifts. We repeat this procedure for all the $\Delta z = 0.04$ redshift bins that we use to measure the SMF. In Table 2, we list $M_{\rm lim}$ values for each of the redshift bins. Furthermore, we present the M_* and redshift relation of BGS Bright galaxies (black) and the stellar mass complete sample (blue) in Figure 12.

REFERENCES

Abareshi B., et al., 2022, Overview of the Instrumentation for the Dark Energy Spectroscopic Instrument (arxiv:2205.10939), doi:10.48550/arXiv.2205.10939

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

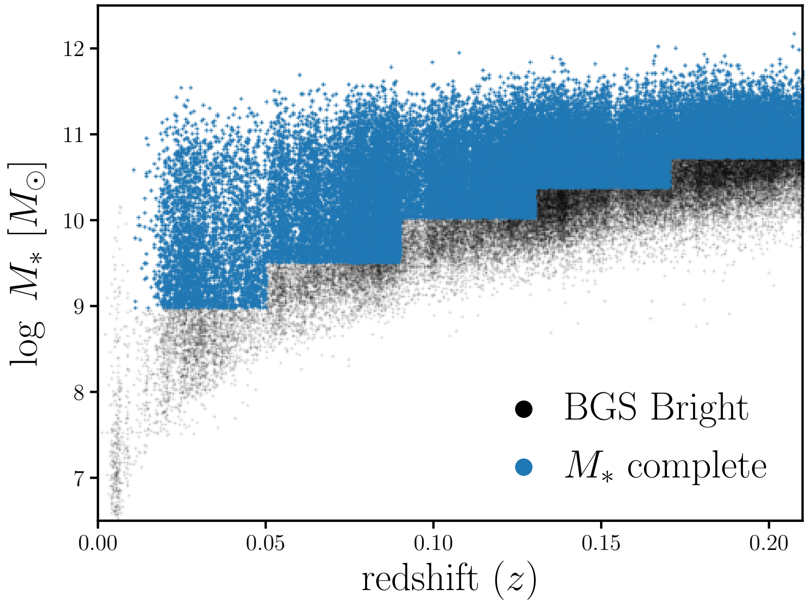

Figure 12. M_* and redshift relation of BGS Bright galaxies in the EDR (black) and the galaxies within the stellar mass completeness limit ($M_* < M_{\rm lim}$; blue). $M_{\rm lim}$ is derived in redshift bins of width $\Delta z = 0.04$. The lowest redshift bin ($0.01 < z < 0.05$) is complete down to $M_* < 10^9 M_\odot$.

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

Benson A. J., 2012, New Astronomy, 17, 175

Bernardi M., Meert A., Sheth R. K., Fischer J. L., Huertas-Company M., Maraston C., Shankar F., Vikram V., 2017, Monthly Notices of the Royal Astronomical Society, 467, 2217

Blanton M. R., Moustakas J., 2009, Annual Review of Astronomy and Astrophysics, 47, 159

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

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

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

Conroy C., White M., Gunn J. E., 2010, The Astrophysical Journal, 708, 58

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

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

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

Donnari M., et al., 2019, Monthly Notices of the Royal Astronomical Society, 485, 4817

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

Driver S. P., et al., 2022, Monthly Notices of the Royal Astronomical Society, p. stac472

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
Hahn C., et al., 2015, The Astrophysical Journal, 806, 162
Hahn C., Tinker J. L., Wetzel A. R., 2017, The Astrophysical Journal, 841, 6
Hahn C., et al., 2019, The Astrophysical Journal, 872, 160
Hahn C., et al., 2021, IQ Collaboratory III: The Empirical Dust Attenuation Framework – Taking Hydrodynamical Simulations with a Grain of Dust
Hahn C., et al., 2022b, DESI Bright Galaxy Survey: Final Target Selection, Design, and Validation
Hahn C., et al., 2022a, The DESI PRObabilistic Value-Added Bright Galaxy Survey (PROVABGS) Mock Challenge
Henriques B. M. B., Thomas P. A., Oliver S., Roseboom I., 2009, Monthly Notices of the Royal Astronomical Society, 396, 535
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
Iovino A., et al., 2010, Astronomy and Astrophysics, 509, A40
Jamieson D., Li Y., Alves de Oliveira R., Villaescusa-Navarro F., Ho S., Spergel D. N., 2022, Field Level Neural Network Emulator for Cosmological N-body Simulations
Karamanis M., Beutler F., 2020, arXiv e-prints, p. arXiv:2002.06212
Kingma D. P., Ba J., 2017, arXiv:1412.6980 [cs]
Kwon K. J., Hahn C., Alsing J., 2022, Neural Stellar Population Synthesis Emulator for the DESI PROVABGS
Leja J., Speagle J. S., Johnson B. D., Conroy C., van Dokkum P., Franx M., 2019, arXiv, p. arXiv:1910.04168
Lejeune T., Cuisinier F., Buser R., 1997, A & A Supplement series, Vol. 125, October II 1997, p.229-246., 125, 229
Lejeune T., Cuisinier F., Buser R., 1998, Astronomy and Astrophysics Supplement, v.130, p.65-75, 130, 65
Li C., White S. D. M., 2009, Monthly Notices of the Royal Astronomical Society, 398, 2177
Lu Y., et al., 2014, The Astrophysical Journal, 795, 123
Malz A. I., Hogg D. W., 2020, How to Obtain the Redshift Distribution from Probabilistic Redshift Estimates
Marchesini D., van Dokkum P. G., Förster Schreiber N. M., Franx M., Labbé I., Wuyts S., 2009, The Astrophysical Journal, 701, 1765
McLachlan G., Peel D., 2000, Finite Mixture Models. Wiley-Interscience
Moustakas J., et al., 2013, The Astrophysical Journal, 767, 50
Muzzin A., et al., 2013, The Astrophysical Journal, 777, 18
Nelson D., et al., 2015, Astronomy and Computing, 13, 12
Noeske K. G., et al., 2007, The Astrophysical Journal Letters, 660, L43
Norberg P., Baugh C. M., Gaztañaga E., Croton D. J., 2009, Monthly Notices of the Royal Astronomical Society, 396, 19
Pacifici C., et al., 2023, The Astrophysical Journal, 944, 141
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
Peng Y.-j., et al., 2010, The Astrophysical Journal, 721, 193
Press W. H., Teukolsky S. A., Vetterling W. T., Flannery B. P., 1992, Numerical Recipes in C (2Nd Ed.): The Art of Scientific Computing. Cambridge University Press, New York, NY, USA
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
Somerville R. S., Davé R., 2015, Annual Review of Astronomy and Astrophysics, 53, 51
Speagle J. S., Steinhardt C. L., Capak P. L., Silverman J. D., 2014, The Astrophysical Journal Supplement Series, 214, 15
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

Villaescusa-Navarro F., et al., 2022, arXiv:2201.01300 [astro-ph]

Vogelsberger M., et al., 2014, Monthly Notices of the Royal Astronomical Society, 444, 1518

Westera P., Lejeune T., Buser R., Cuisinier F., Bruzual G., 2002, Astronomy and Astrophysics, 381, 524

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