

EMPIRICAL MODELING OF THE REDSHIFT EVOLUTION OF THE [N II]/H α RATIO FOR GALAXY REDSHIFT SURVEYS

ANDREAS L. FAISST¹, DANIEL MASTERS¹, YUN WANG^{1,2}, ALEXANDER MERSON^{1,3}, PETER CAPAK¹, SANGEETA MALHOTRA⁴,
AND JAMES E. RHOADS⁴

¹IPAC, Mail Code 314-6, California Institute of Technology, 1200 East California Boulevard, Pasadena, CA 91125, USA

²Homer L. Dodge Department of Physics & Astronomy, University of Oklahoma, 440 W Brooks Street, Norman, OK 73019, USA

³Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA 91109, USA

⁴School of Earth and Space Exploration, Arizona State University, Tempe, AZ 85287, USA

ABSTRACT

We present an empirical parameterization of the [N II]/H α flux ratio as a function of stellar mass and redshift valid at $0 < z < 2.7$ and $8.5 < \log(M/M_{\odot}) < 11.0$. This description can easily be applied to (i) simulations for modeling [N II] λ 6584 line emission, (ii) deblend [N II] and H α in current low-resolution grism and narrow-band observations to derive intrinsic H α fluxes, and (iii) to reliably forecast the number counts of H α emission-line galaxies for future surveys, such as those planned for *Euclid* and the *Wide Field Infrared Survey Telescope* (*WFIRST*). Our model combines the evolution of the locus on the Baldwin, Phillips & Terlevich (BPT) diagram measured in spectroscopic data out to $z \sim 2.5$ with the strong dependence of [N II]/H α on stellar mass and [O III]/H β observed in local galaxy samples. We find large variations in the [N II]/H α flux ratio at a fixed redshift due to its dependency on stellar mass; hence, the assumption of a constant [N II] flux contamination fraction can lead to a significant under- or overestimate of H α luminosities. Specifically, measurements of the intrinsic H α luminosity function derived from current low-resolution grism spectroscopy assuming a constant 29% contamination of [N II] can be overestimated by factors of ~ 8 at $\log(L) > 43.0$ for galaxies at redshifts $z \sim 1.5$. This has implications for the prediction of H α emitters for *Euclid* and *WFIRST*. We also study the impact of blended H α and [N II] on the accuracy of measured spectroscopic redshifts.

Keywords: galaxies: ISM – galaxies: fundamental parameters – cosmology: observations

1. INTRODUCTION

The H α nebular emission line at rest-frame 6563 Å is the most important feature that will be detected by the near-infrared grisms of the upcoming space-based missions *Euclid* (Laureijs et al. 2011; Vavrek et al. 2016) and *WFIRST* (Dressler et al. 2012; Green et al. 2012; Spergel et al. 2015). By measuring redshifts to tens of millions of H α emitters at $1 \lesssim z \lesssim 2$ with low-resolution grism spectroscopy, these surveys will use baryon acoustic oscillations (BAO, e.g., Blake & Glazebrook 2003; Seo & Eisenstein 2003) and redshift space distortions (RSD, Kaiser 1987) analyses to constrain the expansion history of the universe and the growth of structure. Together, these probes will put strong constraints on the nature of dark energy (Guzzo et al. 2008; Wang 2008).

The H α emission of a galaxy depends principally on its star formation rate (SFR), while the H α equivalent-width (EW) is proportional to the specific SFR

(sSFR = SFR/ M , a proxy for the rate of stellar mass increase). The H α emission is, therefore, not only an important tool to study cosmology but also a direct probe of the statistics of cosmic star formation via the H α luminosity function (LF). The LF, in turn, informs predictions of the number counts of galaxies that will be found in the grism surveys of *Euclid* and *WFIRST*.

Current measurements of H α across large redshift ranges up to $z \sim 2$ and over large area on sky come from blind searches in low-resolution *Hubble Space Telescope* (*HST*) grism observations (Atek et al. 2010; Brammer et al. 2012; van Dokkum et al. 2013). Specifically, the best current constraints on the H α luminosity function over the redshifts of interest for *Euclid* and *WFIRST* ($0.4 < z < 2.5$) come from the HST WFC3 Infrared Spectroscopic Parallel Survey (WISPS, Atek et al. 2010). However, these WFC3 grism spectra do not resolve H α from the neighboring [N II] lines at rest-frame 6548 Å and 6584 Å. The resulting uncertainty in the [N II]/H α ratio (in the following, [N II]/H α \equiv

$[\text{N II}]\lambda 6584/\text{H}\alpha$)¹ for the sources measured by WFC3 translates into significant uncertainty in the derived intrinsic H α LF and therefore has a direct impact on shaping future large surveys. Commonly, a constant $[\text{N II}]$ flux contamination fraction of 29% is assumed for such grism surveys according to the average value measured in the local universe for galaxies with H α EW less than 200 Å (Colbert et al. 2013; Mehta et al. 2015). However, such an assumption can introduce luminosity dependent biases because of the dependence of the $[\text{N II}]/\text{H}\alpha$ ratio on several galaxy parameters including the redshift and stellar mass (see also discussion in Pozzetti et al. 2016).

The dark energy figure-of-merit of both *Euclid* and *WFIRST* is very sensitive to the number density of H α emitting galaxies. Due to the sharp exponential fall-off at the bright end of the H α LF that will be probed by these surveys, an uncertainty of a factor of 3 in the $[\text{N II}]/\text{H}\alpha$ flux ratio, for example due to blending, could translate into an uncertainty of a factor of up to 10 in the number counts of bright H α emitting galaxies in the worst case. Furthermore, similar to the grism surveys, most of the H α lines detected by *Euclid* and some detected by *WFIRST* will be blended with $[\text{N II}]$, degrading the accuracy of its measured H α LF. In addition, incorrect (or no) de-blending of these two lines can result in a systematic offset of the H α line centroid of up to 300 km s⁻¹ (depending on the actual $[\text{N II}]/\text{H}\alpha$ ratio) and therefore directly affect the accuracy of redshift measurements (Section 4.3).

A significant variation (by factors of 3–5) in the $[\text{N II}]/\text{H}\alpha$ ratio is expected from the large parameter space spanned by galaxies on the “Baldwin, Phillips & Terlevich” (BPT, Baldwin et al. 1981) diagram, which has $[\text{N II}]/\text{H}\alpha$ as the abscissa. This variation is known to be linked to galaxy properties such as metallicity, SFR, and nitrogen-to-oxygen (N/O) ratio, and, due to changing galaxy demographics, the population averaged $[\text{N II}]/\text{H}\alpha$ ratio is expected to vary as a function of redshift (e.g., Kewley et al. 2013; Masters et al. 2014; Steidel et al. 2014; Shapley et al. 2015; Masters et al. 2016; Strom et al. 2017; Kashino et al. 2017). Simulations may be used to predict the $[\text{N II}]/\text{H}\alpha$ ratio as a function of galaxy parameters (e.g., Hirschmann et al. 2017); however, large uncertainties can arise due to different ingredients and assumptions. Other studies use the H α EW in local galaxy samples as prior for the $[\text{N II}]/\text{H}\alpha$ ratio (see e.g., Villar et al. 2008; Sobral et al. 2009; Ly et al. 2011; Lee et al. 2012; Sobral et al. 2013).

Here we outline an approach to constrain the expected

¹ We assume $[\text{N II}]\lambda 6548 = \frac{1}{3}[\text{N II}]\lambda 6584$ (Acker et al. 1989), and therefore $[\text{N II}]\lambda\lambda 6548, 6584/\text{H}\alpha = \frac{4}{3}[\text{N II}]/\text{H}\alpha$.

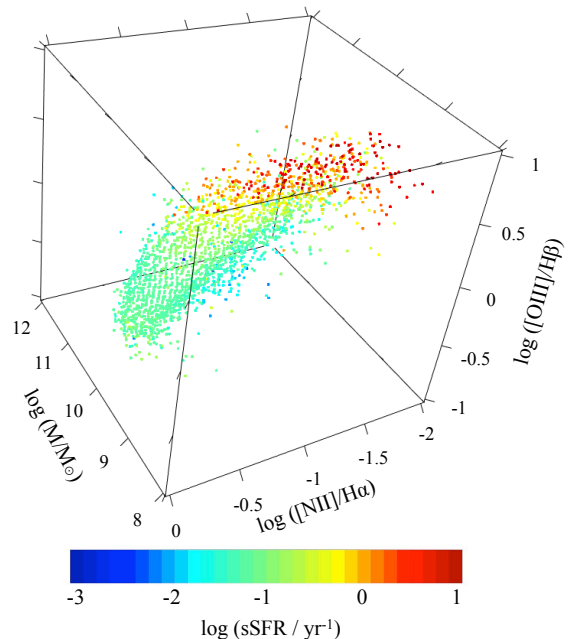


Figure 1. Dependences between $[\text{N II}]/\text{H}\alpha$, $[\text{O III}]/\text{H}\beta$, stellar mass, and sSFR (indicated in color) based on the sample of local galaxies in SDSS. The data is median-binned for visual purposes. An interactive three-dimensional version of this plot built with `plotly` is available in the online Journal (`plotly`: <https://plot.ly/>; a version is also available on the author’s website: http://www.astro.caltech.edu/afaisst/3dplot/plotly_3dplot1.html). The data used to create this figure are available online.

$[\text{N II}]/\text{H}\alpha$ ratios for upcoming surveys using empirical trends in the BPT diagram. Our approach is motivated by the correlations in this diagram illustrated by Masters et al. (2016) and Faisst (2016) (see also Brinchmann et al. 2008). Masters et al. (2016) showed the strong correlation of both stellar mass and SFR density with position on the BPT diagram. Galaxies form a tight locus in the BPT diagram at $z \sim 0$; however, the position of this locus is known to evolve with redshift (Erb et al. 2006; Masters et al. 2014; Steidel et al. 2014; Shapley et al. 2015), which is driven by the changing average sSFR with cosmic time (Daddi et al. 2007; Elbaz et al. 2007; Noeske et al. 2007; Lilly et al. 2013), metallicity (Ly et al. 2016a), ionization parameter (e.g., Nakajima & Ouchi 2014), and electron density of the galaxies.

Here we illustrate that stellar mass and redshift together put strong constraints on the $[\text{N II}]/\text{H}\alpha$ ratio. We parameterize the evolution in the $[\text{N II}]/\text{H}\alpha$ -mass relation with redshift, which allows us to accurately predict the $[\text{N II}]/\text{H}\alpha$ ratio for galaxies with known stellar mass and redshift. The model we present can be used to

- Accurately deblend $[\text{N II}]$ and H α in low-resolution spectroscopic surveys and narrow-band photometric observations,
- Improve the fidelity of the forecasts for the num-

ber counts of $H\alpha$ emitters that will be detected by *Euclid* and *WFIRST*.

First, we outline the idea and motivation of our approach in Section 2. In Section 3, we describe our empirical model that parameterizes the $[N II]/H\alpha$ ratio as a function of stellar mass and redshift. The observational data that is feeding our model is presented in Section 3.1. In the following sections we derive our model and present the final parameterization in Section 3.4. In Section 4, we study in detail the implications of our model on (i) the data interpretation of current surveys (Section 4.1), (ii) the $[N II]$ contamination for *Euclid* and *WFIRST* (Section 4.2), (iii) redshift measurement from blended $H\alpha$ and $[N II]$ lines (Section 4.3), and (iv) the number count predictions for future surveys based on current $H\alpha$ LF determinations (Section 4.4).

Throughout this paper we assume a flat cosmology with $\Omega_{\Lambda,0} = 0.7$, $\Omega_{m,0} = 0.3$, and $h = 0.7$. Furthermore, all stellar masses and star-formation rates (SFR) are scaled to a Chabrier (2003) initial mass function (IMF) and all magnitudes are quoted in AB (Oke 1974).

2. BACKGROUND AND MOTIVATION

We model the $[N II]/H\alpha$ ratio as a function of redshift and stellar mass, two key observational quantities. While other methods to constrain $[N II]$ contamination using optical emission lines could potentially be more accurate (e.g., involving $[O III]$ and $[O II]$, de los Reyes et al. 2015), these cannot be applied to, e.g., narrow-band observations (lacking spectroscopic follow-up) as well as the large samples of future large-area surveys such as with *WFIRST* or *Euclid* due to their relatively low spectroscopic line sensitivity. However, at *WFIRST* and *Euclid* depths², stellar masses will be accurately determined down to at least $\log(M/M_{\odot}) = 9.0 - 9.5$ by covering the 4000 \AA break up to $z = 2$ with $Y-$, $J-$, and $H-$ band imaging (Laureijs et al. 2011; Gehrels et al. 2015). This justifies our approach of using redshift and stellar mass as main quantity to derive the $[N II]/H\alpha$ ratio. Moreover, stellar mass and redshift are well-constrained in the semi-analytical models that are often used to estimate population statistics for $H\alpha$ emitters (e.g., Orsi et al. 2010; Merson et al. 2018). The evolution of the $[N II]/H\alpha$ ratio is empirically constrained both by trends seen in the local SDSS sample as well as by measured $[N II]/H\alpha$ -stellar mass relations out to $z \sim 2$ in the literature. Since the $[N II]/H\alpha$ ratio is a gas-phase metallicity indicator (e.g., Pettini & Pagel 2004), the $[N II]/H\alpha$ -mass relation is effectively the

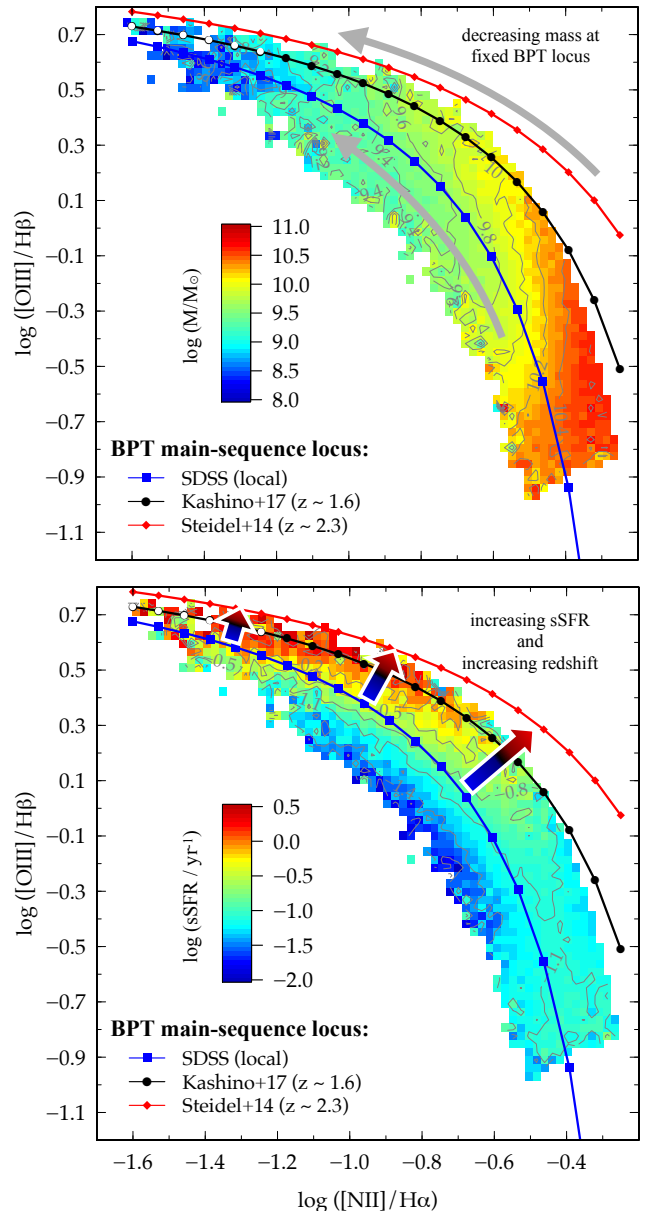


Figure 2. Slices through Figure 1 showing the dependence in stellar mass (*top*) and sSFR (*bottom*) in the local BPT diagram (fit indicated by blue line). High-redshift galaxies occupy sub-regions of the local BPT diagram at higher sSFR (perpendicular to the local BPT locus) as shown by the measured “BPT main-sequence” loci at $z \sim 1.6$ (extrapolated to $\log([N II]/H\alpha) = -1.6$, Kashino et al. 2017) and $z \sim 2.3$ (Steidel et al. 2014; Strom et al. 2017). Stellar mass runs nearly perpendicular to the loci indicating a strong mass dependence. This allows a unique description of emission line ratios as a function of redshift and stellar mass, which is the cornerstone of our model described in the text.

galaxy mass–metallicity (MZ) relation, which is known to evolve with redshift (e.g., Savaglio et al. 2005; Erb et al. 2006; Maiolino et al. 2008; Lilly et al. 2013; Maier et al. 2015; Salim et al. 2015).

Physical galaxy properties such as stellar mass, sSFR, and relative abundance ratios are strongly correlated

² *Euclid* will obtain $Y-$, $J-$, and $H-$ band imaging down to 24AB for a 5σ point source and *WFIRST* will reach ~ 26.5 AB (e.g., Gehrels et al. 2015).

with nebular emission line ratios. This is shown in Figure 1, a 3-dimensional version of the BPT-diagram connecting the line ratios $[\text{N II}]/\text{H}\alpha$ and $[\text{O III}]/\text{H}\beta$ with the more easily accessible observables stellar mass (M) and sSFR (color coded). Figure 2 shows projections of the 3-dimensional figure to better visualize the dependencies with stellar mass (top) and sSFR (bottom). The fitted locus of local galaxies (Kewley & Ellison 2008) is indicated with a blue line.

Trends in the local SDSS data with SFR and stellar mass reflect changes seen in the galaxy population at high redshift. The BPT locus *systematically* shifts with redshift, possibly connected to the overall increase in the global SFR of galaxies. The measured BPT loci of galaxies at $z \sim 1.6$ (Kashino et al. 2017) and $z \sim 2.3$ (Steidel et al. 2014) are indicated in the upper and lower panels of Figure 2, illustrating the pronounced shift in the BPT locus. Indeed, the position on the BPT diagram is very effective to select “high-redshift analogs”, which are a rare sub-sample of local galaxies that resemble high-redshift galaxies in photometric and spectroscopic properties (such as high sSFR or $\text{H}\alpha$ equivalent-width, e.g., Cardamone et al. 2009; Hu et al. 2009; Stanway et al. 2014; Ly et al. 2015; Faisst 2016; Greis et al. 2016; Erb et al. 2016; Ly et al. 2016b).

The locus on the BPT diagram of (roughly) constant sSFR is similar to the locus of the main-sequence on the stellar mass vs. SFR plane. At different redshifts, galaxies populate different distributions of sSFR (e.g., compilations by Lilly et al. 2013; Speagle et al. 2014) and stellar mass ranges, which let the slope and normalization of the main-sequence change across cosmic time. Similarly, galaxies on a “BPT main-sequence” run through a range in $[\text{N II}]/\text{H}\alpha$, and $[\text{O III}]/\text{H}\beta$, and stellar masses (see upper panel of Figure 2) at an sSFR distribution most likely for their redshift. Masters et al. (2016) identified stellar mass and its link to the nitrogen-to-oxygen (N/O) abundance ratio as the main driver for these dependencies, as these quantities strongly vary approximately perpendicular to lines of constant sSFR, i.e., the BPT main-sequence. Once the BPT main-sequence is identified, the $[\text{N II}]/\text{H}\alpha$ (and $[\text{O III}]/\text{H}\beta$) can therefore be uniquely determined from the stellar mass of a galaxy.

The steps for creating our model are therefore as follows:

1. Parameterize the BPT main-sequence as a function of redshift as $O3(N2, z)$,
2. Parameterize the mass dependence on the BPT diagram, i.e., $M(O3, N2)$,
3. Parameterize the $[\text{N II}]/\text{H}\alpha$ ratio as a function of stellar mass and redshift, i.e., $N2(M, z)$, by reversing $M(N2, z)$.

Here, and in the following, we adopt the definitions $N2 = \log([\text{N II}]/\text{H}\alpha)$ and $O3 = \log([\text{O III}]/\text{H}\beta)$. Note, that a parameterization of $N2(M, z)$ can also be derived by directly reversing the observed relation between stellar mass and $[\text{N II}]/\text{H}\alpha$ (which is proportional to the stellar mass – gas-phase metallicity relation). It leads indeed to similar results, however, with a larger uncertainty and ambiguity caused by the $[\text{O III}]/\text{H}\beta$ dependence of the $[\text{N II}]/\text{H}\alpha$ vs. stellar mass relation (see Figure 1). With our approach of parameterizing the entire mass dependence of the BPT diagram, we take this secondary dependence into account, which results in a more accurate and comprehensive model. Moreover this would also allow us to predict the $[\text{O III}]/\text{H}\beta$ ratios in addition to $[\text{N II}]/\text{H}\alpha$ for a given stellar mass and redshift.

3. EMPIRICAL MODEL FOR THE EVOLUTION OF LINE-RATIOS

3.1. Local and high-redshift data

Our model is based on the data of 191 409 local galaxies selected from the Sloan Digital Sky Survey (SDSS, York et al. 2000) using the web-based DR 12 (Alam et al. 2015) query tool³, combined with the observed BPT locus evolution at $0 < z < 2.5$ (Steidel et al. 2014; Masters et al. 2014; Shapley et al. 2015; Kashino et al. 2017).

The stellar masses and emission line measurements in the SDSS catalog are taken from the *Galspec* products provided in the MPA-JHU value added catalog based on the methods of Kauffmann et al. (2003), Brinchmann et al. (2004), and Tremonti et al. (2004). Specifically, the stellar masses are derived from fits to the SDSS *ugriz* total galaxy photometry assuming an exponentially declining star formation history and bursts. In addition, the photometry is corrected for the small contribution of nebular emission using the spectra. The model grids for the fitting are described in Kauffmann et al. (2003) and a Kroupa (2001) IMF is assumed, which has been converted to a Chabrier (2003) IMF. The local galaxies are selected to have a signal-to-noise ratio (S/N) > 5 in the $\text{H}\alpha$ emission line. We note that different S/N thresholds do not have an impact on the subsequent analysis and result. Also, we do not impose a S/N limit on other optical emission lines in order to prevent our sample from any selection bias (e.g., Salim et al. 2014). Galaxies with a significant contribution of an active galactic nucleus (AGN) as suggested by the line ratios on the BPT diagram are removed (about 15%, Kewley et al. 2001; Kauffmann et al. 2003; Brinchmann et al. 2004). The latter introduces an artificial upper boundary on galax-

³ <http://skyserver.sdss.org/dr12/en/tools/search/sql.aspx>

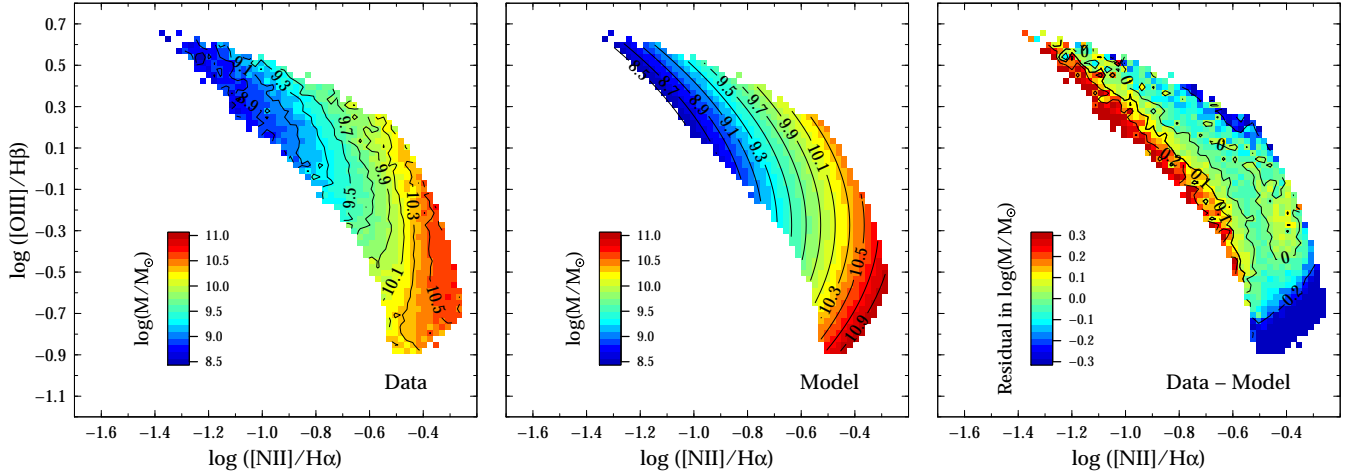


Figure 3. Results of the multi-dimensional fit of $M(O3, N2)$ (Equation 2) to the local galaxy sample from SDSS (*left*: data; *middle*: best-fit model; *right*: residual). For the most parts, our model is able to recover stellar masses to better than 0.1 dex. The best fit parameters are given in Section 3.3.

ies on the BPT diagram at high $[N II]/H\alpha$ and $[O III]/H\beta$ values, but the underlying trends in the distribution are not affected. In addition, we restrict our SDSS sample to $z > 0.05$ in order to minimize the effect of the finite fiber aperture ($3''$ for the SDSS spectra), such that the $3''$ fiber covers at least the central ~ 1.5 kpc of the galaxy. The SQL commands for this selection are provided in Appendix A.

In order to test the redshift dependence of our model, we make use of measurements of $[N II]/H\alpha$, $[O III]/H\beta$, and stellar mass at higher redshifts as presented in the literature. These include galaxies at $z \sim 1.6$ (208 galaxies, Kashino et al. 2017) and $z \sim 2.3$ (in total 360 galaxies, Erb et al. 2006; Genzel et al. 2014; Steidel et al. 2014; Shapley et al. 2015; Strom et al. 2017)⁴. Similar to the SDSS samples, the stellar masses for these samples have been derived from a fit to the total galaxy photometry assuming constant and exponentially declining star formation histories. The photometry includes Spitzer imaging at $> 2 \mu\text{m}$, which covers a similar rest-frame wavelength range as for the local galaxies. In the case of Kashino et al., stellar masses have been converted from a Salpeter (1955) to a Chabrier IMF.

3.2. Parameterization of the BPT main-sequence locus: $O3(N2, z)$

As outlined in Section 2 and indicated in Figure 2, galaxies at a given redshift occupy a defined locus on the BPT diagram (the BPT main-sequence), similar to the $M - \text{SFR}$ main-sequence. This locus shifts towards higher $[O III]/H\beta$ ratios at fixed $[N II]/H\alpha$ (or, alter-

natively, higher $[N II]/H\alpha$ at fixed $[O III]/H\beta$) at high redshifts as shown by many spectroscopic studies (Erb et al. 2006; Masters et al. 2014; Steidel et al. 2014; Shapley et al. 2015; Kashino et al. 2017; Strom et al. 2017). This shift is attributed by Masters et al. (2016) to an increasing ionization parameter and lower metallicity at fixed mass (and thus N/O ratio, which is mostly set by stellar mass as shown by the same study) of the average high-redshift galaxy. The evolution is also captured in compilations of the stellar mass vs. metallicity relation, the latter commonly estimated from these emission line ratios.

In the following, we use the spectroscopic measurements at $z \sim 1.6$ (Kashino et al. 2017)⁵ and $z \sim 2.3$ (Steidel et al. 2014; Strom et al. 2017) to parameterize the BPT main-sequence as a function of redshift (similar to Kewley et al. 2013). We find that a simple shift in $[N II]/H\alpha$ starting from the local locus (Kewley & Ellison 2008) is a good fit to the data (see also Figure 2). We parameterize this shift proportional to cosmic time to obtain the following relation for the redshift dependence of the BPT main-sequence by a least-square fit

$$O3(N2, z) = \frac{0.61}{N2 + \delta - \gamma(1+z)^2} + 1.08, \quad (1)$$

with $\delta = 0.138 \pm 0.005$ and $\gamma = 0.042 \pm 0.005$. Note that this parameterization might not be valid beyond

⁴ Split into 155 galaxies from Steidel et al. (2014), 130 galaxies from Shapley et al. (2015), and 75 galaxies from Genzel et al. (2014).

⁵ Note that the data from this study only covers the range $-1.2 < \log([N II]/H\alpha) < -0.1$ in contrast to the other samples used here. We therefore extrapolate the relation between $[O III]/H\beta$ and $[N II]/H\alpha$ given by that study to reach $\log([N II]/H\alpha) = -1.6$. This extrapolation is well defined as the BPT relation narrows towards high $[O III]/H\beta$ and low $[N II]/H\alpha$. Furthermore, uncertainties in this extrapolation would only affect galaxies stellar masses $\log(M/M_\odot) \lesssim 8.5$, which is below what is considered in the following model.

$z \sim 2.7$ as it cannot be tested with the current data at higher redshifts. Finally, we note that the $z \sim 2.3$ sample by [Steidel et al. \(2014\)](#) is the most comprehensive as it includes published stellar masses, $[\text{N II}]/\text{H}\alpha$, and $[\text{O III}]/\text{H}\beta$ for all individual galaxies, which is crucial for our analysis. The BPT main-sequence has, however, been determined by other studies as well. In particular, we note here the sample at $z \sim 2.3$ by [Shapley et al. \(2015\)](#) based on the MOSFIRE Deep Evolution Field (MOSDEF) survey, which BPT main-sequence is offset by ~ -0.2 dex from the Steidel locus. Using their relation, we obtain for the redshift evolution (Equation 1) parameters $\delta = 0.110 \pm 0.005$ and $\gamma = 0.032 \pm 0.005$. This defines an uncertainty due to sample biases and measurement differences of 0.028 and 0.010 in δ and γ , respectively. In Section 3.5, we discuss in detail the impact of sample biases on the parameterization of the BPT main-sequence and our final model.

3.3. Parameterization of $M(\text{O3}, \text{N2})$

Second, we parameterize the stellar mass distribution on a given BPT main-sequence locus. As argued in [Masters et al. \(2016\)](#) and Section 2, galaxies at high redshifts occupy a distinct region of the local BPT diagram, namely at higher $[\text{O III}]/\text{H}\beta$ for fixed $[\text{N II}]/\text{H}\alpha$. Importantly, they can still be described by the relations found in the SDSS data, although these data become sparse at the location of the high-redshift galaxies and therefore an extrapolation becomes necessary. In the following, we therefore parameterize the relation $M(\text{O3}, \text{N2})$ observed in the local SDSS data to derive the stellar mass distribution by a cut through this plane along a given BPT locus. We find that the following functional form describes best the parabolic shaped (in O3) stellar mass isochrones that are displaced in N2

$$M(\text{O3}, \text{N2}) = A + B(\text{N2} + \alpha) + C(\text{O3} + \beta)^2, \quad (2)$$

where A , B , C , α , and β are determined by a Levenberg–Marquardt algorithm, as part of the `R/minpack.lm` package⁶ ([Elzhov et al. 2016](#)). The fitting to the SDSS data is performed on medians derived from a binning in $[\text{N II}]/\text{H}\alpha$ and $[\text{O III}]/\text{H}\beta$ to increase the S/N in the data (e.g., upper panel in Figure 2). We set the weights for the fitting proportional to the number of galaxies per bin (chosen to be 15 or more). We do not fit galaxies below $\log([\text{O III}]/\text{H}\beta) = -0.6$ that correspond to the most massive galaxies in SDSS and might include post-starburst galaxies and almost quiescent galaxies that could bias our fit. Figure 3 shows the data (*left*) together with the best-fit model (*middle*) and resid-

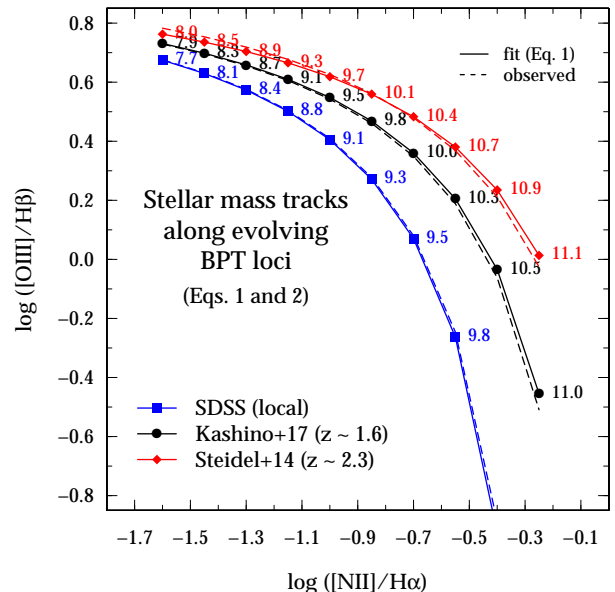


Figure 4. Stellar mass tracks on the BPT diagram for fixed BPT loci at redshifts of $z \sim 0$ ([Kewley & Ellison 2008](#)), $z \sim 1.6$ ([Kashino et al. 2017](#)), and $z \sim 2.3$ ([Steidel et al. 2014](#)). These three data sets are used to derive the shift of the BPT locus as a function of redshift (solid lines, Equation 1). The observed loci at $z \sim 0$, $z \sim 1.6$, and $z \sim 2.3$ are shown as dashed lines for comparison. The stellar mass tracks, $M(\text{N2}, z)$, are parameterized in the final Equation 3, with stellar masses given as $\log(M/M_{\odot})$.

ual (*right*). The best-fit parameters for Equation 2 are $A = 7.689 \pm 0.450$, $B = 3.696 \pm 0.005$, $C = 1.960 \pm 0.015$, $\alpha = 1.126 \pm 0.060$, and $\beta = 0.273 \pm 0.005$. Our best fit describes the stellar masses to better than 0.1 dex for most values of $[\text{O III}]/\text{H}\beta$ and $[\text{N II}]/\text{H}\alpha$. It under-predicts the stellar mass for galaxies on the lower crest of the local BPT main-sequence by up to 0.3 dex. This is mostly an effect of the weighting, which is chosen to minimize the residual at the local BPT main-sequence as well as on the upper crest of the BPT locus where galaxies with higher sSFR (or redshift) are located.

In Figure 4 we show the BPT main-sequence loci at $z \sim 0$, $z \sim 1.6$, and $z \sim 2.3$ from Equation 1 with indicated stellar masses from Equation 2. Note that at a fixed stellar mass the $[\text{N II}]/\text{H}\alpha$ ratio decreases and the $[\text{O III}]/\text{H}\beta$ ratio increases with redshift, which is essentially the stellar mass vs. metallicity relation changing across cosmic time as quoted by many studies in the literature (e.g., [Maiolino et al. 2008](#); [Lilly et al. 2013](#))⁷. In this sense, the above equations also give an empirical parameterization of the evolution of the mass–metallicity

⁶ <https://cran.r-project.org/web/packages/minpack.lm/index.html>

⁷ Recall that the gas-phase metallicity, $12 + \log(\text{O}/\text{H})$, is inversely proportional to $[\text{O III}]/\text{H}\beta$ and proportional to $[\text{N II}]/\text{H}\alpha$ (e.g., [Maiolino et al. 2008](#)).

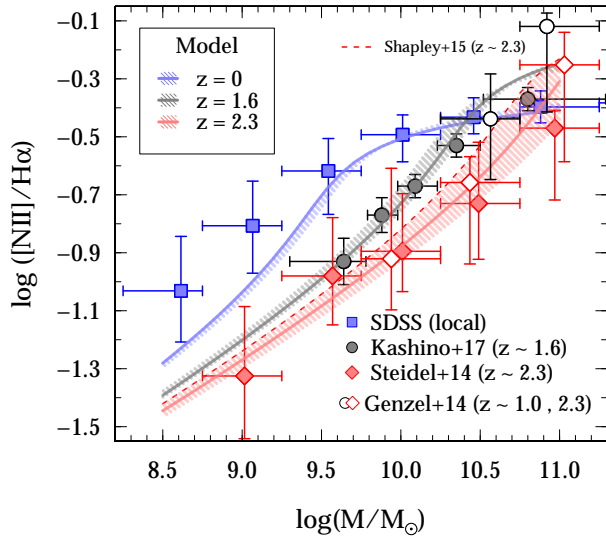


Figure 5. Relation between stellar mass and $[\text{N II}]/\text{H}\alpha$ ratio from observed data at $z \sim 0$ (blue squares), $z \sim 1.6$ (black points), and $z \sim 2.3$ (red diamonds) from SDSS, Kashino et al. (2017), and Steidel et al. (2014), respectively. We also show data from Genzel et al. (2014) at $z \sim 1$ and $z \sim 2.3$ with open symbols for reference (see Section 3.5 for discussion). Our model predictions (Equation 3) at the median redshift of the samples are shown as lines and the hatched regions show the range in model $[\text{N II}]/\text{H}\alpha$ values for the redshift distribution of the observations. Our parameterization reproduces the data within ~ 0.1 dex, well within its scatter. The dashed red line indicates the $z \sim 2.3$ relation from MOSDEF (Shapley et al. 2015) with slightly larger $[\text{N II}]/\text{H}\alpha$ ratios at a given mass compared to the Steidel et al. (2014) relation at the same redshift. This difference can be explained by different selections of the two samples (see Section 3.5).

relation with redshift.

3.4. Final $N2(M, z)$ parameterization

The combination of Equations 1 and 2, allows us to parameterize the $[\text{N II}]/\text{H}\alpha$ ratio as a function of stellar mass and redshift

$$M(N2, z) = 3.696 \xi + 3.236 \xi^{-1} + 0.729 \xi^{-2} + 14.928 + 0.156 (1 + z)^2,$$

with

$$\xi(N2, z) \equiv N2 + 0.138 - 0.042 (1 + z)^2. \quad (3)$$

This equation can be reversed numerically to obtain $N2(M, z)$. For visual clarity, we do not reverse this equation algebraically, but we provide Table 1 for a convenient lookup of $N2(M, z)$.

In the following, we test our model on data at $z \sim 0$, $z \sim 1.6$, and $z \sim 2.3$. Figure 5 shows the relation between stellar mass and $[\text{N II}]/\text{H}\alpha$ at the three different redshifts. For $z \sim 0$ and $z \sim 2.3$, the symbols show

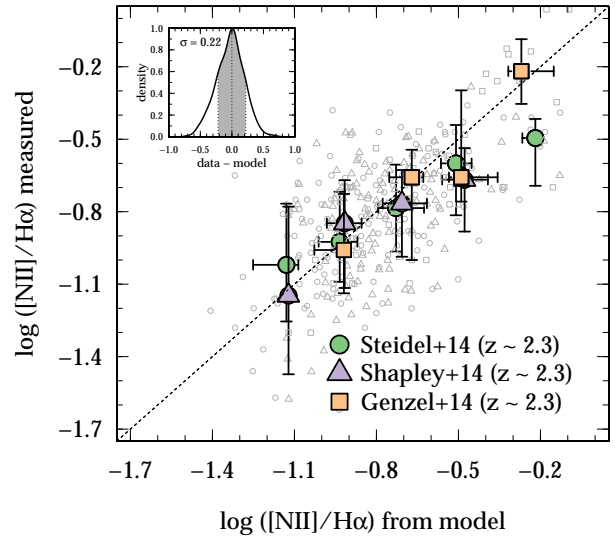


Figure 6. Quantification of scatter in the predicted $[\text{N II}]/\text{H}\alpha$ ratios based on the Steidel et al. (2014) (circles), Shapley et al. (2015) (triangles), and Genzel et al. (2014) (squares) samples at $z \sim 2.3$. The individual measurements are shown in gray, the medians with 1σ scatter from data and model are shown as large filled symbols in green, purple, and orange, respectively. The 1-to-1 relation is indicated by the dotted line. For our model predictions of $[\text{N II}]/\text{H}\alpha$, we expect a scatter of ~ 0.22 dex (see inset), which we find to be constant with stellar mass.

the medians in stellar mass and $[\text{N II}]/\text{H}\alpha$ with 1σ scatter from SDSS and Steidel et al. (2014), for $z \sim 1.6$ we use the median stacks in stellar mass provided by Kashino et al. (2017), thus the errors represent the error on the median and not the actual scatter. We also show the medians at $z \sim 1$ and $z \sim 2.3$ from Genzel et al. (2014) covering the massive end of the galaxy mass function. Our parameterization at $z = 0, 1.6$, and 2.3 using Equation 3 is shown with lines. The hatched regions show the range of our model values for the redshift distribution of the observed samples. Our model predicts the $[\text{N II}]/\text{H}\alpha$ ratio in general within ~ 0.1 dex of the observed data, which is well within the scatter of the data at all redshifts $0 \lesssim z \lesssim 2.3$ and stellar masses $9.5 < \log(M/M_\odot) < 11.0$. This is remarkable since the relation between $[\text{N II}]/\text{H}\alpha$ and stellar mass is solely based on the local SDSS data without information from higher redshifts. However, we note that our parameterization systematically under-predicts the $[\text{N II}]/\text{H}\alpha$ ratios at very low masses ($\log(M/M_\odot) \lesssim 9.5$) and low redshifts ($z \sim 0$) by up to 0.2 dex. Furthermore, we notice that our model over-predicts $[\text{N II}]/\text{H}\alpha$ ratios of the most massive galaxies ($\log(M/M_\odot) \sim 11.0$) in the Steidel et al. (2014) sample systematically by ~ 0.1 dex. The former is likely due to the generally larger residuals in the parameterization of $M(O3, N2)$ for local low-mass galaxies (right panel of Figure 3). The latter can be ex-

plained two-fold. First, Equation 2 is an extrapolation at $\log(M/M_{\odot}) \gtrsim 11.0$ as there are very few star-forming galaxies that are massive and low metallicity in the local SDSS sample. Specifically, there are only 191 star-forming galaxies at $\log(M/M_{\odot}) > 10.8$ with $[\text{O III}]/\text{H}\beta > 1$ (the region on the BPT diagram that is occupied by the high-redshift samples), which represents less than 0.1% of the total sample. At $\log(M/M_{\odot}) > 11$ this amount reduces to 73 galaxies. Second, the statistics of massive high-redshift galaxies in current spectroscopic samples is poor and dominated by sample selection and cosmic variance. Especially, we note that our model predicts almost perfectly the $[\text{N II}]/\text{H}\alpha$ flux ratios of the sample by Genzel et al. (2014) who specifically targeted massive galaxies at $z \sim 2$ (confirmed AGNs removed).

3.5. Scatter in $[\text{N II}]/\text{H}\alpha$ line ratios

Our model provides *median* $[\text{N II}]/\text{H}\alpha$ ratios for a given redshift and stellar mass. This median is mainly defined by the BPT main-sequence locus that we parameterized in Section 3.2. Deviations from this locus will lead to a physical scatter around the median $[\text{N II}]/\text{H}\alpha$ ratios provided by our model. Here, we study the origin and amplitude of this scatter in more detail as well as the impact of measurement uncertainties.

Figure 6 compares the true $[\text{N II}]/\text{H}\alpha$ values to the ones obtained from our model at $z \sim 2.3$ based on the Steidel et al. (2014), Genzel et al. (2014), and Shapley et al. (2015)⁸ samples, for which these measurements ($[\text{N II}]/\text{H}\alpha$ and stellar mass) are published for individual galaxies. Apart from the good agreement on average between model and true $[\text{N II}]/\text{H}\alpha$ values, we measure a (log-symmetric) 1σ scatter of 0.22 dex (inset in Figure 6), which we find to be constant with $[\text{N II}]/\text{H}\alpha$ ratio (hence stellar mass). This scatter is identical for the individual samples at $z \sim 2.3$. The same computation for $z \sim 1.6$ and local galaxies reveal a scatter of 0.21 dex and 0.13 dex, respectively (Appendix B). This scatter is introduced by differences in physical properties of the galaxies as well as measurement uncertainties as discussed below.

3.5.1. Physical scatter due to sSFR

The upper panel in Figure 2 shows that mainly stellar mass determines the position of galaxies on the BPT diagram for a given BPT main-sequence as parameterized in Section 3.2 as a function of redshift. On the other hand (as shown on the lower panel of Figure 2), the sSFR varies mostly *perpendicular* to the BPT main-sequence loci. We therefore argue that sSFR acts as a

secondary parameter defining the location of galaxies on the BPT diagram at a fixed stellar mass and redshift. By fixing a BPT main-sequence for our model, we indirectly assume a median sSFR given by the sample that is used to anchor our model (in our case the average sSFR of the Steidel et al. (2014) sample, which represents well the stellar mass vs. SFR main-sequence at $z \sim 2.3$). Because of the remarkably constant ~ 0.3 dex scatter of the stellar mass versus SFR relation (e.g., Daddi et al. 2007; Noeske et al. 2007; Schreiber et al. 2015; Tomczak et al. 2016), galaxies at a fixed stellar mass and redshift show a range in sSFR, hence inducing a scatter perpendicular to the average BPT main-sequence locus. This (physical) scatter directly translates into the scatter seen in our comparison of true and model $[\text{N II}]/\text{H}\alpha$ ratios.

The effect of selection biases on the BPT main-sequence can be seen by comparing the result of the MOSDEF (Shapley et al. 2015) and Steidel et al. (2014) studies. The BPT main-sequence locus at $z \sim 2.3$ derived from the MOSDEF survey is offset by up to ~ -0.2 dex in $[\text{N II}]/\text{H}\alpha$ from the Steidel et al. (2014) locus, although both samples have almost identical distribution in stellar mass and redshift. Using the MOSDEF locus for our model would therefore result in up to 0.1 – 0.2 dex larger $[\text{N II}]/\text{H}\alpha$ flux ratios at a fixed stellar mass (dashed line in Figure 5). The physical reason for the seeming discrepancy is likely a slight excess of high sSFR galaxies in the Steidel et al. (2014) sample compared to the MOSDEF sample (as also pointed out by Shapley et al. 2015), in agreement with our identification of sSFR as a secondary parameter. This excess in sSFR could be caused by the UV color selection in the case of the Steidel et al. (2014) sample, which favors higher star formation compared to a continuum or stellar mass selected sample as in the case of MOSDEF.

An “emission line complete” sample would allow us to derive the correct average BPT main-sequence locus at a given redshift and hence anchor our model at high redshifts, however, selecting such a sample is almost impossible at these redshifts since there will always be certain selection biases. At $z \sim 2.3$, the likely average locus would be somewhere in between the MOSDEF and Steidel et al. (2014) derivations and therefore not far from our model predictions (see Figure 5).

3.5.2. Scatter due to measurement uncertainties

In addition to differences in stellar mass and sSFR, measurement uncertainties can contribute to the scatter on the BPT diagram for a given sample at a given redshift. This is becoming increasingly more valid at higher redshift where the measurements become lower S/N. These concerns can dilute a clear BPT main-sequence locus.

⁸ Stellar masses and $[\text{N II}]/\text{H}\alpha$ measurements are taken from Sanders et al. (2017). No redshifts are published for individual galaxies, therefore we assume $z = 2.3$ for all galaxies.

Steidel et al. (2014) quote an intrinsic scatter (i.e., corrected for measurement uncertainties) of 0.12 dex on the BPT main-sequence locus (in $[\text{N II}]/\text{H}\alpha$ and $[\text{O III}]/\text{H}\beta$) at $z \sim 2.3$. This is consistent with the measurements by Shapley et al. (2015) at the same redshift, Kashino et al. (2017) at $z \sim 1.6$, and for local galaxy samples (~ 0.11 dex, Kewley & Ellison 2008).

We measure an observed scatter of 0.22 dex, 0.21 dex

and 0.13 dex between true and model $[\text{N II}]/\text{H}\alpha$ ratios at $z \sim 2.3$, $z \sim 1.6$ and $z \sim 0$, respectively (see also Appendix B). Comparing this to the intrinsic scatter in the BPT main-sequence loci given above suggests that roughly half of the uncertainties in the model derived $[\text{N II}]/\text{H}\alpha$ ratios at $z \sim 1.6$ and 2.3 are due to the combined uncertainties in the individual measurements and our model.

Table 1. Look-up table for $[\text{N II}]/([\text{N II}]+\text{H}\alpha)_{\text{tot}}$ flux ratios (including both $[\text{N II}]$ emission lines) given in linear scale from $8.5 < \log(M/M_{\odot}) < 11.1$ and $0 < z < 2.6$ derived by Equation 3[†]. These values can be used for the conversion of the observed $([\text{N II}]+\text{H}\alpha)_{\text{tot}}$ flux to the intrinsic $\text{H}\alpha$ flux.

$\log(M/M_{\odot})$	Redshift														
	0.0	0.2	0.4	0.6	0.8	1.0	1.2	1.4	1.6	1.8	2.0	2.2	2.4	2.6	
8.5	0.07	0.06	0.06	0.06	0.06	0.06	0.05	0.05	0.05	0.05	0.05	0.05	0.04	0.04	
8.7	0.08	0.08	0.08	0.07	0.07	0.07	0.07	0.06	0.06	0.06	0.06	0.05	0.05	0.05	
8.9	0.10	0.10	0.09	0.09	0.08	0.08	0.08	0.07	0.07	0.07	0.07	0.06	0.06	0.06	
9.1	0.12	0.12	0.11	0.11	0.10	0.10	0.09	0.09	0.08	0.08	0.08	0.07	0.07	0.07	
9.3	0.16	0.15	0.14	0.13	0.13	0.12	0.11	0.11	0.10	0.10	0.09	0.09	0.08	0.08	
9.5	0.21	0.20	0.19	0.17	0.16	0.15	0.14	0.13	0.12	0.11	0.11	0.10	0.10	0.09	
9.7	0.26	0.25	0.24	0.23	0.21	0.19	0.17	0.16	0.15	0.14	0.13	0.12	0.11	0.11	
9.9	0.28	0.28	0.28	0.28	0.26	0.24	0.22	0.20	0.18	0.16	0.15	0.14	0.13	0.13	
10.1	0.30	0.30	0.31	0.31	0.31	0.30	0.28	0.25	0.22	0.20	0.18	0.17	0.16	0.15	
10.3	0.31	0.32	0.32	0.33	0.33	0.34	0.33	0.32	0.29	0.25	0.22	0.20	0.19	0.17	
10.5	0.32	0.33	0.34	0.34	0.35	0.36	0.36	0.36	0.35	0.32	0.28	0.25	0.22	0.21	
10.7	0.33	0.34	0.35	0.36	0.36	0.37	0.38	0.39	0.39	0.38	0.35	0.31	0.27	0.24	
10.9	0.34	0.35	0.35	0.36	0.37	0.39	0.40	0.41	0.42	0.42	0.42	0.38	0.33	0.29	
11.1	0.34	0.35	0.36	0.37	0.38	0.40	0.41	0.42	0.44	0.45	0.46	0.45	0.41	0.36	

[†] Note that Equation 3 is basically a parameterization of the stellar mass vs. gas-phase metallicity relation as a function of redshift.

3.6. Large $[\text{N II}]/\text{H}\alpha$ ratios in massive high- z galaxies

Our model predicts larger $[\text{N II}]/\text{H}\alpha$ ratios in massive ($\log(M/M_{\odot}) \sim 11.0$) high-redshift galaxies compared to similar massive galaxies at $z \sim 0$ (Figure 5). This is also suggested by the shifted high-redshift BPT loci to higher $[\text{O III}]/\text{H}\beta$ and $[\text{N II}]/\text{H}\alpha$ ratios (Figure 4). This could be caused by an increasing amount of galaxies with broad line emission at such high stellar masses. In fact, Genzel et al. (2014) study the statistics of broad $\text{H}\alpha$ and $[\text{N II}]$ emission in samples of star-forming galaxies at $1 < z < 3$ and suggest broad nuclear components due to a combination of shocks and photoionization and also AGNs in more than half of these galaxies at $\log(M/M_{\odot}) \sim 11.0$. Specifically, they suggest that the contribution of such galaxies to high-mass samples is at least as large as AGN samples selected with X-ray, optical, infrared, or radio indicators. The median $[\text{N II}]/\text{H}\alpha$ ratio per stellar mass bin from the Genzel et al. (2014) sample is higher than the one of the Steidel et al. (2014) sample and in good

agreement with our model (Figure 5). Along the same lines, Kewley et al. (2013) motivates a shift of the separation line between normal star-forming galaxies and AGNs to higher $[\text{N II}]/\text{H}\alpha$ and $[\text{O III}]/\text{H}\beta$ ratios by means of increased photoionization at higher redshifts.

4. IMPLICATIONS

In the previous section, we have derived an empirical parameterization to predict the $[\text{N II}]/\text{H}\alpha$ flux ratios for galaxies up to $z \sim 3$ and $\log(M/M_{\odot}) = 11.0$. Here, we study in detail the implications of our model on

- Data interpretation of current low spectral resolution surveys (Section 4.1),
- The $[\text{N II}]$ contamination of future flux-limited surveys (Section 4.2),
- Spectroscopic redshift measurements from the blended $\text{H}\alpha$ and $[\text{N II}]$ lines (Section 4.3),

- Expected H α emitter number counts of future surveys derived from current H α LFs (Section 4.4).

For surveys with low spectral resolution, H α is commonly blended with *both* [N II] emission lines (at rest-frame 6548 Å and 6584 Å). Hence, a more useful quantity to quote is the *total [N II] flux contamination fraction*, which we define as $([\text{N II}]\lambda\lambda 6548, 6584)/([\text{N II}]\lambda\lambda 6548, 6584 + \text{H}\alpha)$, in short $[\text{N II}]/([\text{N II}] + \text{H}\alpha)_{\text{tot}}$. In the following we assume for the flux of the second [N II] line, blue-ward of H α , $[\text{N II}]\lambda 6548 = \frac{1}{3}[\text{N II}]\lambda 6584$ (Acker et al. 1989). In Table 1 we provide $[\text{N II}]/([\text{N II}] + \text{H}\alpha)_{\text{tot}}$ flux ratios in linear scale as a function of stellar mass and redshift derived from our Equation 3. This table can be used as a convenient tool to convert observed $([\text{N II}] + \text{H}\alpha)_{\text{tot}}$ fluxes and luminosities (including both [N II]) into intrinsic H α fluxes and luminosities.

4.1. Data interpretation of current surveys at low spectral resolution

In current grism surveys with low spectral resolution where [N II] and H α are not resolved, such as WISPS, a constant total [N II] flux contamination fraction is commonly used to obtain intrinsic H α values from which SFRs or H α luminosity functions are measured (Colbert et al. 2013; Mehta et al. 2015). The value generally applied is 0.29 (i.e., 29%, or $F_{\text{H}\alpha} = 2.5 \times F_{[\text{N II}]}$) that follows from $[\text{N II}]/([\text{N II}] + \text{H}\alpha)_{\text{tot}} = 0.29$ according to the average population of galaxies at $z \sim 0$ with an H α EW of less than 200 Å. However, such an assumption can be misleading, since the true [N II]/H α flux ratio can vary by an order of magnitude across samples depending on redshift and stellar mass (Figure 5). Other studies apply a variable [N II] contamination correction using the relation between [N II]/H α and the H α equivalent-width (\propto sSFR), resulting in a slightly lower median [N II] contamination of $\sim 24\%$ (e.g., Villar et al. 2008; Sobral et al. 2009; Lee et al. 2012; Sobral et al. 2012, 2013). While this approach is more accurate, it could still miss the dependency with stellar mass.

To study the accuracy of a constant correction, we show in Figure 7 the total [N II] flux contamination fraction on linear scaling as a function of redshift at four different stellar masses. Our model is shown as lines with the hatched region corresponding to the approximate, redshift dependent scatter (see Section 3.5). Observed data at $0 < z < 2.3$ binned in redshift and stellar mass ($\Delta \log(M/M_{\odot}) = 0.5$ around stellar masses shown) is shown with symbols and for reference a constant total [N II] flux contamination fraction of 29% is indicated by the horizontal dashed line. We find that the assumption of a constant contamination of 29% is only justified for galaxies at $z < 1$ and $\log(M/M_{\odot}) \sim 10.0$ as

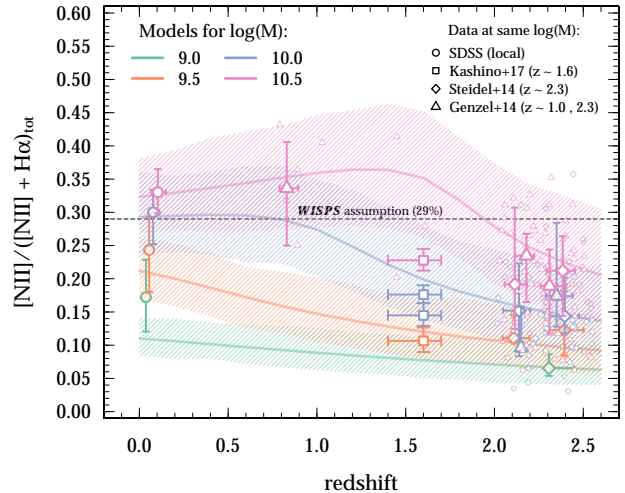


Figure 7. Total [N II] flux contamination fraction as a function of redshift for 4 different stellar masses from our model (lines). Medians of observations are shown with large symbols of the same colors, individual measurements are shown with small symbols. The WISPS assumption (Colbert et al. 2013) of a constant 29% [N II] contamination is shown as dashed line for reference. Note that this assumption is only justified for $\log(M/M_{\odot}) \sim 10.0$ galaxies at $z < 1$ and $\log(M/M_{\odot}) \sim 10.5$ galaxies at $z \sim 2$ and otherwise over- or underestimates the true contamination by a significant factor.

well as $z \sim 2$ and $\log(M/M_{\odot}) \sim 10.5$. Otherwise, our model shows that such an assumption generally overestimates the true contamination, which can be as severe as a factor of 3 for $\log(M/M_{\odot}) \sim 9$ galaxies at all redshifts. But also the [N II] contamination of high-mass ($\log(M/M_{\odot}) > 10.0$) galaxies at $z \sim 1.5 - 2$ is overestimated by factors of 1.5 – 2.

This has important consequences for the intrinsic H α measurements by WISPS. Specifically, since this sample spans a range in stellar mass of $8.5 < \log(M/M_{\odot}) < 10$ at $0.5 < z < 1.5$ (Atek et al. 2010; Henry et al. 2013), we expect the H α luminosities to be systematically underestimated by 20–30% or more on average. Furthermore, since [N II]/H α is a function of stellar mass and therefore (via the SFR) also a function of H α luminosity, the shape of the H α LF is affected, which has an impact on number counts for future surveys (see Section 4.4).

4.2. Total [N II] flux contamination fraction distribution for future flux limited surveys

Future large area surveys at low spectral resolution such as *Euclid* will suffer from [N II] and H α blending. A proper de-blending of these lines is important for measuring physical quantities from H α such as SFR, galaxy kinematics, or dark matter properties, but also for accurate spectroscopic redshift used for cosmology. Here, we present realistic zeroth-order predictions for the total [N II] flux contamination fraction as a function

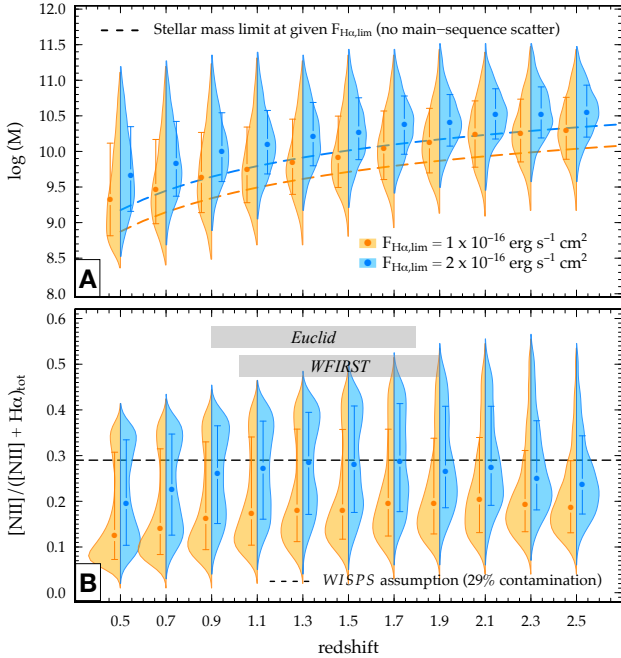


Figure 8. Expected stellar mass (A) and total [N II] flux contamination fraction (B) distributions as a function of redshift for flux limited surveys at $1 \times 10^{-16} \text{ erg s}^{-1} \text{ cm}^{-2}$ (orange) and $2 \times 10^{-16} \text{ erg s}^{-1} \text{ cm}^{-2}$ (blue), similar to expectations for *WFIRST* and *Euclid*, respectively. The medians of the distributions (points with error bars) are slightly displaced in redshift for clarity. Our prediction is based on the empirical stellar mass vs. SFR relation (see text for details on derivation). For a survey like *Euclid* ($0.9 < z < 1.8$) or *WFIRST* ($1 < z < 2$), we expect a large variation in the total [N II] flux contamination fraction of 5–40% ($z = 1$) and 10–45% ($z = 1.8$).

of redshift for flux-limited surveys at $0.5 < z < 2.5$. In the following, we assume observed line flux limits of $1 \times 10^{-16} \text{ erg s}^{-1} \text{ cm}^{-2}$ and $2 \times 10^{-16} \text{ erg s}^{-1} \text{ cm}^{-2}$, similar to expectations for *WFIRST* (5σ for a source of radius $0.3''^9$, see also [Spergel et al. 2015](#)) and *Euclid* (3.5σ for a source with diameter $0.6''$, [Vavrek et al. 2016](#)), around the observed wavelength of $\text{H}\alpha$.

From the observed redshift-dependent stellar mass functions ([Ilbert et al. 2013](#); [Davidzon et al. 2017](#)), we draw 10,000 galaxies, to which we assign a SFR via the observed relation between stellar mass and SFR (main-sequence of star-forming galaxies) as parameterized by [Schreiber et al. \(2015\)](#) including a scatter of 0.3 dex. We then select galaxies above a SFR threshold derived from the line flux limits, which we converted to limiting $\text{H}\alpha$ luminosities (at given redshift) and then SFRs using the [Kennicutt \(1998\)](#) description. The resulting stellar mass distributions for the two flux limits as a function of redshift are shown on the left (orange) and right (blue)

side, respectively, of the “Violin diagram” in panel (A) of Figure 8. Note that the distributions extend across the sharp stellar mass limit derived from the observed line flux limits (dashed lines) because of the scatter of the star-forming main-sequence.

Panel (B) shows the corresponding distributions of the total [N II] flux contamination fraction as a function of redshift. The expected range in $[\text{N II}]/([\text{N II}] + \text{H}\alpha)_{\text{tot}}$ is large because of the wide distribution in the stellar masses. Furthermore it is important to note that the distribution is double-peaked out to $z \sim 1.5$. The first peak is due to the dominant number of low-mass galaxies (with low [N II]/ $\text{H}\alpha$ ratios), while the second peak arises due to the flattening of the stellar mass vs. [N II]/ $\text{H}\alpha$ ratio relation at large stellar masses. This can be seen clearly in Figure 5 where the [N II]/ $\text{H}\alpha$ – M relation at $z = 0$ flattens for stellar masses above approximately $\log(M/M_{\odot}) \sim 10$. Towards higher redshifts, the stellar mass distribution becomes tighter and the [N II] contamination becomes single-peaked because of the increasing luminosity limit.

For a *Euclid*-like survey ($0.9 < z < 1.8$), we expect a large variation in the total [N II] flux contamination fraction of 10–40% ($z = 1$) and 15–45% ($z = 1.8$). For a *WFIRST*-like survey ($1.0 < z < 1.9$), these numbers are lower (5–40% and 10–45%, respectively) due to the higher line sensitivity allowing to probe more galaxies at lower stellar masses, hence lower [N II]/ $\text{H}\alpha$ ratios.

4.3. Spectroscopic redshift measurements from blended $\text{H}\alpha$ and [N II]

The blending of the $\text{H}\alpha$ and [N II] lines can result in biases in the determination of spectroscopic redshifts. Here, we study this bias as a function of stellar mass and redshift via a simple preliminary simulation based on the predictions from our model.

We approximate each of the three emission lines by a Gaussian and assign fluxes relative to the $\text{H}\alpha$, namely, $F([\text{N II}]\lambda 6584) = [\text{N II}]/\text{H}\alpha \times F(\text{H}\alpha)$, and $F([\text{N II}]\lambda 6548) = \frac{1}{3} F([\text{N II}]\lambda 6584)$. The flux ratio [N II]/ $\text{H}\alpha$ is computed from our Equation 3 and we use vacuum wavelengths for the emission lines (6549.86 \AA , 6564.61 \AA , and 6585.27 \AA for [N II] $\lambda 6548$, $\text{H}\alpha$, and [N II] $\lambda 6584$, respectively). For the full width at half maximum (FWHM) of the lines we assume 250 km s^{-1} as commonly measured on average by spectroscopic surveys at $\log(M/M_{\odot}) = 10$. Our final results do not significantly depend on the exact values for the FWHMs, mainly because of the somewhat low spectral resolution of *Euclid*.

We then convolve and bin this input spectrum to the *Euclid* resolution and pixel size, assuming a spectral dispersion of 13.4 \AA/px and a plate scale of $0.3''/\text{px}$ ([Vavrek et al. 2016](#)). The spectral dispersion results in an $R = \lambda/\Delta\lambda$ of 490–735 for $\text{H}\alpha$ at $1 < z < 2$ in a

⁹ *WFIRST* Formulation Science Working Group, 2017, private communication.

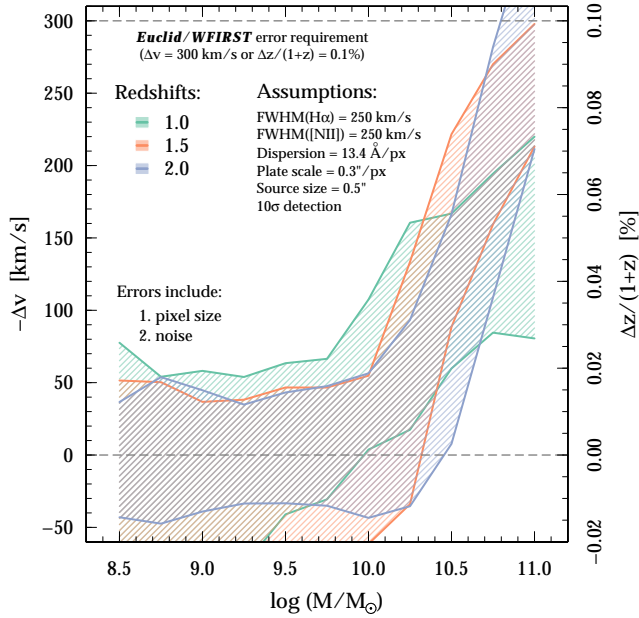


Figure 9. Velocity shifts and redshift biases in the determination of the $H\alpha$ wavelength centroid due to blending with $[N\text{II}]$ for a source size of $0.5''$ and $S/N=10$ on the line. Shown are simulations for a range of stellar masses at $z=1$ (green), 1.5 (orange), and 2.0 (blue). The $H\alpha$ centroid is measured by a Gaussian fit to the observed (i.e., resolution adjusted) spectrum at $6500 - 6600 \text{ \AA}$ in rest-frame (corresponding to $7 - 11$ *Euclid* pixel-pairs at $1 < z < 2$). The hatched area shows the uncertainty due to *Euclid*'s finite spectral resolution and measurement noise. Our simple simulation suggests that the velocity shifts are less than the error requirement for *Euclid* and *WFIRST* ($\Delta v = 300 \text{ km s}^{-1}$ or $\Delta z/(1+z) = 0.1\%$). The bias is increasing steeply with stellar mass at $\log(M/M_\odot) \gtrsim 10$ due to an increasing $[N\text{II}]/H\alpha$ ratio.

2-pixel resolution element for a source of $0.3''$ diameter. In the following, we assume a more realistic source diameter of $0.5''$ and a point spread function (PSF) FWHM of $0.4''$ (Vavrek et al. 2016), which decreases the resolution by a factor ~ 2 , and bin the final observed spectrum to a 2-pixel resolution element. Note that *Euclid*'s resolution element decreases proportional to $(1+z)$ for increasing redshift, i.e., $13.4/(1+z) \text{ \AA/px}$, as the observed spectrum stretches in wavelength. This results in an increase in resolution of 33% from $z=1$ to $z=2$ for a galaxy of fixed apparent size. We also add noise to the output spectrum according to the required 3.5σ flux limit of $2 \times 10^{-16} \text{ erg s}^{-1} \text{ cm}^{-2}$, which we assume is the integrated flux over the blended $H\alpha$ and $[N\text{II}]$ emission lines. In the following, we assume a source detected at 10σ in integrated line flux. This is a good approximation for most of *Euclid*'s detected sources according to predictions of the $H\alpha$ luminosity function (see Section 4.4).

Our preliminary results indicate that *Euclid* will generally not resolve $H\alpha$ and $[N\text{II}]$ for most of the assumed $[N\text{II}]/H\alpha$ values (and thus stellar masses) for a source

of $0.5''$ and a S/N of 10, however, an asymmetry of the blended line caused by the $[N\text{II}]$ red-ward of $H\alpha$ is identifiable. We therefore compute the $H\alpha$ centroid on the final convolved and binned spectrum by fitting a Gaussian at $6500 - 6600 \text{ \AA}$ in rest-frame. This wavelength width encompasses both $[N\text{II}]$ as well as $H\alpha$ and corresponds to $7 - 11$ *Euclid* pixel-pairs (26.8 \AA per pair) at $1 < z < 2$.

Figure 9 shows the resulting centroid shifts in velocity and redshift with respect to the true $H\alpha$ wavelength as a function of stellar mass at $z=1.0, 1.5$, and 2.0 . The hatched area combines the errors from the noise and the finite pixel size of *Euclid*. The latter is obtained by shifting the binning of the final spectrum by up to half a resolution element. The redshift bias increases towards higher stellar masses due to the larger $[N\text{II}]/H\alpha$ flux ratio¹⁰. Furthermore, the bias decreases slightly with increasing redshift at a fixed stellar mass due to the increasing resolution with increasing redshift. While this effect is only small for a source of $S/N=10$, we would expect a much larger reduction of the bias at higher S/N , where $[N\text{II}]$ and $H\alpha$ will likely be resolved at the highest redshifts. However, the amount of detected sources at $S/N > 10$ and high redshift is likely small (see Section 4.4).

In general, we find velocity shifts that are better than the error requirement for *Euclid*, which is $\Delta v = 300 \text{ km s}^{-1}$ or $\Delta z/(1+z) = 0.1\%$ (Vavrek et al. 2016). For a galaxies at $\log(M/M_\odot) < 10$, we expect negligible biases ($|\Delta v| < 50 \text{ km s}^{-1}$), however, the biases increase sharply at $\log(M/M_\odot) > 10$ due to the increasing $[N\text{II}]/H\alpha$ ratio. For a galaxy of $\log(M/M_\odot) = 11$, we expect significant biases around $100 - 300 \text{ km s}^{-1}$, or $\Delta z/(1+z) \sim 0.04 - 0.10\%$. If uncorrected, such a shift will introduce a bias in the BAO measurements in the radial direction. Specifically, at $z=1$ ($z=1.5$), a shift of $\Delta z/(1+z) = 0.04 - 0.10\%$ corresponds to $1.0 - 2.4 \text{ Mpc}$ ($0.7 - 1.8 \text{ Mpc}$) or roughly $0.6 - 1.6\%$ ($0.5 - 1.2\%$) of the BAO scale at $\sim 150 \text{ Mpc}$. This is significant since the BAO peak itself is a few-percent level signal in the galaxy correlation function that needs to be measured at the precision of a few percent or better. Finally, we note that the $[N\text{II}]/H\alpha$ blending may lead to additional systematic effects for BAO/RSD measurements, if metallicity evolution is correlated with density. This will be examined further in future studies.

Our preliminary simulation is very basic and we will use more realistic grism simulations in the future for more detailed investigations. Furthermore, the evolu-

¹⁰ The $[N\text{II}]/H\alpha$ flux ratio is 0.05 (0.03) and 0.40 (0.63) in linear scaling for a galaxy with $\log(M/M_\odot) = 8.5$ and $\log(M/M_\odot) = 11.0$, respectively, for $z=1$ ($z=2$).

tion and distribution of the angular sizes of the galaxies should be taken into account (the combined effect of the increasing cosmological angular diameter distance and the decreasing physical size of the galaxies with redshift).

Finally, we note that H α and [N II] emission lines will be likely resolved for many galaxies detected by *WFIRST* at its spectral resolution of $R \sim 600 - 900$ for H α at $1 < z < 2$ (with a dispersion of $10.85 \text{ \AA}/\text{px}$, [Spergel et al. 2015](#)), therefore much smaller biases are expected.

4.4. Impact on H α luminosity function and number count predictions for Euclid and WFIRST

The dark energy figure-of-merit for both *WFIRST* and *Euclid* is very sensitive to the number density of H α emitting galaxies.

Measurements of the observed *blended* H α LF of low-resolution *HST* grism surveys are used to predict the observed number counts for future large surveys such as *WFIRST* or *Euclid* ([Colbert et al. 2013](#); [Mehta et al. 2015](#)). While these number counts are accurate for the redshift, stellar mass, and sSFR distribution of the grism surveys, any extrapolation beyond that to match *WFIRST*'s and *Euclid*'s parameter space requires the knowledge of the intrinsic H α LF and therefore an accurate assessment of the total [N II] flux contamination fraction (see also discussion in [Pozzetti et al. 2016](#)).

Importantly, future large area galaxy surveys will be predominantly probing the bright-end of the LF at its exponential decline and therefore any uncertainty in the brightest H α luminosities will have a significant impact on the H α number counts. Furthermore, the derivation of the intrinsic H α LF will be important for studying many physical properties of the galaxies, such as their SFRs. This does not only apply to future studies but also current grism spectroscopy and narrow-band photometric observations that do not resolve [N II] and H α . Here we investigate the *relative* change in the intrinsic H α emitter number counts when using different corrections for [N II] contamination. Specifically, we study (i) a constant 29% contamination and (ii) the stellar mass and redshift dependent [N II]/H α flux ratios predicted by our model (Equation 3).

To derive the intrinsic H α LF using our model [N II] contamination (Φ_{model}), we start with the *observed* ([N II]+H α)_{tot} LF ($\Phi_{\text{obs}}(L)$) measured by [Colbert et al. \(2013\)](#) at $0.9 < z < 1.5$. We obtain this LF from their published H α LF ($\log \Phi_* = -2.70$, $\log L_* = 42.18$, and $\alpha = -1.43$) by dividing the luminosities by a factor of $(1 - 0.29)$ to undo *constant* total [N II] flux contamination fraction correction of 29%, which the authors applied. In the following, we treat Φ_{obs} as the true observed LF. Importantly, this LF is not corrected for

[N II] contamination and dust. Φ_{obs} is redshift dependent because of the evolution of the star forming main-sequence, but here we do not model this dependence across $z = 1 - 2$ as we are only interested in the effects of [N II] contamination and not the absolute number of galaxies. On the other hand, the stellar mass and redshift dependent [N II] contamination correction will change the *intrinsic* H α LF across the redshift range studied here. To obtain Φ_{model} , we choose an approach, which only uses the measured $([\text{N II}]+\text{H}\alpha)_{\text{tot}}$ luminosities as input and assumes the most likely underlying stellar mass distribution (robustly determined from other studies) from which we obtain the [N II] contamination from our model. This approach has the advantage that it enables an easy implementation and propagation of a variety of uncertainties into final results. Furthermore, this method results in reliable intrinsic H α LFs even if the mass distribution is poorly measured due to the lack of sufficient multi-wavelength data, as long as the selection function of the galaxy sample is known. Here, we make use of the [Schreiber et al. \(2015\)](#) parameterization of the star-forming main-sequence to derive the underlying stellar mass distribution (we comment below on possible shortcomings). We note that the choice of different parameterizations (e.g., [Tomczak et al. 2016](#)) should not change the following results. We start with a distribution of $([\text{N II}]+\text{H}\alpha)_{\text{tot}}$, which we sample from Φ_{obs} . To obtain stellar masses for these galaxies, we use a “backwards engineering” technique. First, we convert the SFRs of the [Schreiber et al. \(2015\)](#) parameterization into H α luminosities using the [Kennicutt \(1998\)](#) prescription. Thereby we include a dispersion of 0.3 dex measured on the SFR vs. stellar mass main-sequence. Second, we redden the H α luminosities according to the relation between $A_{\text{H}\alpha}$ (H α extinction) and stellar mass robustly derived from the spectra of local galaxy samples in SDSS ([Garn & Best 2010](#)). This relation holds for the WISPS sample at $0.8 < z < 1.5$ as shown in [Domínguez et al. \(2013\)](#). Third, we add the contribution of [N II] to the H α luminosity by using our model. Finally, this translation between dust reddened $([\text{N II}]+\text{H}\alpha)_{\text{tot}}$ luminosities and stellar masses allows us to obtain the underlying stellar mass distribution and intrinsic H α luminosities for the galaxy sample describing Φ_{obs} , from which we are now able to re-compute the intrinsic H α LF Φ_{model} .

Figure 10 shows the three LFs; (i) the total $([\text{N II}]+\text{H}\alpha)_{\text{tot}}$ LF (Φ_{obs}), (ii) the H α LF corrected with constant 29% [N II] correction as published in [Colbert et al. \(2013\)](#) (Φ_{29}), and (iii) the H α LF with redshift and stellar mass dependent [N II] correction from our model (Φ_{model}). The LFs are shown in absolute values (top panels) and relative to Φ_{model} (bottom panels) at redshifts $z = 1$, $z = 1.5$, and $z = 2$. The luminosity limits

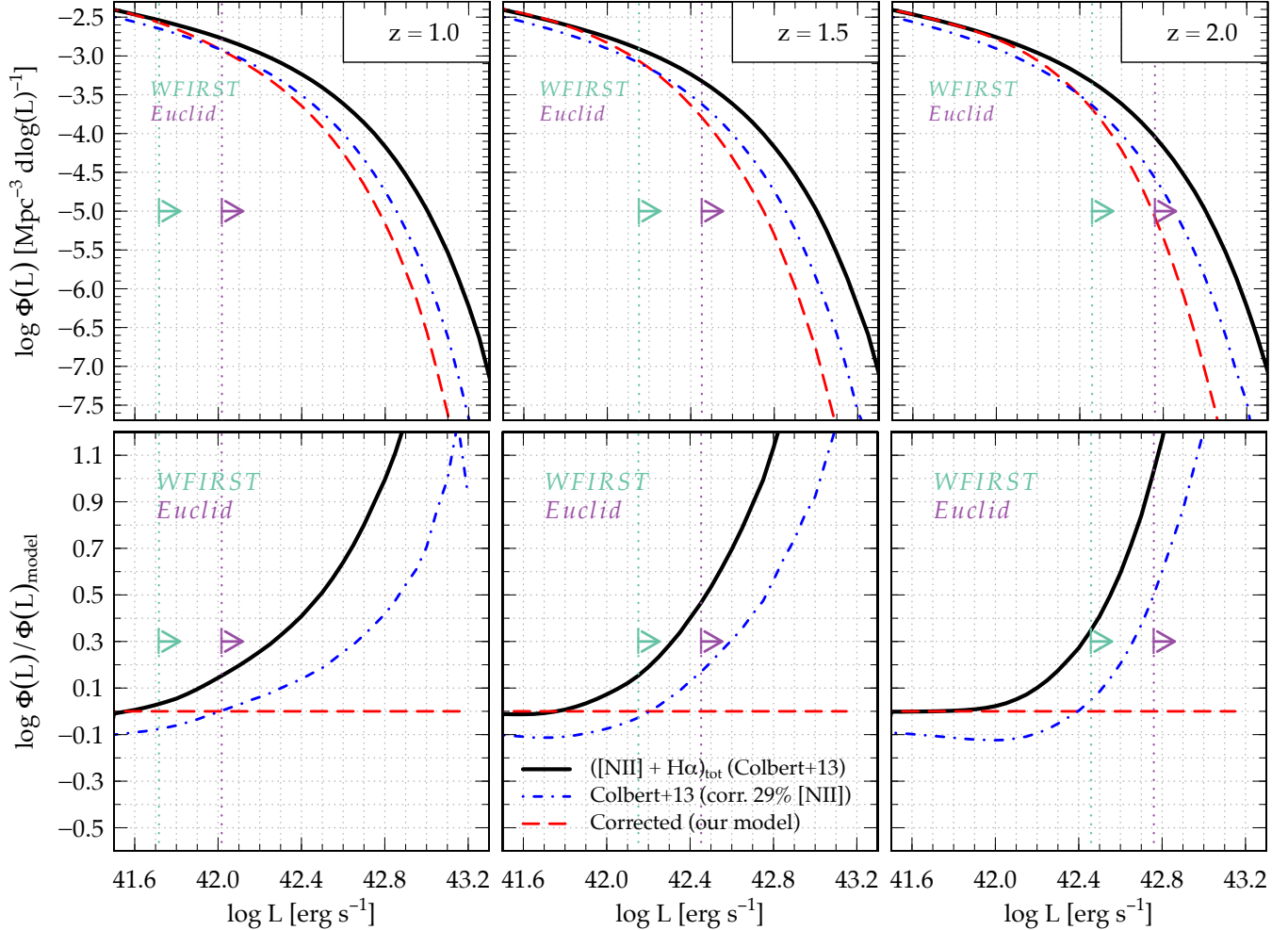


Figure 10. Differential effects of redshift- and stellar mass-dependent total $[\text{N II}]$ flux contamination corrections on the number counts of $\text{H}\alpha$ emitters displayed on the example of the Colbert et al. (2013) luminosity function at $0.9 < z < 1.5$ ($\langle z \rangle = 1.2$). *Top panels:* The black solid line shows the total $([\text{N II}] + \text{H}\alpha)_{\text{tot}}$ LF observed by Colbert et al. (2013) at $0.9 < z < 1.5$ (Φ_{obs}) and the blue dot-dashed line is the $\text{H}\alpha$ LF derived from a correction assuming a constant total $[\text{N II}]$ flux contamination fraction of 29% (Φ_{29}) by the same authors. The red dashed line shows the $\text{H}\alpha$ LF derived from the $([\text{N II}] + \text{H}\alpha)_{\text{tot}}$ LF using our model for $[\text{N II}]$ contamination (Φ_{model}). The luminosity limits for *WFIRST* and *Euclid* are shown as green and purple arrows, respectively. *Bottom panels:* The three LFs relative to our $[\text{N II}]$ corrected $\text{H}\alpha$ LF Φ_{model} with the same color-code as in the top panels. A redshift and stellar mass dependent total $[\text{N II}]$ flux contamination fraction is important to obtain accurate $\text{H}\alpha$ emitter number counts.

for *WFIRST* and *Euclid* (redshift dependent) are shown as arrows for reference. Note that only Φ_{model} changes with redshift due to the redshift-dependent $[\text{N II}]/\text{H}\alpha$ flux ratio, while the other LFs are unchanged.

First of all, it is evident that an accurate $[\text{N II}]$ correction is crucial at the bright-end of the LF where the number counts exponentially drop and the LF is dominated by massive galaxies with large $[\text{N II}]$ corrections (see Figure 7). Similarly, the difference between Φ_{29} and Φ_{model} increases towards the bright-end of the LF due to its steepness and the mass dependence of the $[\text{N II}]$ correction. While at $\log(L_{\text{H}\alpha}) < 42.4$ a constant $[\text{N II}]$ contamination correction generally underestimates the $\text{H}\alpha$ number counts by $\lesssim 0.1$ dex with respect to our model, at higher luminosities the deviation is more severe. For

example, at $z = 1.5$ the number counts of galaxies at $\log(L_{\text{H}\alpha}) = 43.0$ ($\log(M/M_{\odot}) \sim 10.6$) would be overestimated by approximately a factor of 8 (0.9 dex) with respect to using a mass and redshift dependent $[\text{N II}]$ contamination. This factor is expected to be less (factor 5, 0.7 dex) at $z = 1$ and more (factor 15, 1.2 dex) at $z = 2$ at the same $\text{H}\alpha$ luminosity. Such biases are not to be neglected as *Euclid* will probe the high-luminosity part of the $\text{H}\alpha$ LF as indicated by the purple arrows in Figure 10.

Our Equation 3 was also used in Merson et al. (2018) to transform the $[\text{N II}]$ blended $\text{H}\alpha$ flux in the WISPS data into true $\text{H}\alpha$ fluxes for calibrating the semi-analytical galaxy formation code GALACTICUS (Benson 2012), so that reliable forecasts of galaxy number counts

can be obtained for the galaxy redshift surveys planned for *Euclid* and *WFIRST*.

To conclude, we briefly discuss possible caveats of our approach. First, we note that the emission line selected WISPS galaxies may probe a different stellar mass distribution as in the [Schreiber et al. \(2015\)](#) study (which is based on *H* and *K* band continuum selected galaxies). Specifically, we would expect the average stellar mass at a given SFR to be lower in the case of emission line selected galaxies (e.g., [Ly et al. 2012](#)), hence our stellar masses would be overestimated. Assuming conservatively a factor of two lower average stellar masses per SFR would lead to $< 30\%$ lower $[\text{N II}]$ contamination over the mass range $9.5 < \log(M/M_\odot) < 11.0$ (approximately $41.9 < \log(L_{\text{H}\alpha}) < 43.5$) at $z = 1.5$. This translates into < 0.15 dex less overestimation of the $\text{H}\alpha$ emitter counts if using no or a constant 29% $[\text{N II}]$ contamination correction compare to our model. This is negligible compared to the large corrections needed at the bright end of the LF. Second, we note that the [Kennicutt \(1998\)](#) relation to obtain SFRs from $\text{H}\alpha$ luminosities was derived from galaxies with solar metallicity and an electron temperature of the ionized gas of 10^4 K. These assumptions may not be valid at high redshifts. Using the metallicity dependent parameterization of the Kennicutt relation by [Ly et al. \(2016a\)](#), we estimate that the SFR for a given $\text{H}\alpha$ luminosity is ~ 0.2 dex lower for galaxies at $1/5^{\text{th}}$ of solar metallicity. As above, this would lead to similar or less overestimation of stellar mass and $[\text{N II}]$ contamination, respectively, and therefore mostly negligible modifications to our results.

Finally, it should be mentioned that the unknown contribution of AGNs at high stellar masses and redshift (see also Section 3.6) adds an additional uncertainty to the $\text{H}\alpha$ LF that can have similar impacts as inadequate $[\text{N II}]$ contamination corrections. Specifically, [Genzel et al. \(2014\)](#) find that two-thirds of their sample of $z \sim 1-3$ galaxies above $\log(M/M_\odot) = 10.9$ shows broad nuclear emission that could potentially be explained by the occurrence of an AGN. In this case, this would add almost a factor of three (~ 0.5 dex) uncertainty on the number counts at $\log(M/M_\odot) > 10.9$ (approximately $\log(L_{\text{H}\alpha}) > 43.0$ at $z = 1.5$). Hence, this will clearly dominate the uncertainties of the $\text{H}\alpha$ LF at high stellar masses (in comparison, the uncertainties from our model add up to about 10 – 20%). However, compared to the difference in the $\text{H}\alpha$ LF between a constant and our model based $[\text{N II}]$ correction, the uncertainty due to AGN contamination is a factor of two lower (0.5 dex compared to ~ 0.9 dex at $\log(L_{\text{H}\alpha}) = 43.0$ at $z = 1.5$, see Figure 10).

5. SUMMARY AND OUTLOOK

5.1. Summary

We present a parameterization of the $[\text{N II}]/\text{H}\alpha$ flux ratio as a function of stellar mass and redshift from $0 < z < 2.7$ for stellar masses of $8.5 < \log(M/M_\odot) \lesssim 11.0$. Our model encompasses the shift in the BPT locus defined by observed high-redshift data and the dependence of stellar mass on the BPT diagram on the $[\text{N II}]/\text{H}\alpha$ and $[\text{O III}]/\text{H}\beta$ emission line ratios from local galaxies. Our description is easily applicable to simulations for modeling $[\text{N II}]$ emission, current low-resolution grism and narrow-band observations to derive intrinsic $\text{H}\alpha$ fluxes, and to forecast the $\text{H}\alpha$ emission line galaxy number counts of future surveys.

We find large variations in the total $[\text{N II}]$ flux contamination fraction at a fixed redshift due to its dependency on stellar mass. Hence, we emphasize three main implications on current data as well as future surveys.

- The use of a constant $[\text{N II}]$ flux contamination fraction over- and under-predicts the true $[\text{N II}]$ contamination mainly as a function of stellar mass and redshift. This can lead to severe mass and redshift dependent biases in the determination of the intrinsic $\text{H}\alpha$ LF as well as other physical parameters computed from it. For example, a constant $[\text{N II}]$ contamination of 29% overestimates the true value for galaxies at $\log(M/M_\odot) \lesssim 10$ at $z > 0.5$ by a factor of up to 3.
- Intrinsic $\text{H}\alpha$ emitter number counts based on current *HST* grism surveys assuming a constant $[\text{N II}]$ flux contamination fraction of 29% are likely overestimated by 0.9 dex (factors of 8) and more at observed $\log(L) > 43.0$ at $z = 1.5$. Hence, the extrapolation of the *observed* ($[\text{N II}] + \text{H}\alpha$) number counts from these studies to match future surveys such as *WFIRST* and *Euclid*, which probe different redshift and stellar mass distributions, requires a redshift and stellar mass dependent modeling of the $[\text{N II}]$ flux contamination fraction as presented here.
- The blending of $\text{H}\alpha$ and $[\text{N II}]$ leads to a mass- and redshift-dependent systematic bias in the redshift measurement for *Euclid*. Our preliminary simulations indicate a redshift bias $\Delta z/(1+z) \sim 0.04 - 0.10\%$ for the most massive galaxies. This leads to a systematic bias of 0.5 – 1.6%, depending on redshift, in the BAO scale measurement in the radial direction at 150 Mpc.

5.2. Outlook

To examine our results in the context of current galaxy formation theory we plan to compare our model predictions to the predictions from a semi-analytical galaxy formation model (for example the GALACTICUS model). Such a comparison would allow us to further investigate the dependence of the $[\text{N II}]/\text{H}\alpha$ ratio on additional intrinsic galaxy properties, including the sSFR, as well as help test the validity of our model for redshifts $z \gtrsim 3$.

We will deepen our study on the spectroscopic redshift measurement biases by using realistic grism simulations, by applying more accurate noise levels expected for *Euclid*, and by including statistically more detailed properties of the galaxies (such as varying physical size). In addition, the application of our model to large area mock catalogs would allow further examination of how a redshift bias will impact determination of the BAO peak position, as well as subsequent cosmological parameter estimation. In particular, we will study techniques for

correcting this redshift bias.

Finally, we stress that further observational follow-up is needed to tighten our model especially at the massive end. Specifically, a *WFC3* grism filler program targeting massive ($\log(M/M_\odot) > 11$), star forming galaxies at $z \gtrsim 2$ would be useful to understand the line ratios of massive galaxies as well as the contribution of broad-line emission and AGN.

The authors thank J. Colbert for helpful comments and discussions. We also thank the referee for the very useful comments, which helped to improve this paper. A.M. acknowledges sponsorship of a NASA Postdoctoral Program Fellowship. A.M. was supported by JPL, which is run under contract by California Institute of Technology for NASA.

APPENDIX

A. SQL COMMANDS FOR SDSS GALAXY SELECTION

In the following, we list the SQL commands that were used to retrieve our SDSS sample from <http://skyserver.sdss.org/dr12/en/tools/search/sql.aspx>.

```
SELECT top 200000
  p.ObjID, s.fiberID,
  p.ra, p.dec, s.z,
  p.modelMag_u, p.modelMag_g,
  p.modelMag_r, p.modelMag_i,
  p.modelMag_z,
  p.expRad_r, p.expRad_i,
  e.sfr_tot_p50,
  e.lgm_tot_p50,
  g.h.alpha_eqw,
  g.oii_3726_flux,
  g.oii_3726_flux_err,
  g.oii_3729_flux,
  g.oii_3729_flux_err,
  g.neiii_3869_flux,
  g.neiii_3869_flux_err,
  g.h.beta_flux,
  g.h.beta_flux_err,
  g.oiii_5007_flux,
  g.oiii_5007_flux_err,
  g.oii_6300_flux,
  g.oii_6300_flux_err,
  g.h.alpha_flux,
  g.h.alpha_flux_err,
  g.nii_6584_flux,
  g.nii_6584_flux_err,
  g.sii_6717_flux,
  g.sii_6717_flux_err,
  g.sii_6731_flux,
```

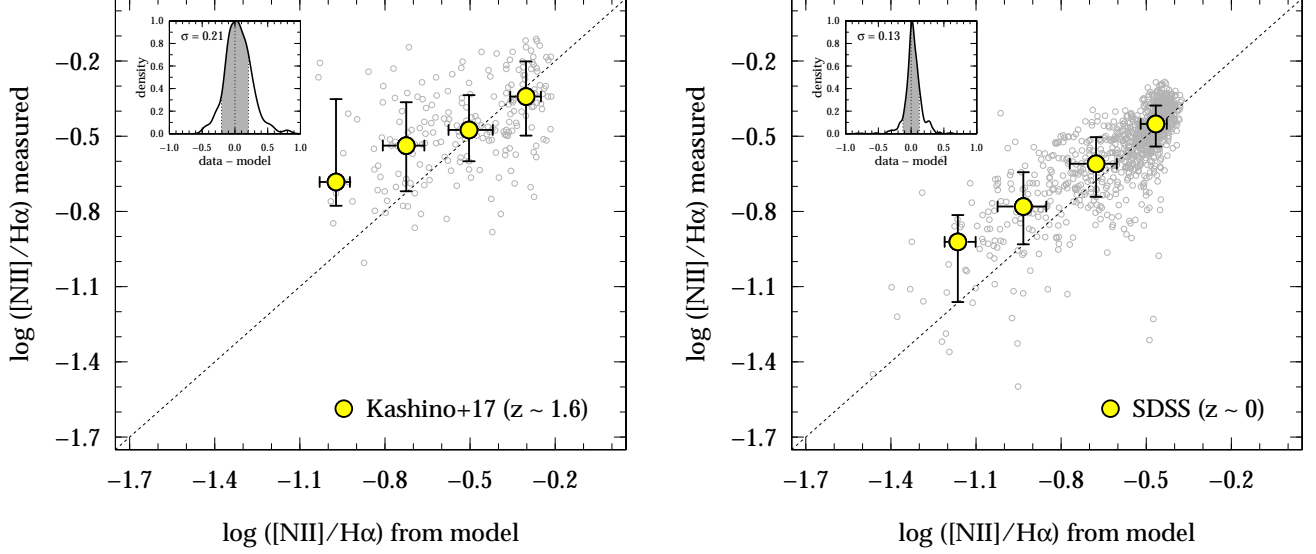



Figure B1. Same as Figure 6 but for $z \sim 1.6$ (left, Kashino et al. 2017) and SDSS galaxies (right).

```

g.sii_6731_flux_err
FROM photoObj p
JOIN specObj s ON s.bestObjID = p.objID
JOIN galSpecLine g ON g.specObjID = s.specObjID
JOIN galSpecExtra e ON e.specObjID = g.specObjID
WHERE
  s.class = 'galaxy'
  and e.bptclass = 1
  and s.zWarning = 0
  and g.h_alpha_flux / nullif(g.h_alpha_flux_err,0) > 5

```

B. SCATTER IN $[N II]/H\alpha$ RATIOS AT $Z \sim 1.6$ AND $Z \sim 0$

Figure 6 compares the true (i.e., measured) $[N II]/H\alpha$ ratios to the ones provided by our model at $z \sim 2.3$. In Figure B1, we show the same figure for $z \sim 1.6$ and $z \sim 0$ for reference. The scatter between true and model $[N II]/H\alpha$ ratios is 0.21 dex for $z \sim 1.6$ (similar to $z \sim 2.3$) and 0.13 dex for $z \sim 0$.

REFERENCES

- Acker, A., Köppen, J., Samland, M., & Stenholm, B. 1989, *The Messenger*, 58, 44
- Alam, S., Albareti, F. D., Allende Prieto, C., et al. 2015, *ApJS*, 219, 12
- Atek, H., Malkan, M., McCarthy, P., et al. 2010, *ApJ*, 723, 104
- Baldwin, J. A., Phillips, M. M., & Terlevich, R. 1981, *PASP*, 93, 5
- Benson, A. J. 2012, *NewA*, 17, 175
- Blake, C., & Glazebrook, K. 2003, *ApJ*, 594, 665
- Brammer, G. B., van Dokkum, P. G., Franx, M., et al. 2012, *ApJS*, 200, 13
- Brinchmann, J., Charlot, S., White, S. D. M., et al. 2004, *MNRAS*, 351, 1151
- Brinchmann, J., Pettini, M., & Charlot, S. 2008, *MNRAS*, 385, 769
- Cardamone, C., Schawinski, K., Sarzi, M., et al. 2009, *MNRAS*, 399, 1191
- Chabrier, G. 2003, *PASP*, 115, 763
- Colbert, J. W., Teplitz, H., Atek, H., et al. 2013, *ApJ*, 779, 34
- Daddi, E., Dickinson, M., Morrison, G., et al. 2007, *ApJ*, 670, 156
- Davidzon, I., Ilbert, O., Laigle, C., et al. 2017, *A&A*, 605, A70
- de los Reyes, M. A., Ly, C., Lee, J. C., et al. 2015, *AJ*, 149, 79
- Domínguez, A., Siana, B., Henry, A. L., et al. 2013, *ApJ*, 763, 145
- Dressler, A., Spergel, D., Mountain, M., et al. 2012, *ArXiv e-prints*, arXiv:1210.7809
- Elbaz, D., Daddi, E., Le Borgne, D., et al. 2007, *A&A*, 468, 33
- Elzhov, T. V., Mullen, K. M., Spiess, A.-N., & Bolker, B. 2016, *minpack.lm: R Interface to the Levenberg-Marquardt Nonlinear Least-Squares Algorithm Found in MINPACK, Plus Support for Bounds*, v.R package version 1.2-1, , , r package version 1.2-1
- Erb, D. K., Pettini, M., Steidel, C. C., et al. 2016, *ApJ*, 830, 52
- Erb, D. K., Shapley, A. E., Pettini, M., et al. 2006, *ApJ*, 644, 813
- Faisst, A. L. 2016, *ApJ*, 829, 99
- Garn, T., & Best, P. N. 2010, *MNRAS*, 409, 421

- Gehrels, N., Spergel, D., & WFIRST SDT Project. 2015, in *Journal of Physics Conference Series*, Vol. 610, *Journal of Physics Conference Series*, 012007
- Genzel, R., Förster Schreiber, N. M., Rosario, D., et al. 2014, *ApJ*, 796, 7
- Green, J., Schechter, P., Baltay, C., et al. 2012, *ArXiv e-prints*, arXiv:1208.4012
- Greis, S. M. L., Stanway, E. R., Davies, L. J. M., & Levan, A. J. 2016, *MNRAS*, 459, 2591
- Guzzo, L., Pierleoni, M., Meneux, B., et al. 2008, *Nature*, 451, 541
- Henry, A., Scarlata, C., Domínguez, A., et al. 2013, *ApJL*, 776, L27
- Hirschmann, M., Charlot, S., Feltre, A., et al. 2017, *MNRAS*, 472, 2468
- Hu, E. M., Cowie, L. L., Kakazu, Y., & Barger, A. J. 2009, *ApJ*, 698, 2014
- Ilbert, O., McCracken, H. J., Le Fèvre, O., et al. 2013, *A&A*, 556, A55
- Kaiser, N. 1987, *MNRAS*, 227, 1
- Kashino, D., Silverman, J. D., Sanders, D., et al. 2017, *ApJ*, 835, 88
- Kauffmann, G., Heckman, T. M., White, S. D. M., et al. 2003, *MNRAS*, 341, 33
- Kennicutt, Jr., R. C. 1998, *ARA&A*, 36, 189
- Kewley, L. J., Dopita, M. A., Sutherland, R. S., Heisler, C. A., & Trevena, J. 2001, *ApJ*, 556, 121
- Kewley, L. J., & Ellison, S. L. 2008, *ApJ*, 681, 1183
- Kewley, L. J., Maier, C., Yabe, K., et al. 2013, *ApJL*, 774, L10
- Kroupa, P. 2001, *MNRAS*, 322, 231
- Laureijs, R., Amiaux, J., Arduini, S., et al. 2011, *ArXiv e-prints*, arXiv:1110.3193
- Lee, J. C., Ly, C., Spitler, L., et al. 2012, *PASP*, 124, 782
- Lilly, S. J., Carollo, C. M., Pipino, A., Renzini, A., & Peng, Y. 2013, *ApJ*, 772, 119
- Ly, C., Lee, J. C., Dale, D. A., et al. 2011, *ApJ*, 726, 109
- Ly, C., Malhotra, S., Malkan, M. A., et al. 2016a, *ApJS*, 226, 5
- Ly, C., Malkan, M. A., Kashikawa, N., et al. 2012, *ApJ*, 757, 63
- Ly, C., Malkan, M. A., Rigby, J. R., & Nagao, T. 2016b, *ApJ*, 828, 67
- Ly, C., Rigby, J. R., Cooper, M., & Yan, R. 2015, *ApJ*, 805, 45
- Maier, C., Ziegler, B. L., Lilly, S. J., et al. 2015, *A&A*, 577, A14
- Maiolino, R., Nagao, T., Grazian, A., et al. 2008, *A&A*, 488, 463
- Masters, D., Faisst, A., & Capak, P. 2016, *ApJ*, 828, 18
- Masters, D., McCarthy, P., Siana, B., et al. 2014, *ApJ*, 785, 153
- Mehra, V., Scarlata, C., Colbert, J. W., et al. 2015, *ApJ*, 811, 141
- Merson, A., Wang, Y., Benson, A., et al. 2018, *MNRAS*, 474, 177
- Nakajima, K., & Ouchi, M. 2014, *MNRAS*, 442, 900
- Noeske, K. G., Weiner, B. J., Faber, S. M., et al. 2007, *ApJL*, 660, L43
- Oke, J. B. 1974, *ApJS*, 27, 21
- Orsi, A., Baugh, C. M., Lacey, C. G., et al. 2010, *MNRAS*, 405, 1006
- Pettini, M., & Pagel, B. E. J. 2004, *MNRAS*, 348, L59
- Pozzetti, L., Hirata, C. M., Geach, J. E., et al. 2016, *A&A*, 590, A3
- Salim, S., Lee, J. C., Davé, R., & Dickinson, M. 2015, *ApJ*, 808, 25
- Salim, S., Lee, J. C., Ly, C., et al. 2014, *ApJ*, 797, 126
- Salpeter, E. E. 1955, *ApJ*, 121, 161
- Sanders, R. L., Shapley, A. E., Kriek, M., et al. 2017, *ArXiv e-prints*, arXiv:1711.00224
- Savaglio, S., Glazebrook, K., Le Borgne, D., et al. 2005, *ApJ*, 635, 260
- Schreiber, C., Pannella, M., Elbaz, D., et al. 2015, *A&A*, 575, A74
- Seo, H.-J., & Eisenstein, D. J. 2003, *ApJ*, 598, 720
- Shapley, A. E., Reddy, N. A., Kriek, M., et al. 2015, *ApJ*, 801, 88
- Sobral, D., Best, P. N., Matsuda, Y., et al. 2012, *MNRAS*, 420, 1926
- Sobral, D., Smail, I., Best, P. N., et al. 2013, *MNRAS*, 428, 1128
- Sobral, D., Best, P. N., Geach, J. E., et al. 2009, *MNRAS*, 398, 75
- Speagle, J. S., Steinhardt, C. L., Capak, P. L., & Silverman, J. D. 2014, *ApJS*, 214, 15
- Spergel, D., Gehrels, N., Baltay, C., et al. 2015, *ArXiv e-prints*, arXiv:1503.03757
- Stanway, E. R., Eldridge, J. J., Greis, S. M. L., et al. 2014, *MNRAS*, 444, 3466
- Steidel, C. C., Rudie, G. C., Strom, A. L., et al. 2014, *ApJ*, 795, 165
- Strom, A. L., Steidel, C. C., Rudie, G. C., et al. 2017, *ApJ*, 836, 164
- Tomczak, A. R., Quadri, R. F., Tran, K.-V. H., et al. 2016, *ApJ*, 817, 118
- Tremonti, C. A., Heckman, T. M., Kauffmann, G., et al. 2004, *ApJ*, 613, 898
- van Dokkum, P., Brammer, G., Momcheva, I., et al. 2013, *ArXiv e-prints*, arXiv:1305.2140
- Vavrek, R. D., Laureijs, R. J., Lorenzo Alvarez, J., et al. 2016, in *Proc. SPIE*, Vol. 9911, *Modeling, Systems Engineering, and Project Management for Astronomy VI*, 991105
- Villar, V., Gallego, J., Pérez-González, P. G., et al. 2008, *ApJ*, 677, 169
- Wang, Y. 2008, *JCAP*, 5, 021
- York, D. G., Adelman, J., Anderson, Jr., J. E., et al. 2000, *AJ*, 120, 1579