The Incomplete Conditional Stellar Mass Function: Unveiling the Stellar Mass Functions of Galaxies at 0.1 < Z < 0.8 from BOSS Observations

Hong Guo 1, Xiaohu Yang 2,3, and Yi Lu 1

1 Key Laboratory for Research in Galaxies and Cosmology, Shanghai Astronomical Observatory, Shanghai 200030, People’s Republic of China; guohong@shao.ac.cn
2 Department of Astronomy, Shanghai Key Laboratory for Particle Physics and Cosmology, Shanghai Jiao Tong University, Shanghai 200240, People’s Republic of China
3 IFSA Collaborative Innovation Center, and Tsung-Dao Lee Institute, Shanghai Jiao Tong University, Shanghai 200240, People’s Republic of China

Received 2018 February 9; revised 2018 April 4; accepted 2018 April 4; published 2018 May 3

Abstract

We propose a novel method to constrain the missing fraction of galaxies using galaxy clustering measurements in the galaxy conditional stellar mass function (CSMF) framework, which is applicable to surveys that suffer significantly from sample selection effects. The clustering measurements, which are not sensitive to the random sampling (missing fraction) of galaxies, are widely used to constrain the stellar–halo mass relation (SHMR). By incorporating a missing fraction (incompleteness) component into the CSMF model (ICSMF), we use the incomplete stellar mass function and galaxy clustering to simultaneously constrain the missing fractions and the SHMRs. Tests based on mock galaxy catalogs with a few typical missing fraction models show that this method can accurately recover the missing fraction and the galaxy SHMR, hence providing us with reliable measurements of the galaxy stellar mass functions. We then apply it to the Baryon Oscillation Spectroscopic Survey (BOSS) over the redshift range of 0.1 < z < 0.8 for galaxies of $M_*>10^{11} M_\odot$. We find that the sample completeness for BOSS is over 80% at z < 0.6 but decreases at higher redshifts to about 30%. After taking these completeness factors into account, we provide accurate measurements of the stellar mass functions for galaxies with $10^{11} M_\odot < M_*> < 10^{12} M_\odot$, as well as the SHMRs, over the redshift range 0.1 < z < 0.8 in this largest galaxy redshift survey.

Key words: cosmology: observations – cosmology: theory – galaxies: distances and redshifts – galaxies: halos – galaxies: statistics – large-scale structure of universe

1. Introduction

The connection between the galaxy properties and those of the dark matter has been investigated in depth in the past decades for the local and high-redshift galaxies (see, e.g., Norberg et al. 2001; Zehavi et al. 2002, 2005, 2011; Yang et al. 2003, 2004, 2007; Zheng et al. 2007; Moster et al. 2010, 2013; Coupon et al. 2012; Leauthaud et al. 2012; Guo et al. 2014; McCracken et al. 2015). The galaxy stellar–halo mass relation (SHMR), in particular, provides important constraints to the galaxy formation and evolution models (see, e.g., Yang et al. 2009, 2012; Moster et al. 2010, 2013; Behroozi et al. 2013a; Beutler et al. 2013; Reddick et al. 2013; Zu & Mandelbaum 2015, 2016; Lin et al. 2016; Wang et al. 2018), as it probes the joint evolution of the galaxy and halo mass growth histories and is directly related to the cosmic star formation histories (see, e.g., Behroozi et al. 2013a; Yang et al. 2013).

There are multiple ways of determining the halo masses for galaxies of different stellar masses. The common methods include the halo occupation distribution (HOD; see, e.g., Jing et al. 1998; Peacock & Smith 2000; Seljak 2000; Scoccimarro et al. 2001; Berlind & Weinberg 2002; Berlind et al. 2003; Zehavi et al. 2005; Zheng et al. 2005, 2007, 2009; Guo et al. 2014, 2015; McCracken et al. 2015) and conditional stellar mass function (CSMF; see, e.g., van den Bosch et al. 2008, 2013; Yang et al. 2009, 2012, 2017a; More 2012; Cacciato et al. 2013; More et al. 2013; Reddick et al. 2013) modeling of the galaxy clustering measurements, the subhalo abundance matching models (see, e.g., Rodríguez-Puebla et al. 2012, 2017; Behroozi et al. 2013a; Moster et al. 2013; Guo et al. 2016), and the direct weak gravitational lensing measurements (see, e.g., Mandelbaum et al. 2006; Miyatake et al. 2015; Zu & Mandelbaum 2015). The central galaxy SHMR has been extensively studied with these methods for galaxy samples at different redshifts and is found to follow a broken power-law relation with a steep slope for low-mass halos of $M < 10^{12} M_\odot$, and becoming much flatter for more massive halos (Yang et al. 2009, 2012; Moster et al. 2010, 2013; Wang & Jing 2010). Within all these probes, a crucial measurement is the galaxy stellar mass function (SMF).

Thanks to the large-scale galaxy redshift surveys, e.g., the 2dF Galaxy Redshift Survey (2dFGRS; Colless 1999) and the Sloan Digital Sky Survey (SDSS; York et al. 2000), at the low redshifts of $z < 0.2$, we are able to accurately measure the galaxy stellar mass function (see, e.g., Bell et al. 2003; Li & White 2009; Bernardi et al. 2010, 2013) and the SHMR (see, e.g., Wang et al. 2007; Yang et al. 2009, 2012; Moster et al. 2010, 2013; Behroozi et al. 2013a; Rodríguez-Puebla et al. 2017). At higher redshifts, the measurements of the galaxy SMFs in the literature have much larger uncertainties compared to the low-redshift ones, because the galaxy SMFs at high redshifts are mostly derived from deep photometric surveys covering small sky area, where the sample variance effect would dominate the error budget (Davidzon et al. 2013, 2017; Ilbert et al. 2013; Maraston et al. 2013; Moustakas et al. 2013; Muzzin et al. 2013; Tomczak et al. 2014; Santini et al. 2015). For example, as summarized in Table 1 of Rodríguez-Puebla et al. (2017) see also Table 3 of Behroozi et al. 2013a, the galaxy SMF at 0.2 < z < 1 is measured with the largest volume in Moustakas et al. (2013)
using the PRism MUlti-object Survey (PRIMUS; Coil et al. 2011), covering an area of 9 deg$^2$.

As constraints to the galaxy SHMR are generally obtained from fitting the galaxy SMFs with or without the spatial clustering measurements, models of the SHMR for higher-redshift galaxies thus have significant differences among different studies, although they agree with each other within errors for low-redshift galaxies (see, e.g., Shankar et al. 2014). The difference is more significant for massive galaxies of $M_*>10^{11}M_\odot$. For example, the SHMR model of Yang et al. (2012) would predict a high-mass-end slope of $M_*\propto M_z^{0.2575}$ for galaxies at $z=0.5$, while the model of Behroozi et al. (2013a) would have $M_*\propto M_z^{0.265}$. However, the total number densities of galaxies with $M_*>10^{11}M_\odot$ for these two models only differ by about 20%, as the halo mass function is decreasing very fast toward the high-mass end. Therefore, discriminating the different SHMR models would require accurate measurements of the galaxy SMF at the massive end, which can only be achieved with wide-area galaxy surveys. Moreover, accurate clustering measurements to constrain the SHMR models would also require a large sample volume.

Currently, the largest galaxy redshift survey at $0.2<z<0.8$ is the SDSS-III Baryon Oscillation Spectroscopic Survey (BOSS; Dawson et al. 2013). The latest data release 12 of the BOSS galaxy sample covers an area of 10,252 deg$^2$ (Reid et al. 2016), which is about 1100 times larger than that of PRIMUS. Therefore, the dominating errors on the BOSS galaxy SMF measurements come from the systematic errors on the galaxy stellar mass measurements, rather than the sample variance effect, which only has a minor contribution. The BOSS galaxy sample would potentially provide the most accurate galaxy SMF and SHMR measurements at these intermediate redshifts. However, the main science drive for BOSS is to measure the baryonic acoustic oscillation signals in order to constrain the cosmological parameters (see, e.g., Alam et al. 2017), so the galaxy-targeting strategy is to select intermediate-redshift, luminous galaxies that cover a large enough volume. Due to the complicated selection criteria of both the apparent magnitude and color, the resulting galaxy sample is, however, neither volume limited nor stellar mass complete. Thus, the measured galaxy SMFs in BOSS cannot be directly compared with those from other studies of volume-limited samples. Furthermore, the BOSS galaxy sample is also not a homogeneous sample of luminous red galaxies (LRGs) at the intermediate redshifts, but also purposely includes a significant fraction of blue galaxies (Maraston et al. 2013).

Efforts have been made to estimate the stellar mass completeness for the BOSS galaxies, using either galaxies selected with wider color cuts (Tinker et al. 2017) or deeper imaging observations of those massive galaxies in the Stripe 82 region (Leauthaud et al. 2016; Saito et al. 2016), with the conclusion that the high-redshift BOSS galaxies are significantly affected by the mass incompleteness. In fact, the mass incompleteness caused by the complicated target selection is not just the concern of BOSS; it has become a common issue for many large-scale galaxy redshift surveys targeting high-redshift objects. For example, the ongoing SDSS-IV extended Baryon Oscillation Spectroscopic Survey (eBOSS; Dawson et al. 2016), targeting LRGs (Prakash et al. 2016) and emission-line galaxies (Raichoor et al. 2017) in the redshift range of $0.7<z<1.1$, adopts various apparent magnitude and color cuts to select objects for further spectroscopic observations. Similar issues also happen for many of the next-generation galaxy redshift surveys, such as the Dark Energy Spectroscopic Instrument (DESI; DESI Collaboration et al. 2016) and Prime Focus Spectrograph (Takada et al. 2014).

Naturally, it is more reliable to constrain the missing fraction of galaxies caused by the complicated target selections without resorting to external measurements from other surveys. Montero-Dorta et al. (2016) presented a forward-modeling technique to quantify the completeness of BOSS galaxies in the red sequence at $z\sim0.55$. By matching the observed color–magnitude distributions with reasonable analytical parametric models convolved with the photometric errors and selection effects, it is possible to derive the intrinsic color–magnitude distribution and therefore estimate the completeness as a function of magnitude. In this paper, we propose a novel method to constrain the completeness by making use of a clustering measurement property—the galaxy clustering is insensitive to the random sampling (missing fraction) of galaxies. On the other hand, within the CSMF framework, accurate clustering measurements can be used to constrain the SHMR. We can therefore incorporate a missing fraction component in the SHMR, so that both the SHMR and the missing fraction of galaxies can be simultaneously constrained. We used mock galaxy catalogs with a few typical missing fraction models to demonstrate the reliability of such a method and then apply it to the BOSS galaxy sample, providing so far the most accurate measurements of the galaxy SMFs and SHMRs at the massive end in the redshift range of $0.1<z<0.8$.

The structure of the paper is as follows. In Section 2, we describe the galaxy samples and the simulation used in the modeling. We introduce our modeling method in Section 3 and test it in Section 4. We present the modeling results for the BOSS galaxies in Section 5 and discuss our models in Section 6. We summarize our results in Section 7. Throughout this paper, we assume a spatially flat $\Lambda$ cold dark matter cosmology, with $\Omega_m=0.307$, $h=0.678$, $\Omega_b=0.048$, and $\sigma_8=0.823$, consistent with the constraints from Planck (Planck Collaboration 2014) and with the parameters used in the simulation adopted for our modeling (see Section 2). For the stellar mass estimates, we assume a universal Chabrier (2003) initial mass function (IMF), the stellar population synthesis (SPS) model of Bruzual & Charlot (2003), and the time-dependent dust attenuation model of Charlot & Fall (2000). All masses are in units of $M_\odot$.

2. Data

2.1. BOSS Galaxy Sample

We use the Data Release 12 of the BOSS galaxy sample, with redshifts for 1,372,737 galaxies over an area of 10,252 deg$^2$ (Reid et al. 2016). The detailed descriptions of the survey can be found in Eisenstein et al. (2011) and Dawson et al. (2013). The BOSS galaxy sample is formally divided into two subsamples with different target selections focusing on galaxies at low and high redshifts. The target selections are based on the following set of combinations of model magnitudes:

$$c_{\parallel} = 0.7(g-r) + 1.2(r-i-0.18), \quad (1)$$

$$c_i = (r-i) - (g-r)/4 - 0.18, \quad (2)$$

Guo, Yang, & Lu

The Astrophysical Journal, 858:30 (17pp), 2018 May 1
where the \( g - r \) and \( r - i \) colors are based on the model magnitudes.

The low-redshift sample, denoted as LOWZ, is an extension of the SDSS-II LRG sample (Eisenstein et al. 2001) to fainter magnitudes in \( 0.15 < z < 0.43 \), with the selection cuts of

\[
|c_i| < 0.2, \quad r < 13.6 + c_l/0.3, \quad 16 < r < 19.6, \tag{4}
\]

where \( r \) is based on the \texttt{cmodel} magnitude. As clearly shown in Figure 11 of Eisenstein et al. (2001), most of the galaxies selected with the above cuts in LOWZ would be LRGs.

The higher-redshift sample, denoted as CMASS, targets galaxies that roughly follow a constant stellar mass cut in the redshift range of \( 0.43 < z < 0.7 \) (Maraston et al. 2013). The corresponding selection cuts are

\[
|d_\parallel| > 0.55, \quad i < 19.86 + 1.6(d_\parallel - 0.8), \quad 17.5 < i < 19.9, \tag{5}
\]

where \( i \) is the \textit{i}-band \texttt{cmodel} magnitude.

To increase the stellar mass completeness at different redshifts, we use the combined sample of LOWZ and CMASS. The redshift and angular distribution of the combined BOSS galaxy sample are presented in Reid et al. (2016, their Figures 8 and 11). The sky coverage of the LOWZ sample in the northern Galactic cap (NGC) is slightly smaller than that of CMASS, due to the removal of data observed in the first 9 months that have an incorrect star–galaxy separation cut applied. At \( z < 0.4 \), as there is only a small fraction of CMASS galaxies in this redshift range, we include all LOWZ galaxies and those CMASS galaxies that fall into the LOWZ geometry mask (covering about 9000 deg\(^2\)). At \( z > 0.4 \), the different angular distributions of LOWZ and CMASS are taken into account by combining the random catalogs for the two samples when measuring the spatial clustering.

As we have a combined sample covering a large redshift range of \( 0.1 < z < 0.8 \), we divide the galaxies into different redshift intervals with a bin size of \( \Delta z = 0.1 \) to study the evolution of these massive galaxies. In total, we have seven redshift bins, with the detailed information displayed in Table 1, where the total number of galaxies, \( N_{\text{tot}} \), the average sample stellar mass with the corresponding scatter, \( \langle \log M_\ast \rangle \), and the mean number density, \( n_\ast \), are shown for each sample. As the galaxy SMF and SHMR have been studied extensively in \( 0.1 < z < 0.2 \) with the SDSS DR7 Main galaxy sample, in this paper we focus on the measurements in \( 0.2 < z < 0.8 \) and use the low-redshift measurements as a consistency check with the literature.

![Figure 1](https://wwwmpa.mpa-garching.mpg.de/SDSS/DR7/)

The galaxy stellar mass used in this paper is estimated in Chen et al. (2012) by fitting the galaxy spectra over the rest-frame wavelength range of 3700–5500 Å with the principal component analysis (PCA) method. We use the stellar mass obtained by applying the SPS models of Maraston & Strömbäck (2011) with the IMF of Kroupa (2001) and the dust attenuation model of Charlot & Fall (2000). The total galaxy stellar mass is obtained by applying the mass-to-light ratio within the fiber aperture to the whole galaxy. We refer the readers to Chen et al. (2012) for details.

At low redshifts, the BOSS sample (especially LOWZ) has a significant overlap with the SDSS DR7 Main galaxy sample (Abazajian et al. 2009), where the galaxy stellar masses have been derived in the MPA-JHU catalog following the method of Kauffmann et al. (2003) by applying the SPS model of Bruzual & Charlot (2003). In order to compare with the literature for low-redshift measurements of the galaxy SHMR and SMFs, we cross-match the two galaxy samples and show the comparisons of the stellar masses in the different catalogs in Figure 1. The stellar masses with the PCA method are systematically overestimated by 0.105 dex compared to the MPA-JHU stellar masses, due to the truncated star formation histories and a smaller fraction of galaxies with recent bursts (see Figures 12 and 13 of Chen et al. 2012, and discussions therein). We also apply the standard correction of a constant shift of \(-0.05\) dex (Bernardi et al. 2010) to convert from a Kroupa (2001) IMF to that of Chabrier (2003). Therefore, in this paper we reduce the galaxy stellar masses in Chen et al. (2012) by 0.155 dex to be consistent with the literature.

Table 1: \textbf{Samples of Different Redshift Bins}.

| Redshift Range | \( N_{\text{tot}} \) | \( \langle \log M_\ast / M_\odot \rangle \) | \( n_\ast / h^3 \text{Mpc}^{-3} \) |
|----------------|-----------------|-----------------|-----------------|
| \( 0.1 < z < 0.2 \) | 84404 | 11.02 ± 0.12 | 5.71 × 10^{-4} |
| \( 0.2 < z < 0.3 \) | 117822 | 11.10 ± 0.12 | 3.25 × 10^{-4} |
| \( 0.3 < z < 0.4 \) | 179726 | 11.30 ± 0.12 | 2.84 × 10^{-4} |
| \( 0.4 < z < 0.5 \) | 294291 | 11.31 ± 0.14 | 2.80 × 10^{-4} |
| \( 0.5 < z < 0.6 \) | 398565 | 11.28 ± 0.15 | 2.85 × 10^{-4} |
| \( 0.6 < z < 0.7 \) | 172251 | 11.37 ± 0.16 | 0.99 × 10^{-4} |
| \( 0.7 < z < 0.8 \) | 30954 | 11.64 ± 0.69 | 0.15 × 10^{-4} |

\[
d_\parallel = (r - i) - (g - r)/8, \tag{3}
\]
SMF in a broad redshift range, we only focus on the massive galaxies with \( M_\ast > 10^{11} M_\odot \) in this paper. As will be demonstrated in the following sections, although red and blue galaxies could have different selection functions, the derived total galaxy SMFs and the SHMRs using the whole BOSS sample are not affected, even if we only select homogeneous red galaxy samples in our measurements. Because the red galaxies dominate the contribution to the SMF for \( M_\ast > 10^{11} M_\odot \) (see, e.g., Figure 10 of Moster et al. 2013), the small fraction of blue galaxies in BOSS would not have any significant effect on the SHMR for the whole galaxy sample, even though red and blue galaxies tend to have slightly different SHMRs.

### 2.2. Dark Matter Simulation

In order to evaluate our method in constraining the missing fraction of galaxies, we use the dark matter halos extracted from the BigMultidark Planck simulation (BigMDPL\(^5\); Klypin et al. 2016), with the cosmological parameters of \( \Omega_m = 0.307, \Omega_b = 0.048, h = 0.678, n_s = 0.96, \) and \( \sigma_8 = 0.823. \) The simulation has a box size of 2.5 \( h^{-1} \) Gpc and a mass resolution of \( 2.4 \times 10^{10} h^{-1} M_\odot. \) It has been used to construct mock galaxy catalogs for the CMASS sample (Rodríguez-Torres et al. 2016). The dark matter halos and subhalos in the simulation are identified with the ROCKSTAR phase-space halo finder (Behroozi et al. 2013b).

We use seven different redshift outputs of \( z = 0.152, 0.246, 0.347, 0.453, 0.547, 0.655, \) and 0.759 from BigMDPL, roughly corresponding to the median redshifts of the seven BOSS galaxy samples in the different redshift bins. The galaxy clustering and stellar mass function measurements from the different redshift bins are modeled separately using the corresponding halo catalogs in order to constrain their evolution.

### 3. Method

In this paper, we incorporate an incompleteness component into the CSMF (hereafter ICSMF) framework (Yang et al. 2012) to predict the accurate galaxy total SMF and SHMR from incomplete galaxy samples such as those in BOSS.

---

\(^5\) https://www.cosmosim.org/cms/simulations/bigmpl/
would be underestimated when using and SMFs to is the CSMF for the satellite galaxies. Weñrñrñrñ

true. We have ignored the possible evolution of CSMF between satellite galaxies, can be estimated as (Yang et al. 2012)

\[ \langle N_c(M) \rangle = \int_{M_{\star,1}}^{M_{\star,2}} \frac{dM_s}{M} \frac{d\Phi_c(M_s|M) c_1(M_s)}{dM_s} \] (13)

\[ \langle N_s(M) \rangle = \int_{M_{\star,1}}^{M_{\star,2}} \frac{dM_s}{M} \frac{d\Phi_s(M_s|M) c_\Pi(M_s)}{dM_s}, \] (14)

where \( \Phi_c(M_s|M) \) is the CSMF for the satellite galaxies. We assume that satellite galaxies have the same CSMF as the centrals when they are distinct halos at the last accretion epoch, as widely used in the subhalo abundance matching algorithm (see Yang et al. 2017b, and references therein). Then the total satellite galaxy distribution in halos of mass \( M \) can be described as

\[ \phi_s(M_s|M) = \int dM_{\text{acc}} \phi_s(M_{\text{acc}}|M) n_s(M_{\text{acc}}|M), \] (15)

where \( M_{\text{acc}} \) is the subhalo mass at the last accretion epoch and \( n_s(M_{\text{acc}}|M) \) is the corresponding subhalo mass function in host halos of mass \( M \).

Note that within this framework, we have assumed the same SHMR for central and satellite galaxies, which is not exactly true. We have ignored the possible evolution of CSMF between the accretion epoch and the redshift of interest. We also ignore the growth and tidal stripping effects for the satellite galaxy stellar mass after accretion. As found by recent studies of Guo et al. (2016) and Yang et al. (2017b), the central and satellite galaxies tend to have somewhat different galaxy–halo relations. As shown in Figure 2 of Guo et al. (2016), the projected 2PCF \( w_p(r_p) \) would be underestimated when using \( M_{\text{acc}} \) as the subhalo mass by assuming the same SHMR as the central galaxies. The agreement with the observed \( w_p(r_p) \) can be improved by adopting different SHMRs for central and satellite galaxies (Figure 6 in Guo et al. 2016).

In general, one may assume different SHMRs for central and satellite galaxies and use the observed \( w_p(r_p) \) and SMFs to make constraints on both of them. However, assuming different SHMRs for satellite galaxies will introduce another four free parameters \( (M_{\star,0}, M_1, \alpha, \text{ and } \beta) \). Since the SHMRs of the central and satellite galaxies are not completely independent, the observation would be overfitted by adopting completely independent SHMRs for central and satellite galaxies. A more consistent and reasonable model is to include the redshift evolution in the central galaxy SHMR and the stellar mass evolution of satellite galaxies after accretion, such as those carried out in Yang et al. (2012). We will defer such a sophisticated model of galaxy stellar mass evolution based on BOSS observation to a subsequent paper. Although we are not modeling the evolutionary trajectories of satellite galaxies, these effects only have minimal influence on the total SMF, as the majority of galaxies in these massive galaxy samples are central galaxies (White et al. 2011; Parejko et al. 2013; Guo et al. 2014). In addition, as these effects might compromise at certain levels, the SHMRs obtained from central and satellite galaxies are still quite similar (Yang et al. 2017b).

In practice, since we are using halos/subhalos from simulations for our study, the satellite occupation function in subhalos of mass \( M_{\text{acc}} \) can be directly estimated as

\[ \langle N_s(M_{\text{acc}}) \rangle = \int_{M_{\star,1}}^{M_{\star,2}} \frac{dM_s}{M} \frac{d\phi_s(M_{\text{acc}}|M) c_\Pi(M_s)}{dM_s} dM_s, \] (16)

With these HOD definitions, we can then split the galaxy sample into different stellar mass ranges and calculate the clustering measurements for these subsamples to constrain the ICSMF model parameters as in Yang et al. (2012). We note that the effect of \( \sigma_c \) on the galaxy clustering is automatically taken into account in Equations (13) and (16).

3.1.2. Two-point Correlation Function and Stellar Mass Function

With the ICSMF for central and satellite galaxies, we measure the galaxy 2PCF by applying the simulation-based method of Zheng & Guo (2016) with the halo catalogs in the BigMDPL simulations, which overcomes the difficulty of modeling the halo exclusion effect, the scale-dependent halo bias, and the residual redshift-space distortion (RSD) effect. By tabulating the clustering measurements of the halos and subhalos, this method is equivalent to, but significantly more efficient than, directly populating halos and subhalos in the simulations with galaxies using the occupation function, \( \langle N_c(M) \rangle \) and \( \langle N_s(M_{\text{acc}}) \rangle \). In detail, The 3D galaxy 2PCF \( \xi(r) \) is measured in the BigMDPL simulations as follows (Zheng & Guo 2016):

\[ \xi(r) = \sum_{i,j} \frac{n_h(M_i) n_h(M_j) \langle N_c(M_i) \rangle \langle N_c(M_j) \rangle \xi_{\text{hh}}(r; M_i, M_j)}{n_g^2} + \sum_{i,j} \frac{n_h(M_i) n_s(M_j) \langle N_s(M_i) \rangle \langle N_s(M_{\text{acc},j}) \rangle \xi_{\text{hs}}(r; M_i, M_{\text{acc},j})}{n_g^2} + \sum_{i,j} \frac{n_s(M_i) n_{\text{acc},j} \langle N_{\text{acc},i} \rangle \langle N_s(M_{\text{acc},j}) \rangle \xi_{\text{ss}}(r; M_{\text{acc},i}, M_{\text{acc},j})}{n_g^2}, \] (17)

where \( n_h(M) \) and \( n_s(M_{\text{acc}}) \) are the halo mass function and subhalo mass function at the last accretion epoch, respectively. The predicted galaxy number density \( n_g \) can be calculated as

\[ n_g = \sum_i \left( \langle N_c(M_i) \rangle n_h(M_i) + \langle N_s(M_{\text{acc},i}) \rangle n_s(M_{\text{acc},i}) \right). \] (18)

The 3D 2PCFs \( \xi_{\text{hh}}(r; M_i, M_j) \), \( \xi_{\text{hs}}(r; M_i, M_{\text{acc},j}) \), and \( \xi_{\text{ss}}(r; M_{\text{acc},i}, M_{\text{acc},j}) \) are the tabulated 2PCFs of the halo–halo, halo–subhalo, and subhalo–subhalo pairs between the different mass bins from the simulation. With the help of halo and subhalo catalogs, we can accurately measure \( \xi(r; M_i, M_j) \)
directly from the simulations, rather than applying any theoretical models for the halo bias $b_h(M)$ (Mo et al. 1996).

To reduce the effect of RSD caused by galaxy peculiar velocities, we focus on the measurements of the projected 2PCF $w_p(r_p)$ (Davis & Peebles 1983), defined as

$$w_p(r_p) = 2 \int_{0}^{r_{\text{max}}} \xi(r_p, r_c) dr_c,$$

(19)

where $r_c$ and $r_p$ are the separations of galaxy pairs along and perpendicular to the line of sight (LOS). $r_{\text{max}}$ is the maximum LOS distance to achieve the best signal-to-noise ratio.

The observed (incomplete) galaxy SMF of the BOSS sample can be predicted as

$$\Phi(M) = \int dM \Phi_\xi(M|\mathcal{M}) c_1(M) n_0(M) + \int dM_{\text{acc}} \Phi_\xi(M_{\text{acc}}|\mathcal{M}) c_\Pi(M_{\text{acc}}) n_0(M_{\text{acc}}).$$

(20)

In summary, we have four free parameters ($M_\text{c,0}$, $M_\text{i,0}$, $\sigma_0$, and $\beta$) for the SHMR (Equation (11)) and another six parameters ($f_0$, $M_{\text{c,0,min}}$, $M_{\text{i,0,min}}$, and $\sigma_0$) for the incompleteness component (Equation (12)). The predictions of $\Phi(M_\text{c})$ and $w_p(r_p)$ can then be compared with those measured in the observed galaxy sample in order to constrain the best-fitting model parameters.

With the best-fitting model parameters, we can infer the intrinsic galaxy total SMF $\tilde{\Phi}(M_\text{c})$ at each redshift interval as

$$\tilde{\Phi}(M_\text{c}) = \int dM \Phi_\xi(M_\text{c}|\mathcal{M}) n_0(M) + \int dM_{\text{acc}} \Phi_\xi(M_{\text{acc}}|\mathcal{M}) n_0(M_{\text{acc}}).$$

(21)

3.2. Observational Measurements

With the above-listed models for the galaxy projected 2PCFs and SMFs, we can use the related observational measurements to constrain the model parameters.

We measure the projected 2PCF $w_p(r_p)$ for BOSS galaxies through the Landy–Szalay estimator (Landy & Szalay 1993). In practice, we integrate $\xi(r_p, r_c)$ in Equation (19) to a maximum LOS distance of $r_{\text{max}} = 100 h^{-1}$ Mpc to achieve the best signal-to-noise ratio, and the same value is adopted in the model calculation. We choose logarithmic $r_p$ bins with a width $\Delta \log r_p = 0.2$ from 1 to 63.1 $h^{-1}$ Mpc and linear $r_c$ bins of width $\Delta r_c = 2 h^{-1}$ Mpc from 0 to 100 $h^{-1}$ Mpc. We measure the projected 2PCFs for the two stellar mass ranges of $10^{10} M_\odot < M_\text{c} < 10^{11.5} M_\odot$ and $10^{11.5} M_\odot < M_\text{c} < 10^{12} M_\odot$.

The observed galaxy SMF is measured for galaxy stellar mass in the range of $10^{10}$ to $10^{12} M_\odot$ with a logarithmic width of $\Delta \log M_\text{c} = 0.1$. Therefore, we have 28 data points in total, with 18 for $w_p(r_p)$ and 10 for $\Phi(M_\text{c})$. We estimate the error covariance matrices for $w_p(r_p)$ and $\Phi(M_\text{c})$ using the jackknife resampling technique with 100 subsamples (Guo et al. 2013, 2014). We note that the cross-covariance between the $w_p(r_p)$ measurements for the two stellar mass bins is also taken into account in the covariance matrix. The jackknife resampling method provides a reasonable way to estimate the sample variance effect (see agreement with the errors estimated from mock catalogs in Appendix B of Guo et al. 2013).

3.3. Model Constraints

In order to fully explore the probability distribution of the model parameters, we apply a Markov Chain Monte Carlo (MCMC) method. The likelihood surface is determined by $\chi^2$, contributed by the observed galaxy stellar mass function $\Phi(M_\text{c})$ and the projected 2PCF $w_p(r_p)$,

$$\chi^2 = \chi^2_{w_p} + \frac{(\Phi - \Phi^s)^2}{\sigma^2_\Phi},$$

(22)

$$\chi^2_{w_p} = (w_p - w_p^s)T C_{w_p}^{-1}(w_p - w_p^s),$$

(23)

where $C_{w_p}$ is the full error covariance matrix of $w_p(r_p)$. The quantity with (without) a superscript “s” is the one from the data (model). The degree of freedom ( dof ) of the model is 18, i.e., $\text{dof} = 28 - 10$.

Here we only adopt the diagonal elements of the covariance matrix of $\Phi$, because the uncertainties from systematic effects of stellar mass measurements are hard to estimate (Mitchell et al. 2013). The contribution of Poisson noise to the observed SMF is added in quadrature to $\sigma^2_\Phi$.

4. Test the Performance of Our Method

4.1. Tests with Mock Catalogs

Before we apply our method to the BOSS galaxies, we perform a validity test on a mock galaxy catalog in the redshift range of $0.1 < z < 0.2$ that has the same geometry as the BOSS sample. We first assign each halo in the BigMDPL simulation a value of galaxy stellar mass according to the following model parameters of Equation (11):

$$\log M_{\text{c,0}} = 10.459, \log M_\text{i} = 10.844, \alpha = 0.309, \beta = 8.077,$$

which are adopted from Yang et al. (2012) as the best-fitting parameters at the simulation output redshift of $z = 0.152$ (see Yang et al. 2012, for more parameters for different sets of cosmologies). We note that satellite galaxies are included in the mock catalogs by applying the SHMR with the subhalo mass at the accretion epoch. Once the dark matter halos and subhalos are populated with galaxies of different stellar masses, we then place a virtual observer at the center of the simulation volume and calculate the R.A., decl., and redshift for each galaxy. The galaxy peculiar velocity is taken into account when calculating the redshift.

We first test our method for the case of the same stellar mass completeness function $c(M_\text{c})$ for central and satellite galaxies, i.e., $c_0(M_\text{c}) = c_\Pi(M_\text{c})$. We constructed three mock catalogs by applying the following three simple selection models on the central and satellite galaxies:

$$c_1(M_\text{c}) = 0.6,$$

(24)

$$c_2(M_\text{c}) = 0.3,$$

(25)

$$c_3(M_\text{c}) = (\log M_\text{c} - 10.5)/2.$$  

(26)

The models $c_1(M_\text{c})$ and $c_2(M_\text{c})$ are random selections with different rates. The model $c_3(M_\text{c})$ is a linear selection with higher rates for more massive galaxies.

We then measure $w_p(r_p)$ and $\Phi(M_\text{c})$ for these mock catalogs and run the MCMC chains to find the best-fitting model parameters. The results are displayed in Figure 3, with blue, red, and magenta symbols (and lines) for the three mock catalogs with $c_1(M_\text{c})$, $c_2(M_\text{c})$, and $c_3(M_\text{c})$, respectively. The
top left panel shows the projected 2PCFs $w_p(r_p)$, with the crosses and circles for the stellar mass ranges of $10^{11} - 10^{11.5} M_\odot$ and $10^{11.5} - 10^{12} M_\odot$, respectively. For clarity, the measurements for the $c_2(M_*)$ and $c_3(M_*)$ models are shifted upward by 1 and 2 dex, respectively. The best-fitting models are shown as the solid lines. Top right: the three different selection functions, with the dotted lines as the input models and the solid lines with shaded regions as the median and 1σ uncertainties of the best-fitting models. Bottom left: comparisons of the stellar–halo mass relations, with the circles for the best-fitting models and the dotted line for the input model. Bottom right: observed and total galaxy SMFs, with the crosses for the observed galaxy SMFs $\Phi(M_*)$ and the circles for the predicted total SMFs $\Phi(M_*)$ from the best-fitting models. The solid lines are the best-fitting models for $\Phi(M_*)$. The black dotted line is for the input model of $\Phi(M_*)$. The measurements for $\Phi(M_*)$ are shifted upward by 0.5 dex for clarity.

Before we apply our ICSMF model to the BOSS observation, it is important to check to what extent the true incompletenesses for central and satellite galaxies can be recovered if they are different. To this end, we input $c_1(M_*) = 0.3$ for central galaxies and $c_2(M_*) = 0.6$ for satellite galaxies and perform the same procedure as above. As shown in Figure 4, both the input completeness functions, the SHMR, and the total galaxy SMF are very well recovered, demonstrating the validity of our ICSMF model to accurately predict the galaxy SMFs from incomplete samples with complicated target selections. In the following sections, we will use separate completeness functions for central and satellite galaxies to model the real BOSS galaxy samples.
method to the BOSS galaxies in this redshift range and compare with the literature.

We show in Figure 5 the modeling results for the $w_p(r_p)$ (top left panel), $c(M_*)$ (top right panel), SHMR (bottom left panel), and total galaxy SMF (bottom right panel) at $0.1 < z < 0.2$ as in Figure 3. The low-redshift sample of BOSS is more than 85% complete for $M_*>10^{11.4}M_\odot$.

In the bottom left panel, we compare the SHMR for our best-fitting model (circles) with the predictions of different theoretical models of Yang et al. (2012), Behroozi et al. (2013a), Moster et al. (2013), and Rodríguez-Puebla et al. (2017) (lines of different colors). The high-mass-end slope of the SHMR is not well constrained in previous literature, where significant differences exist among the models. These differences mainly come from three origins: (1) the stellar mass functions in these studies are somewhat different, (2) the adopted cosmological parameters are different, and (3) the scatter $\sigma_c$ used in these studies is somewhat different as well. Our measurements here agree best with the model of Behroozi et al. (2013a) at this redshift interval.

Our predicted galaxy total SMF from the BOSS sample is shown as the black circles in the bottom right panel. We also compare them with the galaxy SMFs measured in Li & White (2009) using SDSS DR7 and in Moustakas et al. (2013) using the joint sample of SDSS and Galaxy Evolution Explorer (Martin et al. 2005). We use the stellar mass measurements estimated with the $r$-band model magnitude from Li & White (2009) to account for the fact that the typical $r$-band Petrosian magnitudes would result in an underestimation of the galaxy stellar mass (Guo et al. 2010). We add a constant offset of $-0.05$ dex to the stellar mass of the Moustakas et al. (2013) to account for the Flexible Stellar Population Synthesis models used in their stellar mass estimates (see their Figure 19 for comparisons of different SPS models). We find very good agreement between our measurements and those from the literature.

The above two sets of tests demonstrate that our method is robust in recovering the missing fraction of galaxies, as well as providing unbiased measurements of the galaxy SMFs and SHMRs from large-scale surveys with complicated target selections. In the following sections, we will apply our method to the BOSS galaxies in $0.2 < z < 0.8$ to predict the galaxy total SMFs and SHMRs.

5. Results

5.1. ICSMF Model Constraints

We measure the projected 2PCFs and the incomplete SMFs from BOSS observation within a redshift range of $0.2 < z < 0.8$ according to the methods outlined in Section 3.2. Here galaxies are divided into different samples according to the criteria listed in Table 1. These measurements are then used to constrain the ICSMF model parameters using the algorithm outlined in Section 3.3.
We show in Figure 6 the projected 2PCF measurements for the BOSS galaxies at $0.2 < z < 0.8$. The blue and red symbols are for galaxies in the stellar mass ranges of $10^{11} - 10^{11.5} M_{\odot}$ and $10^{11.5} - 10^{12} M_{\odot}$, respectively. Our best-fitting models are shown as the solid lines of different colors. The clustering amplitude of $w(r_p)$ for the more massive sample is about twice that of the lower-mass sample, which is consistent with the galaxy bias measurements of Tinker et al. (2017). As these BOSS galaxies generally live in halos of masses larger than $10^{12} M_{\odot}$, the halo bias is quickly increasing with mass for these massive halos (Mo et al. 1996; Tinker et al. 2005). Our best-fitting models reasonably fit all the observed $w(r_p)$, except for the high-mass sample at $0.2 < z < 0.3$. While the best-fitting $\chi^2$ (41.7) for this sample is still reasonable for a dof of 18, the slight deviation is possibly related to the assumption of the same SHMR for central and satellite galaxies and the constant scatter $\sigma_c$. The current BOSS measurements are not accurate enough to fully constrain those sophisticated models with more freedom.

Figure 7 shows our best-fitting models for the observed galaxy SMFs at different redshifts, with the circles for the measured galaxy SMFs in BOSS and solid lines for our best-fitting models. The agreement between the data and models for galaxy samples at different redshifts is remarkably good, implying that the sample selection functional forms used in this study are suitable.

The best-fitting stellar mass completeness functions for the different galaxy samples are shown in Figure 8, with the red and blue solid lines for central and satellite galaxies, respectively. The shaded area represents the model uncertainties. The BOSS galaxy samples are more than 80% complete for massive galaxies of $M > 10^{11.5} M_{\odot}$ at $z < 0.6$, but the completeness decreases very fast for lower-mass galaxies, causing the galaxy SMF to decrease at the low-mass end in Figure 7. At higher redshifts of $z > 0.7$, even those very massive galaxies are only about 30% complete. For comparison, we also display the completeness estimates from Leauthaud et al. (2016) as the circles in the redshift range of $0.2 < z < 0.7$. As their estimates are for the whole BOSS sample, they are more comparable to the completeness functions of our central galaxies, because the central galaxies dominate the BOSS sample. Our predictions are generally in good agreement with their estimates at all redshifts.

The completeness of satellite galaxies at the massive end, $f_{\text{cl}}$, is not as well constrained as that of central galaxies, as seen from the $1\sigma$ distribution. This is related to the fact that central galaxies dominate the massive end of the SMF (Yang et al. 2009). But at the low-mass end, the completeness of satellite galaxies is quite well constrained and generally larger than that of the central galaxies. It is caused by the $i$-band sliding cut ($i < 19.86 + 1.6(d_{\perp} - 0.8))$ of the BOSS CMASS.
target selections to remove the fainter and bluer galaxies (see, e.g., Figure 1 of Guo et al. 2013), while the low-mass red galaxies have higher satellite fractions (see, e.g., Zehavi et al. 2011; Guo et al. 2014; Saito et al. 2016; Yang et al. 2017b). Therefore, the color selection of the BOSS galaxy samples at a given stellar mass is also taken into account in our separate modeling of the completeness functions for the central and satellite galaxies.

5.2. Central Galaxy SHMR

We show in Figure 9 the best-fitting central galaxy SHMRs for the different redshift samples (circles with errors), in comparison to five other theoretical models from Leauthaud et al. (2012), Yang et al. (2012), Behroozi et al. (2013a), Moster et al. (2013), and Rodríguez-Puebla et al. (2017) (lines of different colors). For fair comparisons, we have corrected the
predicted halo masses for the different cosmologies used in these models. The halos in Leauthaud et al. (2012), Yang et al. (2012), and Moster et al. (2013) are defined as 200 times the background mass density, 180 times the background density, and 200 times the critical density, respectively. We correct these definitions to our definition of the virial halo mass in the ROCKSTAR halo finder using the average offsets in the BigMDPL simulation. At $z > 0.5$, the halo definition of 200 times the background density is very similar to the virial mass definition, with a correction at the level of about 0.01 dex.

Generally, we find a steeper slope of the central galaxy SHMR for these massive galaxies at $z > 0.3$, compared to other models. In these previous models, the galaxy SHMR is generally derived by fitting models to a set of different observations of the galaxy SMFs, clustering measurements, or galaxy weak-lensing measurements at different redshifts.

\[ \text{Figure 8. Stellar mass completeness functions for galaxies at different redshift ranges. The solid lines with the shaded area are the median and 1σ uncertainties of the best-fitting models. The red and blue lines are for the central and satellite galaxies, respectively. The circles with errors are the estimates of the stellar mass completeness of all galaxies at } 0.2 < z < 0.7 \text{ from Leauthaud et al. (2016).} \]

\[ \text{Figure 9. Comparisons of our best-fitting models of central galaxy SHMRs (circles with errors) with those of Leauthaud et al. (2012), Yang et al. (2012), Behroozi et al. (2013a), Moster et al. (2013), and Rodríguez-Puebla et al. (2017) (lines of different colors).} \]
measurements of z neighboring bins. The measurements of Moustakas et al. represent the previous measurements of Pérez-González et al. White (in Rodríguez-Puebla et al. 2009) attenuation model following the various offset values suggested assumptions about the IMF, the SPS model, and the dust redshifts. Therefore, we expect the existence of reasonable consistent estimates of the corresponding halo masses at different estimated galaxy stellar masses, Total galaxy stellar mass functions, Figure 10.

As the high-mass-end galaxy SMF measurements would be significantly affected by the Malmquist bias, the shape of the galaxy SHMR may be somewhat dependent on the assumption of σe. We will discuss the effect of different σe values in Section 6. We show in Table 2 the best-fitting model parameters. We note that the low-mass-end slope α+β is not well constrained by our samples, as we are only using galaxies more massive than $10^{11} M_\odot$. The high-mass-end slope α varies from about 1/4 to 1/2 from $z = 0.1$ to $z = 0.8$.

5.3. Total Galaxy Stellar Mass Functions

As we demonstrated in Figure 4, with the accurate modeling of the observed galaxy SMFs and 2PCFs, we are able to make accurate predictions of the total galaxy SMFs as well. Taking advantage of the large galaxy samples of BOSS, we show with black circles in Figure 10 the total SMFs for galaxies in different redshift bins. For reference, the related values are listed in Table 3. The measurement errors have included the fractional errors of the observed galaxy SMFs in Figure 7,
which are added in quadrature to the model uncertainties from
the MCMC chains. For comparison, we also display the galaxy
SMFs obtained by Pérez-González et al. (2008), Moustakas et al.
(2013), and Leauthaud et al. (2016) in similar redshift ranges.
The measurements of Li & White (2009) at \( z \sim 0.1 \) are shown in
each panel for easy comparison of the SMF evolution. The observed
(corrected) BOSS galaxy SMF \( \Phi(M) \) is displayed as
the dotted line in each panel. The measurements of Pérez-
González et al. (2008) were obtained with the IMF of Salpeter
(1955), the PEGASE SPS model (Fioc & Rocca-Volmerange
1997), and the dust model of Charlot & Fall (2000). As
suggested by Bernardi et al. (2010, their Table 2), we apply a
correction of \( -0.22 \) dex \((-0.25 \) dex for IMF and \( 0.03 \) dex for
the SPS model) to their stellar masses in order to be consistent with
our measurements (see also Rodríguez-Puebla et al. 2017). We
note that Rodríguez-Torres et al. (2016) also presented the expected
total SMF for CMASS galaxies at \( 0.55 < z < 0.65 \) (their Figure 3).
It was constructed by combining the observed CMASS SMF for
\( M_\ast > 2.5 \times 10^{11} \) \( M_\odot \) and the SMF from Guo et al.
(2010) for lower masses. Since we are using the same set of
CMASS galaxies, we do not show it here.

The measurements of Pérez-González et al. (2008; surveyed
area \( \sim 0.184 \) deg\(^2\)), Moustakas et al. (2013; surveyed area
\( \sim 9 \) deg\(^2\)), and Leauthaud et al. (2016; surveyed area
\( \sim 139.4 \) deg\(^2\)) are limited by the small survey volumes and suffer
from significant sample variance effects. From the comparisons,
our SMF measurements are basically consistent with the trend
seen in the previous ones, but we present the most accurate measurements for massive galaxies of \( M_\ast > 10^{11} \) \( M_\odot \). At
\( z > 0.6 \), where significant differences between the SMF measurements appear in the literature, our measurements tend
to agree better with those of Pérez-González et al. (2008). Since
the galaxy SMF measurements are generally used to infer the
galaxy SHMR, it is therefore important to use the accurate galaxy
SMFs to derive the galaxy SHMR at the massive end. Finally, we
note that as the overall completeness for the \( z > 0.7 \) galaxy
sample is very low in BOSS, we expect that future larger surveys
will obtain more accurate measurements of the galaxy total SMFs
at these higher redshifts.

To ease comparisons with the literature, we fit our measurements of the galaxy total SMF with a standard single
Schechter function (Schechter 1976),

\[
\Phi(M_\ast) = (\ln 10)\Phi^\ast \exp \left( -\frac{M_\ast}{M_\ast^\ast} \right) \left( \frac{M_\ast}{M_\ast^\ast} \right)^{1+\alpha^\ast}.
\]  
(27)
are only slightly affected by the existence of blue halo masses are estimated to be around 0.1 redshifts of $z < 0.2$ see, e.g., Tinker et al. 2013. Although galaxies with the same stellar mass but different colors are found to occupy halos of different masses at low redshifts of $z < 0.2$ (see, e.g., More et al. 2011, 2013; Paranjape et al. 2015; Rodríguez-Puebla et al. 2015; Mandelbaum et al. 2016; Zu & Mandelbaum 2016), the differences in the average halo masses are estimated to be around 0.1–0.2 dex at higher redshifts of $z > 0.2$ (see, e.g., the right panel of Figure 7 in Tinker et al. 2013).

In this paper, we are using the whole BOSS galaxy sample to derive the completeness as a function of the stellar mass, which includes the contributions from different populations. Although we estimate the stellar mass incompleteness with the function of Equation (12), there would be possibly remaining effects of the target selections coming from the mix of the different populations of red and blue galaxies, which potentially have different selection functions. As estimated in Masters et al. (2011) and Montero-Dorta et al. (2016), blue galaxies make up about 25% of the CMASS sample, while the majority of LOWZ galaxies are LRGs. Masters et al. (2011) proposed to use the color cut of $g - i > 2.35$ to select the red (elliptical) galaxies in CMASS (see also Maraston et al. 2013). In order to test whether our derived galaxy total SMF and the SHMR will be affected by the existence of blue galaxies, we impose the color cut of $g - i > 2.35$ on the typical BOSS sample at $0.5 < z < 0.6$ to select a roughly homogeneous population of red galaxies. The fraction of blue galaxies removed in this sample is about 21%. Due to the existence of the photometric errors, such a simple color cut does not ensure that we have a purely red galaxy sample, but it still serves as a simple test of the effect of the blue population.

We show in Figure 12 the comparisons between the measurements and models for all galaxies (our fiducial model, shown as black symbols and lines) and those of red galaxies (red symbols and lines). As the red galaxies dominate the SMF at $M_\star > 10^{11} M_\odot$, the observed clustering measurements of $w_\perp(r_p)$ are only slightly affected by the existence of blue galaxies for the lower stellar mass bin, while the derived SHMR and the total galaxy SMF are quite consistent with those obtained from all BOSS galaxies.

The most relevant change for our red galaxy sample is in the stellar mass completeness shown in the top right panel of Figure 12, where the completeness functions of central and satellite galaxies are displayed as the dotted and dashed lines, respectively. Although there seem to be larger decreases of satellite completeness at the massive end, this does not imply that we removed more blue satellite galaxies. Note that at the very massive end, according to comparison between the observed SMFs of all and red galaxies in the bottom right panel, there are almost no blue galaxies at all. Then at the relatively low mass end with $M_\star < 10^{11.3} M_\odot$, both the central and satellite galaxy completeness functions are slightly reduced by a roughly similar amount (or slightly larger for the satellite galaxies) owing to the removal of blue galaxies. However, considering that the satellite fraction is only about 10% (White et al. 2011; Guo et al. 2014) at this mass range, the majority of the removed blue galaxies should still be central galaxies. Indeed, the fractional decrease of the observed SMF at the low-mass end for the red galaxies is consistent with the decrease of amplitude for the central galaxy completeness.

The above test demonstrates that even though we are using the whole BOSS sample to derive the SHMR and the total galaxy SMF for $M_\star > 10^{11} M_\odot$, our results are not affected by the existence of blue galaxies. This is caused by the fact that the red galaxies dominate the massive end of the SMF. As discussed in Tinker et al. (2013), the possible differences between the SHMRs of red and blue galaxies are not large enough to significantly change the clustering for these massive galaxies. However, we note that if one goes to much lower mass galaxies, where red and blue galaxy populations may have quite different relations of stellar to halo mass, one may include a separate component to represent the fractions of red and blue galaxies, respectively. This separation is beyond the scope of the current paper and will be explored in a future work.

Considering the construction of the theoretical model, we assumed a double power-law functional form for the central galaxy SHMR, and the galaxy selection function is characterized...
by Equation (11). A more flexible functional form for the SHMR has been proposed by Behroozi et al. (2010) with a five-parameter model. But as tested in Behroozi et al. (2013a), compared to the four-parameter double power-law model, the additional free parameter mainly helps improve the fits to the SMF at the low-mass end at $M_\ast < 10^9 M_\odot$. Adopting the five-parameter SHMR model has minor effects on our results, as we are focusing on the most massive galaxies. From the goodness of our fits to the clustering measurements and observed SMFs in Figures 6 and 7, we conclude that the functional form for the stellar mass completeness is flexible enough for modeling the completeness of the BOSS galaxies, which has already been shown in Leauthaud et al. (2016) and Montero-Dorta et al. (2016) for the completeness as a function of both stellar mass and magnitude, respectively.

Lastly, we assumed the scatter $\sigma_c$ between the stellar and halo masses, which includes both the contribution from the intrinsic scatter and the statistical errors on the stellar mass estimates, to be a constant value $\sigma_c = 0.173$. This value is set according to the stellar mass distribution of central galaxies in groups or halos of given masses (Yang et al. 2009). We have tested that if we set the amount of the total scatter to be larger or smaller than this value (e.g., with $\sigma_c = 0.2$ and 0.15, respectively), the SHMRs thus constrained at the high-mass end will be somewhat lower or higher. However, the SMFs recovered are quite independent of this $\sigma_c$ value, i.e., roughly consistent within 1$\sigma$ errors. In addition, although we do not have strong constraints on the value of $\sigma_c$, the inferred $\chi^2$ of different redshift samples tend to favor our fiducial model of $\sigma_c = 0.173$ for the BOSS galaxies. Thus, we expect that the total galaxy SMFs obtained in this work are robust. We note that the assumed scatter does not include the uncertainties in the systematic effects in the stellar mass estimates caused by the different IMFs, SPS models, and dust attenuation laws, which can be reasonably corrected for using constant offsets (see, e.g., Moustakas et al. 2013; Muzzin et al. 2013; Rodríguez-Puebla et al. 2017).

7. Conclusions

In this paper, we have introduced an incomplete conditional stellar mass function (ICSMF) model, which is applicable to large-scale galaxy surveys with complicated target selections. By assuming suitable functional forms for the stellar mass completeness function and the galaxy SHMR, we are able to predict the observed galaxy clustering measurements and the incomplete galaxy SMFs, and vice versa, constraining the ICSMF model parameters using these observational measurements. We tested our method using mock galaxy catalogs and

---

**Figure 12.** Similar to Figure 5, but for the comparisons between the measurements and models for all galaxies (our fiducial model, shown as black symbols and lines) and those of red galaxies (red symbols and lines). We show the completeness functions of central and satellite galaxies in the top right panel as the dotted and dashed lines, respectively. The SHMR of the red galaxies is displayed as the red circles in the bottom left panel, while that of all galaxies is shown as the solid line. The observed SMFs of all and red galaxies are shown as the black and red dotted lines in the bottom right panel, respectively. The predicted total SMF of the red galaxies is shown as the circles, and that of all galaxies is represented by the solid line.
then applied it to the BOSS galaxy survey in the redshift range of $0.1 < z < 0.8$. We then predicted the galaxy total SMF measurements and central galaxy SHMRs for $10^{10} M_{\odot} < M_{*} < 10^{12} M_{\odot}$, which are useful for studying the star formation history and evolution of these massive galaxies.

Our main conclusions are summarized as follows:

1. Based on tests using mock galaxy catalogs, we show that the ICSMF model can accurately recover the incompleteness factors, the SHMRs, and the galaxy total stellar mass functions. The incompleteness factors thus obtained are independent and more consistent than the methods of involving external measurements from other surveys, which might introduce additional systematics from different surveys.

2. By applying our ICSMF model to the BOSS galaxy samples, we find that the BOSS galaxies are more than 80% complete at the massive end for $0.1 < z < 0.5$, while for higher redshifts the completeness decreases very fast to around 30% at $z \sim 0.75$.

3. We obtain accurate measurements of the central galaxy SHMRs for the BOSS galaxies with $M_{*} > 10^{10} M_{\odot}$ from $z = 0.1$ to $z = 0.8$. We find that the high-mass-end slope of the SHMR is generally steeper than the previous measurements in the literature, varying from about 1/4 to 1/2 in $0.1 < z < 0.8$.

4. We provide accurate measurements of the total galaxy SMFs for BOSS galaxies within mass range $10^{11} M_{\odot} < M_{*} < 10^{12} M_{\odot}$ and redshift range $0.1 < z < 0.8$, which will provide tighter constraints to the evolution of these massive galaxies.

This work is supported by the National Key Basic Research Program of China (nos. 2015CB857002 and 2015CB857003) and the National Science Foundation of China (nos. 11233005, 11621303, 11655002, and 11773049). H.G. acknowledges the support of the 100 Talents Program of the Chinese Academy of Sciences. This work is also supported by a grant from the Science and Technology Commission of Shanghai Municipality (grant no. 16DZ2260200).

We thank the anonymous referee for the helpful comments that significantly improved the presentation of this paper. We thank H. J. Mo and Zheng Zheng for helpful discussions. We gratefully acknowledge the use of the High Performance Computing Resource in the Core Facility for Advanced Research Computing at Shanghai Astronomical Observatory. We acknowledge the Gauss Centre for Supercomputing e.V. (www.Gauss-centre.eu) and the Partnership for Advanced Supercomputing in Europe (PRACE-CE, www.prace-eu.eu) for funding the MultiDark simulation project by providing computing time on the GCS Supercomputer SuperMUC at Leibniz Supercomputing Centre (LRZ, www.lrz.de).

ORCID iDs
Hong Guo https://orcid.org/0000-0003-4936-8247
Xiaohu Yang https://orcid.org/0000-0003-3997-4606
Yi Lu https://orcid.org/0000-0003-0770-8485

References
Abazajian, K. N., Adelman-McCarthy, J. K., Aguirres, M. A., et al. 2009, ApJS, 182, 543
Alam, S., Ata, M., Bailey, S., et al. 2017, MNRAS, 470, 2617
Behroozi, P. S., Conroy, C., & Wechsler, R. H. 2010, ApJ, 717, 379
