Disentangling the physical parameters of gaseous nebulae and galaxies

Daichi Kashino,1* Akio K. Inoue,2
1 Department of Physics, ETH Zürich, Wolfgang-Pauli-strasse 27, CH-8093 Zürich, Switzerland
2 Department of Environmental Science and Technology, Faculty of Design Technology, Osaka Sangyo University, S-1-1, Nakagaito, Daito, Osaka 574-8550, Japan

Accepted XXX. Received YYY; in original form ZZZ

Abstract

We present an analysis to disentangle the connection between physical quantities that characterize the conditions of ionized H II regions – metallicity (Z), ionization parameter (U), and electron density (n_e) – and the global stellar mass (M_*) and specific star formation rate (sSFR = SFR/M_*) of the host galaxies. We construct composite spectra of galaxies at 0.027 ≤ z ≤ 0.25 from Sloan Digital Sky Survey, separating the sample into bins of M_* and sSFR, and estimate the nebular conditions from the emission line flux ratios. Specially, metallicity is estimated from the direct method based on the faint auroral lines [O III]λ4363 and [O II]λ7320,7330. The metallicity estimates cover a wide range from 12 + log O/H ~ 7.6-8.9. It is found that these three nebular parameters all are tightly correlated with the location in the M_*–sSFR plane. With simple physically-motivated ansatze, we derive scaling relations between these physical quantities by performing multi regression analysis. In particular, we find that U is primarily controlled by sSFR, as U ∝ sSFR^{0.43}, but also depends significantly on both Z and n_e. The derived partial dependence of U ∝ Z^{-0.36} is weaker than the apparent correlation (U ∝ Z^{-1.55}). The remaining negative dependence of U on n_e is found to be U ∝ n_e^{-0.29}. The scaling relations we derived are in agreement with predictions from theoretical models and observations of each aspect of the link between these quantities. Our results provide a useful set of equations to predict the nebular conditions and emission-line fluxes of galaxies in semi-analytic models.

Key words: galaxies: ISM – HII regions

1 INTRODUCTION

Ionized nebulae in star-forming galaxies show a wide range of physical conditions. The properties of H II regions are characterized by a set of physical quantities, including density (or pressure), gas-phase metallicity (hereafter metallicity or Z), and the ionization parameter U. These parameters are connected through various fundamental physical processes taking place in the gas, such as photoionization, collisional excitation, and radiative cooling. Interestingly, the properties of the interstellar medium (ISM) appear also strongly correlated with the global properties of the host galaxies, such as stellar mass (M_*) and star formation rate (SFR). This implies that the history and the on-going activities of the galaxies affect the conditions and the processes occurring in the resident nebulae. Disentangling the connections between these parameters is thus essentially important to understand the astrophysical processes govern the evolution and behavior of baryons in the galaxies.

The correlation between M_* and Z (the “mass–metallicity” (MZ) relation) has been well established at low redshifts (z ~ 0.1) based on a large sample from the Sloan Digital Sky Survey (SDSS) (e.g., Tremonti et al. 2004), and also at higher redshifts (z ≥ 1) as well (e.g., Erb et al. 2006; Zahid et al. 2011; Yabe et al. 2012; Zahid et al. 2013, 2014; Sanders et al. 2015; Kashino et al. 2017, 2018). Ellison et al. (2008) pointed out the presence of an anticorrelation between Z and SFR at fixed M_*, which has then further explored by many authors (e.g., Lara-López et al. 2010; Yates et al. 2012; Lara-López et al. 2013; Mannucci et al. 2010; Andrews & Martini 2013). In particular, Mannucci et al. (2010) proposed the so-called fundamental metallicity relation (FMR) in the M_*–SFR–Z space. In these studies, the scatter of the MZ relation is explained by the result of varia-

* E-mail: kashinod@phys.ethz.ch

© 2015 The Authors
tions in the inflowing rate of circumgalactic medium into the system; inflowing gas drives formation of new stars while di-
luting the gas-phase metallicity. In this context, it has also
been argued that the relation between $M_*$, $Z$ and the gas
content is more fundamental, rather than SFR (e.g., Both-
well et al. 2013, 2016a,b). However, the shape of the FMR
and its redshift evolution are still under debate (e.g., Yates
et al. 2012; Andrews & Martini 2013; Kashino et al. 2016;
Telford et al. 2016).

A difficulty in measuring metallicity is that in most
cases we cannot use the “direct” method and thus must
rely on the so-called “strong-line” method. The direct method
utilizes the flux ratio between auroral and strong lines, such
as $\lambda4363/\lambda5007$, to measure the electron tempera-
ture $T_e$ of the ionized gas. Once the temperature is
determined, the absolute oxygen abundance can be esti-
mated from the $T_e$-dependent ratios of strong lines, such as
$\lambda4372$ and $\lambda5007$, relative to the $H\beta$ line. On the
other hand, the strong-line method is a technique that uses a
ratio between strong line ratios (e.g., $R_{33} = (\lambda4372 +$
$\lambda4959, 5007)/H$β) and metallicity, calibrated theoretically
or observationally (e.g., Kewley & Dopita 2002; Pettini
& Pagel 2004; Nagao et al. 2006; Maiolino et al. 2008;
Kewley & Ellison 2008; Marino et al. 2013). Since the de-
tection of the auroral lines is usually challenging due to
their faintness, especially at high metallicity, the strong-line
method has been routinely used. However, the strong-line
ratios are functions not only of metallicity, but also of other
parameters such as the ionization parameter. Such system-
tic uncertainties could result in a misleading comparison
between different galaxy populations, e.g., low and high red-
shifts galaxies (e.g., Kewley et al. 2015; Telford et al. 2016).

The electron density $n_e$ of the ionized gas is routinely
estimated from the intensity ratio of the $[S\,\text{II}]/\lambda6717,6731$
or $[O\,\text{III}]/\lambda4372,3729$ doublet lines. The typical electron densities
of $H\,\text{II}$ regions in low redshift galaxies have been measured to
be $n_e \sim 50$–$120$ cm$^{-3}$ (Brinchmann et al. 2008). The densities
in high redshift ($z \gtrsim 1$) galaxies, however, have found to
be elevated relative to low redshifts, ranging from $10^2$
$\gtrsim$ to a few $10^3$ cm$^{-3}$ (e.g., Shirazi et al. 2014; Masters et
al. 2014; Shimakawa et al. 2015; Sanders et al. 2016; Kashino
et al. 2017; Kaasinen et al. 2017). In particular, Shimakawa
et al. (2015) found that $n_e$ is correlated with SFR and/or SFR
surface density. Such correlation could naturally explain the
observed evolution in the typical $n_e$ with redshift. Higher
ionization parameters are often measured in high redshift and/or high sSFR galaxies (e.g., Hainline et al. 2009; Nakajima
& Ouchi 2014; Shapely et al. 2015; Strom et al. 2017). These studies routinely use the $[O\,\text{III}]/H\beta$ ratio as a proxy of $U$. Its effective
ness has been established theoretically (e.g., Dopita et al.
2000; Kewley & Dopita 2002), while it is known that the
line ratio is also sensitive to metallicity at fixed $U$. Pioneering
studies (Evans & Dopita 1985; Dopita & Evans 1986) claimed the presence of a strong inverse correlation between $U$
and $Z$ from comparisons between their theoretical pho-
toionization model and observations. Dopita et al. (2006b)
have theoretically shown that stellar ionizing photon flux il-
luminating the gas reduces at higher metallicity because of
increasing absorption due to a higher opacity in the stel-
lar winds, and due to more efficient conversion from radia-
tion energy into kinetic energy of the winds. Recent observa-
tions of high redshift galaxies (hence higher sSFR and lower
$Z$) have also found anticorrelations between $U$ and $Z$ via
the strong-line method (Nakajima & Ouchi 2014; Onodera
et al. 2016). In contrast, however, Kaasinen et al. (2018) re-
cently concluded that the ionization parameter $U$ is directly
linked to sSFR, while little or probably not with metallicity,
from comparison between samples of local and $z \sim 1.5$ star-
forming galaxies. It remains under debate whether $U$ and $Z$
is intrinsically correlated or not, and if it is the case, what
physical mechanisms work behind the relationship.

The physical quantities that characterize the conditions of
galaxies and the ionized nebulae are connected with each
other in complicated ways. As mentioned above, each as-
pect of this big puzzle has been studied by many authors.
Now we need to disentangle their connection from a view-
point beyond focusing on a correlation between two or three
quantities simultaneously. In this paper, we will focus on
the integrated global properties of galaxies (i.e., $M_*$ and
SFR) and the basic parameters that define the conditions of
the ionized gas locally: metallicity $Z$ (O/H), ionization
parameter $U$, and electron density $n_e$. To achieve the goal,
accurate and precise measurements of oxygen abundance is
essentially important, and thus we use the direct method
that requires a detection of faint auroral lines of oxygen
ions such as $[O\,\text{III}]/\lambda4363$, and $[O\,\text{III}]/\lambda5809, 5007$.
To overcome the faintness of these lines, we employ techniques presented in Andrews & Martini (2013) and Curti
et al. (2017), who stacked thousands of galaxy spectra from SDSS. This en-
ables us to use the direct method even for low SFR and
metal-rich galaxies.

This paper is organized as follows. We describe our sam-
ple of galaxies and the methods of spectral stacking and
emission-line measurement in Section 2. Derivation of the
physical properties of ionized nebulae (i.e., $Z$, $U$ and $n_e$)
indicated in Section 3. In Section 4, we describe the set
of simple, physically-motivated ansätze for the relations be-
tween these quantities, and analyses of the correlations of
the stacked measurements. We present results in Section 5,
and discussions in Section 6. Summary and the conclusions
are given in Section 7. Throughout the paper, the stellar
masses and SFRs are computed by using a Kroupa (2001)
universal initial mass function (IMF). The metallicity $Z$
denotes the gas-phase oxygen abundance, which is defined as
the abundance ratio of oxygen atoms to $10^{12}$ hydrogen
atoms, as $12 + \log(O/H)_{\odot} = 8.69$, or $Z_\odot = 0.014$ for metallic mass fraction (Asplund et al. 2009).

2 DATA AND METHOD

2.1 Sample selection

The parent sample of galaxies used for this study was con-
structed following the method described in Andrews & Mar-
tini (2013). We here present a brief overview of the sample
selection. We utilize the public MPA-JHU catalog, which
contains $\sim 930,000$ galaxies from the SDSS Data Release 7
(Abazajian et al. 2009). The MPA-JHU catalog contains for
each galaxy the estimates of stellar mass and total SFR from
the aperture-corrected H$\alpha$ luminosity, with a Kroupa IMF
(Brinkmann et al. 2004; Salim et al. 2007).
The galaxies are limited to have a spectroscopic redshift between 0.027 ≤ z ≤ 0.25 with an uncertainty of σz < 0.001. For this redshift range, both the most blue [O\textsc{ii}] λ3727 line and the most red [O\textsc{ii}] λλ7320,7330 lines fall within the observed wavelength range of 3800–9200 Å. We further imposed on the sample the detection of Hβ, Hα, and [N\textsc{ii}] λ6584 all at S/N ≥ 5. Possible AGNs were removed by using the Baldwin-Phillips-Terlevich (BPT; Baldwin et al. 1981; see also Veilleux & Osterbrock 1987) diagram with an empirical classification curve derived by Kauffmann et al. (2003). Objects with possible poor constraints of photometry were removed, based on the warning flags (DEBLEND_NOPEAK or DEBLENDED_AT_EDGE), or visual inspection. The final sample consists of 194,005 galaxies and the median redshift is z = 0.081. Figure 1 shows the distribution of our selected galaxies in the M∗–sSFR plane by contours. The ridge line of the contours is well matched to the so-called main sequence of star-forming galaxies derived by Elbaz et al. (2007).

2.2 Spectral stacking

To determine the metallicity via the direct method, it is required to measure the fluxes of the faint auroral emission lines of oxygen ions, [O\textsc{iii}]λ4363 and [O\textsc{ii}]λλ7320,7330. However, the intensity of these lines rapidly declines below the noise level of the single spectrum as the metallicity increases. To probe a wide range of the parameter space, we improved the S/N ratios by creating composite spectra of galaxies.

Our motivation is to disentangle the connections between physical parameters characterizing local conditions in ionized nebulae and global properties of host galaxies. The stellar mass and SFR are both extensive variables, and hence are naturally correlated with each other. The interpretation would be more clarified by using M∗ as a single indicator of the size of the system, while taking sSFR, an intensive variable, rather than SFR. We therefore separated the sample galaxies into the grid in the M∗–sSFR plane with binsizes of 0.20 dex in M∗ and 0.25 dex in sSFR. Figure 1 shows the numbers of galaxies in each box that represents a single stack.

We utilized the public spectral data from the SDSS DR7, which have been processed with the SDSS Spec2D pipeline (Stoughton et al. 2002). Before creating composite spectra, the individual spectra were 1) corrected for the Milky Way reddening with the extinction values from Schlegel et al. (1998) and adopting a Cardelli et al. (1989) extinction law, 2) shifted to the rest frame based on the spectroscopic redshift, 3) resampled with a universal wavelength grid with a spacing of ∆ log10λ = 5 × 10−5, and 4) normalized by the Hβ flux. The individual spectra have then been co-added by taking the mean flux between the 25th and the 75th percentiles at each wavelength grid, similarly to Curti et al. (2017).

2.3 Stellar continuum subtraction

To measure the fluxes of the nebular emission lines, it is important to fit and subtract the stellar component from the composite spectra. In particular, the Balmer lines and [O\textsc{iii}]λ4363 in the vicinity of Hγ, are highly impacted by
the stellar atmospheric absorption. For the stellar continuum subtraction, we generated a synthetic spectrum for each composite spectrum. We performed the fit using the IDL version of the pPXF package (Cappellari 2017) with the built-in library of single stellar population (SSP) spectral templates. The library includes 150 model spectra based on the MILES library of stellar templates (Sánchez-Blázquez et al. 2006), computed for 25 different ages and 6 different metallicities using the code presented in Vazdekis et al. (2010). The templates cover a wavelength range of 3540–7410 Å with a resolution of 2.51 Å (full width at half maximum; FWHM).

We performed the stellar template fitting first using the entire spectral range between 3650 Å and 7360 Å to determine the best-fit synthesized spectrum, generated as an arbitrary superposition of the 150 simple stellar population spectra in the library. We then fit the single best-fit template in each subrange of a few hundred angstroms around or low-order polynomial fitting around the lines. Mean- 

Figure 2 shows that the absorption features are prominent at wavelengths of the Balmer series lines. These Balmer absorption lines arise from the absorption of photons by hydrogen excited at the n = 2 level in the photospheres of stars. This effectively reduces the observed fluxes of the Balmer lines if one simply excludes the continuum based on a linear or low-order polynomial fitting around the lines. Meanwhile, it is not always possible to perform the stellar continuum subtraction based on the stellar population synthesis fitting, as done in this paper. Therefore, it would be useful to derive a simple empirical relation between the effect of Balmer absorption and the galaxy properties. In Appendix A, we present the amount of the Balmer absorption as a function of $M_*$ and sSFR, and provide empirical relations for absorption correction.

2.4 Line flux measurement

We performed emission-line fitting using the MPFIT package for IDL (Markwardt 2009) to measure the fluxes of the emission lines in the stellar-subtracted composite spectra. Each emission line was fit with a single gaussian profile. Note that all the doublet lines \([\text{O \ II}] \lambda\lambda 3726,3729, [\text{O \ III}] \lambda \lambda 4959,5007, [\text{N \ II}] \lambda \lambda 6548,6584, [\text{S \ II}] \lambda \lambda 6717,6731, \text{and} \ [\text{O \ II}] \lambda \lambda 7320,7330\) were fit with independent amplitudes. We disregarded lines that were poorly fit with low S/N (<5), which is computed for each line as a ratio of measured flux and the uncertainty based on the errors on the fitting parameters returned by the MPFIT procedure. The measured fluxes of all emission lines were then corrected for dust extinction by adopting a Cardelli et al. (1989) extinction law. The absolute level of extinction was estimated based on the observed Hα/Hβ ratio by assuming the intrinsic values of Hα/Hβ given in Osterbrock & Ferland (2006, Table 4–4) for Case B recombination with $n_e = 10^2$ cm$^{-3}$. The intrinsic Hα/Hβ ratio is very insensitive to the electron density if $n_e < 10^4$ cm$^{-3}$, while being slightly dependent on the electron temperature. For the range of 7000 $< T_e(^{O\*}) < 20000$ K of our data, the Hα/Hβ ratio could vary within $\pm 5\%$, increasing as $T_e$ decreases. Therefore we iteratively estimated the extinction as follows. We first estimate the extinction by assuming the intrinsic Hα/Hβ ratio of 2.86 for $T_e = 10^4$ K and derive the electron temperature from the extinction-corrected fluxes. The estimated $T_e$ is then used to modify the intrinsic Hα/Hβ ratio, and to re-compute the level of extinction. We then repeat the extinction correction of all the line fluxes and derivation of the physical quantities. The change in the resulting log O/H after the iteration is correlated with log O/H, and is negligible at log O/H $\lesssim 8.2$ while gradually being larger at higher Z. The final estimates after the iteration are reduced by $\approx 0.08$ dex at 12 + log O/H = 8.8 relative to the initial estimates with the intrinsic Hα/Hβ of 2.86 assumed.

As pointed out by Andrews & Martini (2013), an emission feature was found at $\approx 4360$ Å, which is blended with the [O III]$\lambda\lambda 4363$ auroral line. Curti et al. (2017) suggested that this feature is reasonably associated with emission lines from Fe II ions. Following their treatment, we simultaneously fit the three line components of the possible [Fe II] lines that are centered at 4357.7, 4358.8, and 4360.6 Å, in addition to Hα and [O III]$\lambda\lambda 4363$. As shown by Curti et al. (2017), we found that the contribution from the possible [Fe II] lines increases as metallicity increases. Following Curti et al. (2017), we disregarded possible detection of [O III]$\lambda\lambda 4363$ when the flux measured for the $\lambda\lambda 4360$ feature is $> 0.5$ times the flux measured for [O III]$\lambda\lambda 4363$.

3 PHYSICAL PARAMETERS

In this work, we investigate the connection among the five physical quantities: two galaxy global properties, $M_*$ and sSFR, and three nebular parameters characterizing the conditions of the H II regions, metallicity Z (12 + log O/H), ionization parameter $U$, and electron density $n_e$. These nebular properties were derived based on the stacked measurements of the emission-line fluxes as described in the following subsections. For this purpose, we used an package of IDL routines called improp. This software offers functions similar to the commonly used IRAF/TEMDEN package. To evaluate the nebular parameters and their uncertainties we generated 200 random realizations for each set of the line fluxes measured from each stack while incorporating the flux errors with Gaussian distributions. We then took the median and the standard deviation for each parameter estimate from the resultant realizations.

3.1 Electron temperature and density

The electron temperatures $T_e(^{O\*})$ and $T_e(^{O^{++}})$ were determined from the line ratios $[\text{O \ II}] \lambda\lambda 3726,3729/[\text{O \ II}] \lambda \lambda 7320,7330$ and $^{[1]}$

\begin{linenomath*}
1 \text{The code has been developed by John Moustakas et al. and is available here: https://github.com/moustakas/improp}
\end{linenomath*}
either the \([\text{O}^\text{iii}]\) ratios given by Curti et al. (2017, see Equation 3). In this work, we utilized the empirical relation using the strong-line \([\text{O}^\text{ii}]\) stacks where we failed to detect the \([\text{O}^\text{iii}]\) (see Equation (14) of Pilyugin et al. 2006). Similarly, for \([\text{O}^\text{iii}]\) derived by Pilyugin et al. (2006), following the method of \([\text{O}^\text{ii}]\) stacks without the \([\text{O}^\text{iii}]\) failed to detect \([\text{O}^\text{iii}]\) (see Equation (14) of Pilyugin et al. 2006). For these stacks, we estimated the temperature \(T_e(\text{O}^\text{++})\) by utilizing the so-called flux-flux (H) relation derived by Pilyugin et al. (2006), following the method in Curti et al. (2017). This empirically-calibrated relation predicts the \([\text{O}^\text{iii}]\) \(4363\) flux relative to \(\text{H}^\beta\) from the strong-line ratios, \([\text{O}^\text{iii}]\) \(4959,5007/\text{H}^\beta\) and \([\text{O}^\text{ii}]\) \(3727/\text{H}^\beta\) (see Equation (14) of Pilyugin et al. 2006). Similarly, for \([\text{O}^\text{iii}]\) stacks where we failed to detect the \([\text{O}^\text{ii}]\) \(4320,7330\) lines, we utilized the empirical relation using the strong-line ratios given by Curti et al. (2017, see Equation 3). In this work, we use the stacks where at least we successfully detect either the \([\text{O}^\text{iii}]\) or \([\text{O}^\text{ii}]\) auroral lines, and disregarded stacks having neither detections of these. We emphasize that, at the metallicity range where the \([\text{O}^\text{ii}]\) \(4363\) line is undetected, the total oxygen abundance is dominated by the singly-ionized oxygen \(\text{O}^+\). This is also the case for the opposite situation, i.e., the oxygen abundance of the stacks without the \([\text{O}^\text{ii}]\) \(4320,7330\) detection is dominated by \(\text{O}^{++}\). Therefore, the effects of the possible uncertainties from the use of these empirical relations are negligible in the total oxygen abundance estimates.

The flux ratio of the auroral and strong lines depends not only on the electron temperature, but also weakly on the electron density. The electron density can be estimated from the \([\text{Si}^\text{ii}]\) \(6717/\text{[Si}^\text{ii}]\) ratio, with a small dependence on the temperature \(T_e(S^+))\). We therefore iteratively estimated both \(T_e\) and \(n_e\) assuming \(T_e(S^+) = T_e(O^+)\). The \([\text{Si}^\text{ii}]\) doublet ratio saturates at \(n_e < 10^3\) cm\(^{-3}\). As a result, the errors in getting lower for \(n_e\), or cannot be constrained for some cases. In our analysis, we set the lower limit of the \(n_e\) range to be \(n_e = 5 \text{ cm}^{-3}\). We substituted \(\log n_e/(\text{cm}^{-3}) = 0.7 \pm 0.7\) for the value of three stacks for which the observed \([\text{Si}^\text{ii}]\) ratio is above the saturation limit. The derived \(n_e\) of the stacks ranges from \(17 \text{ to } 300 \text{ cm}^{-3}\) (the central 90 percentiles), with the median \(n_e = 76 \text{ cm}^{-3}\). Note that this range and the median value do not represent the electron densities of the individual SDSS galaxies since the number of galaxies in the stacks are hugely different from one to another bin. Lastly, we note that the spectral resolution of the SDSS spectra are not sufficient to obtain reasonable estimates of \(n_e\) from the \([\text{O}^\text{iii}]\) \(4372,3729\) doublet ratio.

### 3.2 Direct abundance determination

The ionic oxygen abundance \(\text{O}^+/\text{H}^+\) and \(\text{O}^{++}/\text{H}^+\) were determined from \([\text{O}^\text{ii}]\) \(4372,3727/\text{H}^\beta\), and \([\text{O}^\text{iii}]\) \(4959,5007/\text{H}^\beta\) as well as from the temperature estimates, \(T_e(\text{O}^+)\) and \(T_e(\text{O}^{++})\), respectively. The procedure in the impro package includes a slight dependence on \(n_e\) as well. As described above, if we have constraint only either \(T_e(\text{O}^+)\) or \(T_e(\text{O}^{++})\), we estimated...
the other using the empirical relations based on the strong lines. We remind that, for such cases, the contribution of the ions whose auroral line(s) is not detected is subdominant. We then assumed that the total oxygen abundance is the sum of the ionic abundances of these two ions,

$$\frac{O_2}{H} = \frac{O^+}{H^+} + \frac{O^{++}}{H^{++}}.$$  \hspace{1cm} (1)

We neglected the minimal contribution of neutral atoms from the flux ratio of the same species at different ionization states (i.e., $O^+$ and $O^{++}$). The contribution of $O^{++}$ and higher ionization ions was also neglected, whose ionization potential is 54.9 eV or higher, thus there is little chance to be ionized by stellar radiation.

### 3.3 Ionization parameter

The ionization parameter is a key quantity in this work. It is locally defined as the ratio of ionizing photon flux $S_{\text{ion}}$ and the hydrogen density $n_H$: $q_{\text{ion}} = S_{\text{ion}}/n_H$. The ionization parameter can be made dimensionless by dividing by the speed of light as follows:

$$U = \frac{q_{\text{ion}}}{c} = \frac{n_e}{n_H}$$  \hspace{1cm} (2)

where $n_e$ is the volume density of the ionizing photons. Throughout the paper, we use the dimensionless $U$, instead of $q_{\text{ion}}$. We note that the electron density $n_e$ is almost equivalent to $n_H$ in a usual H II region.

We estimated $U$ from a sensitive indicator, the flux ratio of $[\text{O III}]_{\lambda 5007}/[\text{O II}]_{\lambda 3726,3729}$ (Dopita et al. 2000; Kewley & Dopita 2002). Although the flux ratio is also sensitive to metallicity, we can resolve this degeneracy using the metallicity estimate from the direct method (Section 3.2). The $[\text{O III}]_{\lambda 5007}/[\text{O II}]_{\lambda 3726,3729}$ ratio also slightly varies with $n_e$ at fixed $U$ and $Z$, as well. We consider this effect by using the $n_e$ estimate from the $[\text{S II}]_{\lambda 6716,6731}$ doublet ratio.

We determined the ionization parameters for each stack by interpolating the flux ratios computed for grids of $Z$, $U$, and $n_H$, assuming $n_H = n_e$. We used the CLOUDY (C17.00) photoionization code to compute the emission-line fluxes for given nebular conditions. For the gas phase, we adopted the solar chemical composition, though the chemical composition does little affect the estimates of $U$ since it is estimated from the flux ratio of the same species at different ionization levels (i.e., $O^+$ and $O^{++}$). Photoionization calculations were executed in the pseudo plane-parallel geometry (i.e., inner radius $\gg$ thickness of the ionized shell) for the following values of $Z$, $U$, and $n_H$.

$$12 + \log(O/H) = \begin{cases} 7.49, 7.69, 7.89, 8.09, 8.29, 8.39, \\ 8.49, 8.59, 8.69, 8.798, 9.19 \\ \end{cases}$$

$$\log U = \begin{cases} -4.00, -3.75, -3.50, -3.25, -3.00, \\ -2.75, -2.50, -2.25, -2.00 \\ \end{cases}$$

$$\log n_H/(\text{cm}^{-3}) = \begin{cases} 0.5, 1.0, 1.5, 2.0, 2.5, 3.0 \\ \end{cases}$$

Ionizing spectra used in our CLOUDY calculations were produced by the Starburst99 population synthesis code (Leitherer et al. 1999), adopting an constant SFR history with a Kroupa (2001) IMF (0.1–100 $M_{\odot}$) and the Geneva high mass-loss evolutionary tracks with five values of stellar metallicity: $Z_e = 0.001, 0.004, 0.008, 0.020$, and 0.040, corresponding to $12 + \log(O/H) = 7.54–9.15$. We adopted a set of spectra at 10 Myr, and, for a given gas-phase $Z$, the input ionizing spectrum was produced by interpolating the spectral set at $Z_e = Z$ to match the stellar and gas-phase metallicities. The ionization parameter $U$ is then derived from a $U$–$[\text{O III}]_{\lambda 5007}/[\text{O II}]_{\lambda 3726,3729}$ relation obtained by interpolating the three-dimensional grid data at the estimated $n_H = n_e$ and log O/H. The estimates of $U$ range from log $U = −3.5$ to −2.4.

### 3.4 Derived quantities

We successfully estimated the metallicity for 138 stacks, of which we detected both $[\text{O III}]_{\lambda 4363}$ and $[\text{O II}]_{\lambda 7320,7330}$ for 75 stacks, only $[\text{O II}]_{\lambda 7320,7330}$ for 61 stacks, and only $[\text{O II}]_{\lambda 4363}$ for 2 stacks. Figure 1 shows the derived metallicity by colors in each stacking bin in the $M_*,\text{ssSFR}$ plane. Here a caveat is that the stacks with successful metallicity estimation are biased towards lower $M_*$ and higher SFR, resulting in being biased to the population above the main sequence (dotted line; Elbaz et al. 2007). This is because the intensity of auroral lines rapidly declines with increasing $Z$ as well as with decreasing SFR. As a result, the detection of the auroral lines has often failed at high $M_∗(\gtrsim 10^{10.6} M_{\odot})$, or below the main sequence even if thousands of spectra are stacked. In contrast, the detection succeeded at lower $M_*$ and high SFRs even with a few (or only a single) galaxies in the stacks. However, we would emphasize that the successful stacks cover the whole range of the metallicity of interest ($7.6 \lesssim 12 + \log(O/H) \lesssim 8.9$), proved by most studies of the MZ relation at low and higher redshifts (references in Section 1), including both regimes of extremely metal-poor galaxy populations (e.g., Sánchez Almeida et al. 2016) and the super-solar.
pressed by power-law functions, i.e., we assume linearity for
dian values of individual galaxies within the bins. The data

\[ M \]

All these panels in Figures 1 and 3 clearly show that the

In this section, we attempt to disentangle the connection be-

1 summarize the derived physical properties (full version

shown that

\[ M \]

Notes.

(1) (2) (3) (4) (5) (6) (7) (8) (9) (10)

(1) (2) (3) (4) (5) (6) (7) (8) (9) (10)

(1) (2) (3) (4) (5) (6) (7) (8) (9) (10)

(1) (2) (3) (4) (5) (6) (7) (8) (9) (10)

(1) (2) (3) (4) (5) (6) (7) (8) (9) (10)

In Figure 3, the ionization parameter \( U \) and electron density \( n_e \) are presented in the same manner as metallicty. All these panels in Figures 1 and 3 clearly show that the nebular conditions depend strongly on \( M_* \) and sSFR. It is shown that \( Z \) and \( U \) are more sensitive to sSFR than to \( M_* \). Meanwhile, the electron density \( n_e \) increases both with \( M_* \) at fixed sSFR as well as with sSFR at fixed \( M_* \). Table 1 summarize the derived physical properties (full version including emission line flux measurements available online). The relations between the observed strong-line fluxes and derived physical quantities are presented in Appendix B.

4 CORRELATION ANALYSIS

In this section, we attempt to disentangle the connection be-

between the parameters of the ionized H II regions and the host
galaxies. Figure 4 shows correlations for all possible combina-

tions among the five parameters of interest, \( \log M_* / M_{\odot} \),

\[ \log sSFR/(\text{yr}^{-1}) \], \( 12 + \log O/H \), \( \log U \), and \( \log n_e/(\text{cm}^{-3}) \). The stellar mass and sSFR of each stack are defined as the median values of individual galaxies within the bins. The data points are color-coded by sSFR, and the symbol size denotes the number of galaxies in each stacking bin, as indicated outside the frames.

4.1 Ansätze

We here propose a set of ansätze to figure out the connec-
tions between the parameters. First we assume that the par-
tial correlations between any two of the parameters are ex-

pressed by power-law functions, i.e., we assume linearity for

their logarithms. Although this is undoubtedly an oversim-

plification, it is reasonably supported by the observed re-

lations shown in Figure 4, and indeed a more elaborate parametrization is not justified by our limited data. Still, the power-law slopes that we obtain from the data will be helpful to qualitatively understand the physics controlling the conditions of the nebulae and galaxy evolution.

To shape a course to analyze the observed correlations, we employ further assumptions for the dependencies of the parameters. In Figure 5, we schematically show our assump-
tion about the dependencies of the parameters. The black (white) arrows indicate the direction of possible causation which is expected to lead to a positive (negative) correlation between the two parameters at their starting and ending points. We here consider for convenience an additional parameter, the gas mass ratio \( \mu_{\text{gas}} = M_{\text{gas}}/M_* \), which is as-

sumed to be uniquely determined once the key parameters of interest are given.

One may naturally argue that the SFR of a galaxy in-

crease with the amount of the fuel, i.e., the gas content in the

system. Star formation consumes the gas reservoir while accumulating the stellar mass of the galaxy. Thus, the gas

mass ratio is naturally expected to be anticorrelated with \( M_* \). These causations are illustrated, with labels “Fueling” and “Consumption” respectively, in Figure 5.

Mathematically, the SFR integrated over a galaxy is re-

lated to its total molecular gas mass as \( \text{SFR} = M_{\text{gas}}/\tau_{\text{dep}} \)

where \( \tau_{\text{dep}} \) is the gas depletion timescale, and the star for-

mation efficiency is defined as its inverse \( \text{SFE} = 1/\tau_{\text{dep}} = \text{SFR}/M_{\text{gas}}. \) On the other hand, a simple hypothesis relates the local SFR volume density \( \dot{\rho}_s \) with the gas volume den-
Figure 4. Correlations between two parameters among the five of interest, $\log M_\ast/M_\odot$, $\log$ sSFR/(yr$^{-1}$), $12 + \log(O/H)$, $\log U$, and $\log n_e/(\text{cm}^{-3})$. Filled circles indicate stacked measurements. The color and size of the data points indicate the median sSFR and the number of galaxies in each stack, respectively, as shown outside the frame.

Figure 4. Correlations between two parameters among the five of interest, $\log M_\ast/M_\odot$, $\log$ sSFR/(yr$^{-1}$), $12 + \log(O/H)$, $\log U$, and $\log n_e/(\text{cm}^{-3})$. Filled circles indicate stacked measurements. The color and size of the data points indicate the median sSFR and the number of galaxies in each stack, respectively, as shown outside the frame.

D. Kashino & A. K. Inoue

Figure 4. Correlations between two parameters among the five of interest, $\log M_\ast/M_\odot$, $\log$ sSFR/(yr$^{-1}$), $12 + \log(O/H)$, $\log U$, and $\log n_e/(\text{cm}^{-3})$. Filled circles indicate stacked measurements. The color and size of the data points indicate the median sSFR and the number of galaxies in each stack, respectively, as shown outside the frame.

D. Kashino & A. K. Inoue

Figure 4. Correlations between two parameters among the five of interest, $\log M_\ast/M_\odot$, $\log$ sSFR/(yr$^{-1}$), $12 + \log(O/H)$, $\log U$, and $\log n_e/(\text{cm}^{-3})$. Filled circles indicate stacked measurements. The color and size of the data points indicate the median sSFR and the number of galaxies in each stack, respectively, as shown outside the frame.

D. Kashino & A. K. Inoue

Figure 4. Correlations between two parameters among the five of interest, $\log M_\ast/M_\odot$, $\log$ sSFR/(yr$^{-1}$), $12 + \log(O/H)$, $\log U$, and $\log n_e/(\text{cm}^{-3})$. Filled circles indicate stacked measurements. The color and size of the data points indicate the median sSFR and the number of galaxies in each stack, respectively, as shown outside the frame.

D. Kashino & A. K. Inoue

Figure 4. Correlations between two parameters among the five of interest, $\log M_\ast/M_\odot$, $\log$ sSFR/(yr$^{-1}$), $12 + \log(O/H)$, $\log U$, and $\log n_e/(\text{cm}^{-3})$. Filled circles indicate stacked measurements. The color and size of the data points indicate the median sSFR and the number of galaxies in each stack, respectively, as shown outside the frame.

D. Kashino & A. K. Inoue

Figure 4. Correlations between two parameters among the five of interest, $\log M_\ast/M_\odot$, $\log$ sSFR/(yr$^{-1}$), $12 + \log(O/H)$, $\log U$, and $\log n_e/(\text{cm}^{-3})$. Filled circles indicate stacked measurements. The color and size of the data points indicate the median sSFR and the number of galaxies in each stack, respectively, as shown outside the frame.

D. Kashino & A. K. Inoue

Figure 4. Correlations between two parameters among the five of interest, $\log M_\ast/M_\odot$, $\log$ sSFR/(yr$^{-1}$), $12 + \log(O/H)$, $\log U$, and $\log n_e/(\text{cm}^{-3})$. Filled circles indicate stacked measurements. The color and size of the data points indicate the median sSFR and the number of galaxies in each stack, respectively, as shown outside the frame.

D. Kashino & A. K. Inoue

Figure 4. Correlations between two parameters among the five of interest, $\log M_\ast/M_\odot$, $\log$ sSFR/(yr$^{-1}$), $12 + \log(O/H)$, $\log U$, and $\log n_e/(\text{cm}^{-3})$. Filled circles indicate stacked measurements. The color and size of the data points indicate the median sSFR and the number of galaxies in each stack, respectively, as shown outside the frame.

D. Kashino & A. K. Inoue

Figure 4. Correlations between two parameters among the five of interest, $\log M_\ast/M_\odot$, $\log$ sSFR/(yr$^{-1}$), $12 + \log(O/H)$, $\log U$, and $\log n_e/(\text{cm}^{-3})$. Filled circles indicate stacked measurements. The color and size of the data points indicate the median sSFR and the number of galaxies in each stack, respectively, as shown outside the frame.
could be naturally expected as the consequence of the history of mass assembly and accompanying chemical enrichment. The current metallicity, however, also encloses the history of gas inflows and outflows. In particular, the inverse correlation seen between $Z$ and SFR at fixed $M_\ast$ has been commonly interpreted as the result of the so-called “dilution” effect: enhanced inflow of metal-poor gas in the intergalactic medium would dilute the metallicity in the ISM, while boosting star formation. This naturally predicts the anticorrelation between $Z$ and the gas content $\mu_{\text{gas}}$ through the gas consuming mass assembly, $Z$ is naturally expected to be anticorrelated with $\mu_{\text{gas}}$ even if we do not consider the effects of dilution due to an enhanced inflow rate. The dilution scenario implicitly assumes that galaxies of the same $M_\ast$ experienced the equivalent history of star formation and created the same amount of metals, and that the dispersions in the current gas content and thus metallicity are caused by recent fluctuations in the inflow rate in a short timescale. Keeping this in mind, we consider the effect of $\mu_{\text{gas}}$ on $Z$, especially at fixed $M_\ast$, and thus named the causation just “Dilution” for simplicity.

Given $n_e$, $Z$, and sSFR, we can then approach the ionization parameter $U$. Returning to the simple definition of $U = n_e/n_H$, we can separate the ionization parameter into two controlling factors, the ionizing photon flux illuminating the gas and the local density of the ionized gas. It would thus be natural to suppose that $U$ is anticorrelated with $n_e$ when fixing the ionizing radiation (labeled as “By definition” in Figure 5). Concerning the numerator, it may be natural to assume that the ionizing photon flux is primarily controlled by the sSFR (labeled as “Radiation” in Figure 5).

However, the increase in the galaxy-wide sSFR would not necessarily lead an increase in the ionizing photon flux in a single H II region, while certainly resulting in an increase in the total amount of the ionizing photons in a whole galaxy. Lastly, we expect that $U$ is also anticorrelated with $Z$. This is because an increase in metals contained in the stellar atmosphere of the central massive stars would result in a higher optical depth in the stellar winds, more efficiently absorbing ionizing photons (e.g., Dopita et al. 2006b). In addition, at higher metallicity, ionizing photons are more efficiently scattered in the photospheres and then their radiation energy are converted to kinetic energy of the stellar winds. We may also expect that an increase in the dust amount in the H II regions, which is expected to correlate with metallicity, would reduce the ionizing photon flux contributing to ionization (Inoue et al. 2001; Inoue 2001). The possible causation between $Z$ and $U$ is labeled as “Absorption/Conversion to kinetic energy” in Figure 5.

With these simple ansätze, we attempt to derive the dependencies of sSFR, $Z$, and $U$ on the other parameters. The ansätze denote $\text{sSFR} = \text{sSFR}(\mu_{\text{gas}}, n_e)$ and $Z = Z(M_\ast, \mu_{\text{gas}})$. However, we need to eliminate $\mu_{\text{gas}}$ because we have no direct information about the gas content ($\mu_{\text{gas}}$) of the galaxies. We have assumed that $n_e$ is uniquely determined by $M_\ast$ and $\mu_{\text{gas}}$ as shown in Figure 5, and thus $\mu_{\text{gas}}$ is uniquely determined once $M_\ast$ and $n_e$ are given. We thus parametrize sSFR and $Z$ as a function of $M_\ast$ and $n_e$ eliminating $\mu_{\text{gas}}$, as well as $U$ as a function of sSFR $Z$, and $n_e$. In this section, we explicitly argued which positive or negative correlation is expected between two parameters. In our analysis, however, we did not impose any limits on the power-law indices and their signs in the following analysis.

### 4.2 Multi regression analysis

Following the ansätze we made, we derived the relations between the quantities ($M_\ast$, sSFR, O/H, $U$, and $n_e$). The physically-motivated relations to be constrained are expressed as follows:

\[
\log \text{sSFR} + 9 = p_0 + p_1(\log M_\ast - 10) + p_2(\log n_e),
\]

\[
4 + \log O/H = q_0 + q_1(\log M_\ast - 10) + q_2(\log n_e),
\]

\[
\log U = r_0 + r_1(\log n_e) + r_2(4 + \log O/H) + r_3(\log \text{sSFR} + 9).
\]

We note that $M_\ast$, sSFR, and $n_e$ are given in units of $M_\odot$, yr$^{-1}$, and cm$^{-3}$, respectively. In addition to these, we also derived the empirical functional forms for the three nebular parameters, Z, $U$, and $n_e$, as a linear function of $\log M_\ast$ and log sSFR. Our fitting procedure fits a linear surface expressed by these equations to the stacked measurements, accounting for all the measurement errors on the relevant variables.

Let us consider a general form of these models as $y = p_0 + \sum_{k=1}^m p_k x_k$. We refer to $y$ and $x_k$ ($k = 1, 2, ..., m$) as objective and explanatory variables, respectively. We estimate the model parameters (i.e., the normalization and coefficients of Equations (3–5)) based on the maximum likelihood estimation. The likelihood $L$ is computed for a given model as follows:

\[
L = \prod_{i=1}^N \frac{1}{\sqrt{2\pi \sigma_{\text{err},i}^2}} \exp \left[ -\frac{1}{2} \frac{(y_{\text{obs,mod,}} - y_{\text{mod,}})^2}{2\sigma_{\text{err},i}^2} \right].
\]
where $N = 138$ is the number of stacked measurements. The error $\sigma_{\text{err},i}$ is obtained for each stack by summing in quadrature the uncertainties expected for the objective ($\delta y_i$) and explanatory ($\delta x_{k,i}$) variables as follows:

$$\sigma_{\text{err},i}^2 = \delta y_i^2 + \sum_{k=1}^{m} (p_k \delta x_{k,i})^2.$$  

(7)

Note that the value $\sigma_{\text{err},i}$ depends on the coefficients of the model.

We sampled the posterior probability distribution of the model parameters by employing a Markov-Chain Monte-Carlo (MCMC) technique, and using the emcee package for Python (Foreman-Mackey et al. 2013). In our fitting procedure, we adopted a uniform prior probability function for the model parameters by employing a Markov-Chain Monte-Carlo technique, and using the emcee package for Python (Foreman-Mackey et al. 2013). In our fitting procedure, we adopted a uniform prior probability function for each parameter.

5 RESULTS

In Table 2, we summarize the coefficients constrained by multi regression analysis. The upper three rows are the results for the physically-motivated models as defined in Section 4.1 (Equations 3–5). The lower three rows show the results obtained from regression on $M_*$ and sSFR. The reduced $\chi^2$ statistics are computed with these best-fit parameters as follows:

$$\chi^2/\text{dof} = \sum_{i=1}^{N} \frac{(y_{i,\text{obs}} - y_{i,\text{model}})^2}{\sigma_{\text{err},i}^2} / (N - m - 1).$$  

(8)

Note that the $(m + 1)$ is the number of parameters to be constrained. We also computed the residual intrinsic scatter assuming that the individual stacked measurements are intrinsically scattered with respect to the best fit following a Gaussian distribution of $\sigma_{\text{int}}$. In practice, we estimated $\sigma_{\text{int}}$ that satisfies the following relation:

$$\sum_{i=1}^{N} \frac{(y_{i,\text{obs}} - y_{i,\text{model}})^2}{\sigma_{\text{err},i}^2 + \sigma_{\text{int}}^2} / (N - m - 1) = 1.$$  

(9)

The reduced $\chi^2$ statistics for the fits with the physically-motivated ansätze are reasonably close to unity, with the residual scatter below 0.1 dex. This indicates that these models well fit the data with little residual information. For $Z$ and $U$, the $\chi^2/\text{dof}$ values are larger for the results from the regression on only $M_*$ and sSFR. Generally, the $\chi^2/\text{dof}$ statistics can be used to evaluate which model is better fitting the same data. However, in the case here, the $\chi^2/\text{dof}$ values do not immediately indicate that, for example for $Z$, the fit with $M_*$ and $n_e$ is better than the fit with $M_*$ and sSFR because the different models use the different variables. The intrinsic scatters are anyway smaller than 0.1 dex for the fits with $M_*$ and sSFR, similarly to the fits with the ansätze. Hereafter we focus our attention to the results for the physically-motivated models.

We found the best-fit relations for the sSFR, metallicity $Z$, and the electron density $n_e$, as follows (the errors are given in Table 2).

$$\log \text{sSFR} = -12.661 - 0.627 (\log M_* - 10) + 1.753 \log n_e.$$  

(10)

$$12 + \log O/H = 9.238 + 0.228 (\log M_* - 10) - 0.401 \log n_e.$$  

(11)

$$\log U = -2.316 - 0.360 (\log O/H + 4)$$

$$- 0.292 \log n_e + 0.428 (\log \text{sSFR} + 9).$$  

(12)

All these coefficients are well constrained, fully rejecting zero coefficients, i.e., the null hypothesis of no dependence on the parameter. In particular, we highlight that the scaling relation of $U(Z, n_e, \text{sSFR})$ is fairly constrained even with three variables. The signs of these coefficients are all consistent with our initial assumption described in Section 4.1 (see Figure 5). Note that the magnitude of the coefficients does not immediately indicate the relative strength of the dependence on the corresponding parameter since the entire range covered by the parameters are different from each other.

In Figure 6, we show the sSFR as a function of $n_e$, color-coded by $M_*$, in comparison with the derived linear relation of Equation (10) shown in the background. Similarly, Figure 7 shows $Z$ as a function of $M_*$, color-coded by $n_e$ (upper panel), and as a function of $n_e$, color-coded by $M_*$ (lower panel). For $U$, we computed the residuals after subtracting the sSFR term (0.428 + log sSFR + 9)). Figure 8 shows these residuals as a function of $Z$, color-coded by $n_e$ (upper panel), and as a function of $n_e$, color-coded by $Z$ (lower panel). In all these panels of Figures 6–8, the overall trends both of the correlation between the $x$- and $y$-axis variables and the color trend of the third axis variables are reasonably reproduced by the best-fit scaling relations. Meanwhile, there are several outliers possibly due to their large errors on the estimates of $n_e$ and/or $Z$. In the next section, we attempt to interpret the global trends and discuss the physical mechanisms behind the scaling relations, not caring about minor deviations from these power-law relations.
Table 2. Coefficients of the linear surface fitting between the parameters

|                       | Intercept | log $M_\star$($M_\odot$) - 10 | log $n_e$(cm$^{-3}$) | 4 + log(O/H) | log sSFR( yr$^{-1}$) + 9 | $\chi^2$/dof | $\sigma_{\text{int}}$ |
|-----------------------|-----------|-------------------------------|---------------------|--------------|--------------------------|-------------|-----------------|
| log sSFR + 9          | $-3.661^{+0.124}_{-0.135}$ | $-0.627^{+0.028}_{-0.029}$ | 1.753$^{+0.063}_{-0.066}$ | - | - | 0.99 | - |
| 4 + log(O/H)          | 1.238$^{+0.036}_{-0.022}$ | 0.228$^{+0.011}_{-0.011}$ | $-0.401^{+0.024}_{-0.024}$ | - | - | 1.61 | 0.0633 |
| log $U$               | $-2.316^{+0.004}_{-0.009}$ | - | $-0.292^{+0.002}_{-0.002}$ | $-3.600^{+0.007}_{-0.007}$ | $0.428^{+0.014}_{-0.014}$ | 1.89 | 0.0585 |
| log $n_e$             | 2.066$^{+0.011}_{-0.012}$ | 0.310$^{+0.021}_{-0.019}$ | - | - | - | 0.402$^{+0.020}_{-0.019}$ | 0.85 |
| 4 + log(O/H)          | 0.427$^{+0.002}_{-0.002}$ | 0.136$^{+0.006}_{-0.006}$ | - | - | - | $-0.208^{+0.008}_{-0.008}$ | 5.71 |
| log $U$               | $-3.073^{+0.002}_{-0.002}$ | $-0.137^{+0.002}_{-0.002}$ | - | - | - | $0.372^{+0.002}_{-0.002}$ | 26.07 |
|                       |           |                               |                     |              |                          |             | 0.0874 |

We note that $M_\star$, sSFR, and $n_e$ are given in units of $M_\odot$, yr$^{-1}$, and cm$^{-3}$, respectively.

b Reduced chi-square statistic.

c Intrinsic scatter with respect to the relevant objective variable.

Figure 7. Upper panel: Metallicity $Z$ as a function of $M_\star$, color-coded by $n_e$. Lower panel: $Z$ as a function of $n_e$, color-coded by $M_\star$. The best-fit relation expressed by Equation 11 is shown in the background.

Figure 8. Upper panel: Residual of $U$ after subtracting the sSFR dependence ($0.428(\log\text{sSFR} + 9)$) as a function of $Z$, color-coded by $n_e$. Lower panel: $U - 0.428\log\text{sSFR}$ as a function of $n_e$, color-coded by $Z$. The best-fit relation expressed by Equation 12 is shown in the background.

6 DISCUSSIONS

We successfully constrained the dependencies of the physical quantities, sSFR($M_\star, n_e$), Z($M_\star, n_e$), and $U$(sSFR, Z, $n_e$), based on the simplified ansätz. In this section, we attempt to interpret the results with our physically-motivated assumptions, and compare our findings to other theoretical predictions and observational results in the literature.
6.1 Specific SFR and gas content

Our result from regressing sSFR on $M_\ast$ and $n_e$ indicates that sSFR is a strong function of $n_e$ with a power-law slope of $\approx 1.75$ at fixed $M_\ast$. As noted above, this relatively large dependence includes the underlying correlations between $n_e$ and $\mu_{\text{gas}}$. The sSFR of galaxies can be expressed in terms of SFE ($=\text{SFE}/M_{\text{gas}}$) and the gas content as follows:

$$\text{sSFR} = \text{SFE} \cdot \frac{M_{\text{gas}}}{M_\ast} = \text{SFE} \cdot \mu_{\text{gas}}. \quad (13)$$

As we mentioned in Section 4.1, it would be natural to assume that the SFE is proportional to the inverse of the free-fall timescale $\tau_f$ of gas clouds, thus we assume $\text{SFE} \propto n_e^{1/2}$. This simple assumption and the observed scaling of sSFR (Equation 10) yield

$$\text{SFE} \propto \text{sSFR}^{0.29} M_\ast^{0.14}. \quad (14)$$

Observationally, Sargent et al. (2014) found that SFE increases as a weak function of sSFR: $\text{SFE} \propto \text{sSFR}^{0.19}$, which may be in broad agreement with our result. We can also obtain scaling relations involving the gas content as follows.

$$\mu_{\text{gas}} \propto M_{\text{gas}}^{-0.63} n_e^{1.3} \quad \text{or} \quad M_{\text{gas}} \propto M_\ast^{0.17} n_e^{1.3} \quad (15)$$

implies that the electron density $n_e$ in H\textsc{ii} regions sensitively increases with the gas content at fixed $M_\ast$. We also obtain a relation between SFR and $M_{\text{gas}}$ as

$$\text{SFR} \propto M_{\text{gas}}^{1.40} M_\ast^{-0.15}. \quad (17)$$

The scaling relation $\text{SFR} \propto M_{\text{gas}}^{1.40}$ (with minor dependence on $M_\ast$) agrees very well with observed integrated SK relation based on direct measurements of the gas content of galaxies (e.g., Daddi et al. 2010; Genzel et al. 2010; Tacconi et al. 2013; Santini et al. 2014), as well as the conventional SK relation that correlates the surface densities of SFR and gas mass, for which observations have measured indices $N \approx 1.4$ (e.g., Kennicutt 1998). It is worth emphasizing that the scaling relation between sSFR and $\mu_{\text{gas}}$ with a slope slightly smaller than 1.5, as supported by other observations, is derived by assuming $\text{SFE} \propto \tau_f^{-1} \propto n_e^{1/2}$. In doing so, $n_e$ is regarded as being proportional to the density of the molecular clouds where stars form, though this is not trivial because $n_e$ is the measure of the electron density of H\textsc{ii} regions. On the other hand, the agreement between our result ($\text{SFR} \propto M_{\text{gas}}^{1.40}$) and other observations of the SK relations indicate that this assumption would be reasonable.

6.2 Metallicity

We found positive and negative dependences of metallicity $Z$, respectively, on $M_\ast$ and $n_e$. In the previous subsection, we obtained the scaling relation between $n_e$ and gas content (Equation 16) based on a simple assumption of $\text{SFE} \propto \tau_f^{-1} \propto n_e^{1/2}$ and the observed scaling of sSFR (Equation 10). Using Equation 16, we can eliminate $n_e$ from the $Z(M_\ast, n_e)$ relation (Equation 11), and rewrite $Z$ in terms of $M_\ast$ and gas content as follows:

$$Z \propto M_{\text{gas}}^{0.23} n_e^{-0.40} \propto \mu_{\text{gas}}^{-0.32} M_\ast^{0.027}. \quad (18)$$

or, more simply,

$$Z \propto \mu_{\text{gas}}^{-0.32} M_\ast^{0.35}. \quad (19)$$

The former equation indicates that the correlation between $Z$ and gas-to-stellar mass ratio $\mu_{\text{gas}}$ would be almost independent of $M_\ast$ (and other parameters). This is indeed what we expect from simple, classical models of chemical evolution of galaxies (e.g., Schmidt 1963). The latter form, however, indicates that the metallicity increases with $M_\ast$, as naturally expected as the result of the chemical evolution along the galaxy mass assembly history. Meanwhile, the anticorrelation with gas content is as expected from the “dilution” scenario as described in Section 4.1.

Anticorrelations between $Z$ and the gas content of galaxies have been directly measured by observations. Bothwell et al. (2013) analyzed a galaxy sample with the measurement of the H\textsc{i} gas content, and found that the metallicity (via the strong-line method) decreases with the H\textsc{i} gas mass $M_{\text{HI}}$ at fixed $M_\ast$, approximately as $Z \propto M_{\text{HI}}^{-0.17}$ (see Figure 2 of the reference). Furthermore, Davé et al. (2013) attempted to model the global properties of the H\textsc{i} gas content in galaxies using cosmological simulations. Their preferable outflow model, which is able to reasonably reproduce the observed H\textsc{i} mass function, predicts that the metallicity is inversely correlated with the H\textsc{i} gas mass with a slope of $-0.26$ at a given $M_\ast$. Our result is in reasonable agreement with these dependencies found by both observations and numerical simulations.

The inverse correlation between $Z$ and sSFR (at fixed $M_\ast$) has been also characterized by regressing $Z$ on the $M_\ast$–sSFR plane (Table 2 lower part) as follows:

$$12 + \log(O/H) = 8.427 + 0.136 (\log M_\ast - 10) - 0.208 (\log \text{sSFR} + 9). \quad (20)$$

We note that this relation is nearly equivalent to what is obtained by eliminating $n_e$ from $Z(M_\ast, n_e)$ and $n_e(M_\ast, \text{sSFR})$ (see Table 2), with a slight difference in the coefficient of $\log M_\ast$. The relation between these three parameters ($M_\ast$, $Z$, sSFR), i.e., the FMR, has been commonly characterized by introducing a variable $\mu_\alpha$ defined as follows:

$$\mu_\alpha = \log M_\ast - \alpha \log \text{sSFR}. \quad (21)$$

where the parameter $\alpha$ is determined to minimize the scatter in terms of metallicity around the average relation. Our result from a simple plane fitting provides $\alpha = 0.208(0.136+0.208) = 0.60 \pm 0.01$. This is in good agreement with the value ($\alpha \approx 0.66$) found by Andrews & Martini (2013). Meanwhile, such values of $\alpha$ are larger than those ($\alpha \approx 0.3$) as reported by studies relying on the strong-line metallicity indicators (e.g., Mannucci et al. 2010). Note that Equation (20) can be rewritten using $\mu_{0.61}$ as follows:

$$12 + \log(O/H) = (5.19 \pm 0.06) + (0.344 \pm 0.006) \mu_{0.60}. \quad (22)$$

The slope of 0.34 is lower than the slope of 0.43 found by Andrews & Martini (2013), and the difference is not distinguished if we adopt $\alpha = 0.66$.

6.3 Ionization parameter

Our results indicate that the ionization parameter $U$ varies as a strong function of the sSFR of galaxies, but also depends significantly on both $Z$ and $n_e$. The former dependence implies that the ionization photon flux in a single H\textsc{ii} region increases on average with increasing total (i.e., galaxy-wide)
Physical conditions of ionized gas

sSFR of the host galaxies. To directly examine this hypothesis, we need to resolve the individual H II regions in galaxies across wide ranges of \( M_* \) and sSFR.

There are other ideas to account for the correlations between sSFR and \( U \). The observed \([\text{O} \text{III}]/[\text{O} \text{II}]\) ratio, a proxy of \( U \), could effectively increase for a “density-bounded” H II region, which is fully ionized and whose size is determined by the cloud size, instead of the Strömgren radius. Since \( O^{++} \) dominates at the inner region of the ionized sphere, close to the radiation source, whereas \( O^+ \) becomes more dominant at the outer region, density-bounded H II regions are expected to have a significantly smaller [O II] zone, resulting in higher \([\text{O} \text{III}]/[\text{O} \text{II}]\). Specially, Binette et al. (1996) considered the effects of density-bounded (“matter-bounded” in the reference) ionized gas clouds to the changes of the emission-line ratios, and showed that the line ratios (especially, \([\text{O} \text{III}]/[\text{O} \text{II}]\)) vary along the increasing fraction of the density-bounded clouds in a very similar way to the variation with increasing \( U \) (see Figure 6 in the reference). A larger contribution of density-bounded H II regions has proposed to account for observed higher \([\text{O} \text{III}]/[\text{O} \text{II}]\) ratios (Brinchmann et al. 2008; Nakajima & Ouchi 2014). In addition, it is recently thought that extremely high \([\text{O} \text{III}]/[\text{O} \text{II}]\) ratios are often incident to non zero escape fraction, i.e., the fraction of ionizing photons that escape into intergalactic space (e.g., Nakajima & Ouchi 2014; Nakajima et al. 2016; Izo et al. 2016, 2018). Similarly, if an H II region overlaps an adjacent H II region, the effective ionization parameter would increase locally because the overlap region is illuminated by both ionizing sources. It could be naturally expected that such situations tend to happen more frequently in systems with higher sSFRs because the number density of H II regions may increase with increasing sSFR. We note that our estimate of the ionization parameter assumes a simple ionization-bounded plane-parallel geometry in order to limit the number of free parameters, and that our estimation could thus be inaccurate for some cases. We defer further investigations of the relations between \( U \) and the emission-line indicators, such as \([\text{O} \text{III}]/[\text{O} \text{II}]\), including more realistic geometric conditions to future work.

We now turn to the dependence on metallicity. There have been observational studies that indicate the presence of an anticorrelation between \( U \) and Z based on a strong-line method (Nakajima & Ouchi 2014; Onodera et al. 2016). Shapley et al. (2015) and Kojima et al. (2017) mentioned the anticorrelation by using the stacked measurements based on the direct method from Andrews & Martini (2013). In Figure 4, our data show that \( U \) is apparently strongly anticorrelated with \( Z \): a simple power-law fit yields \( U \propto Z^{-1.52\pm0.04} \). Compared with this apparent relation, however, our result from the four-dimensional fitting indicates a shallower slope for the partial correlation \( U \propto Z^{-0.38} \) (Equation 12) at fixed sSFR and \( n_e \). The apparent stronger \( Z \)-dependence can be attributed to the anticorrelation between sSFR and \( Z \) (Section 6.2). Our result thus indicates necessity of physical mechanisms to generate the intrinsic moderate inverse correlation between \( U \) and \( Z \). Some explanations have been suggested to account for the \( U-Z \) anticorrelation as mentioned in Section 4.1 (e.g., Dopita & Evans 1986; Dopita et al. 2006a; Kewley et al. 2013).

Dopita et al. (2006b) have theoretically found that \( U \) correlates with \( Z \) as \( U \propto Z^{-0.8} \) including the following two effects. First, at higher metallicity, ionizing photons emitted by a star are more absorbed by the stellar winds due to an increased optical depth. Second, at higher metallicity, ionizing photons are more efficiently scattered in the photospheres and their radiation energy is more converted to kinetic energy in the stellar wind base region (Leitherer et al. 1992). It may be worth noting that the line-blanketing effects could also be considered as an origin of the \( U-Z \) anticorrelation: metal elements stored in the stellar photosphere block the radiated ionizing photons, reducing the emergent flux at wavelengths of the metallic lines. However, the back-warming effect occurs immediately following the blocking, where metals re-emit photons at lower frequencies, raising the adjacent continuum. Kudritzki (2002) demonstrated that these effects have little influence on the number of hydrogen ionizing photons (\( \geq 13.6\ eV \)) because the two effects almost entirely cancel each other.

Another explanation is that the \( U-Z \) correlation is in part caused by dust absorption in the H II region. Inoue et al. (2001) established a theoretical framework for the dust absorption of ionizing photon in an H II region, and found observationally an anticorrelation between the fraction \( f_d \) of the ionizing photons contributing to ionization and metallicity (or dust-to-gas mass ratio) for individual H II regions in the Milky Way. Similarly, Inoue (2001) found signs of such anticorrelation also for nearby galaxies, with \( f_d \) ranging from \( \sim 40-70\% \).

The \( U-Z \) anticorrelation could also be related to the possible correlation between the IMF and metallicity (Evans & Dopita 1985): a top-heavy IMF would yield more ionizing photons per SFR. Recently, Martín-Navarro et al. (2015) found that the local IMF is tightly related to the local metallicity, becoming top-heavier at lower metallicity. Such correlation may be explained if the fragmentation scales of the molecular clouds are determined by the Jeans mass, which increases with increasing temperature (i.e., towards lower metallicity). A change in the IMF, however, may also change the estimations of \( M_* \) and sSFR, as well as the ionization parameter because the shape of the ionizing spectrum changes. We defer further investigation including the possible dependence of the IMF to future work.

Turning to the \( n_e \)-dependence, our four-dimensional fitting found a negative partial correlation between \( U \) and \( n_e \) as \( U \propto n_e^{-0.29} \), at fixed sSFR and \( Z \). As opposed to their apparent positive correlation (with the Spearman’s rank correlation coefficient \( \rho = 0.476 \) with \( p \)-value \( < 10^{-8} \)) as shown in the relevant panel in Figure 4. The apparent correlation is caused by the positive correlation between sSFR and \( n_e \) and the tight correlation between \( U \) and sSFR. Indeed, the bottom-leftmost panel in Figure 4 shows the presence of an inverse correlation between \( U \) and \( n_e \) at fixed sSFR (i.e., at fixed color). The observed dependence is weaker than what is expected from the definition of \( U \), i.e., \( U \propto n_e^{-0.1} \). This may indicates that the differential electron densities between relatively inner regions, where the ionization parameter is defined, and outer regions, which may dominate the \([\text{S} \text{II}]\) fluxes, in the H II regions.

2 This study specially focused on the conditions in the narrow-line region of active galactic nuclei, but can also applied to star-forming H II regions.
The weak negative dependence of $U$ on $n_e$ is in good agreement with what was derived by Dopita et al. (2006b), who found that the theoretical ionization parameter is weakly related to the density $n_0$ of the ambient ISM as $U \propto n_0^{-1/5}$ by modeling the evolution of an expanding H$\text{II}$ region, pressurized by the stellar winds from the central star cluster. In their model, the internal density $n_0$ of an isothermal spherical H$\text{II}$ region is related to the ambient density, as $n_0 \propto n^{-1/5}$, as well as the inner radius $R_0$, which is also linked to it as $R_0 \propto n^{-1/5}$ (Oey & Clarke 1997). For a fixed ionizing luminosity of the central source, the ionization parameter depends on these quantities as $U \propto 1/(n_0 R_0^2)$, yielding the relation between $U$ and the ambient density as above.

Lastly, we note that we recently obtained a conclusion that the ionization parameter $U$ is directly linked to sSFR, while little or probably not with metallicity (Kasen et al. 2018), from comparison between the high-redshift sample of the main-sequence star-forming galaxies at $z \sim 1.6$ (a subset from Kashino et al. 2017, 2018) and the “sSFR-matched” sample of the SDSS galaxies. These two samples that have the equivalent average sSFRs present similar levels of the ionization parameter, whereas the sSFR-matched low-redshift sample shows relatively lower metallicities. However, our results in this paper are indeed not against what is found by Kasen et al. (2018) within the uncertainties. The difference in the metallicity of these two samples is found to be $\sim 0.1$ dex. From our result, this metallicity difference induces only a very small difference $(\lesssim 0.05$ dex) in $U$ at fixed sSFR, which is equivalent to the level of statistical uncertainties presented in Kasen et al. (2018). We also note that our analysis has been applied only for the low redshift galaxies, whereas the attention in Kasen et al. (2018) was more focused on the evolution over a wide range of cosmic time. The investigation of redshift evolution of the relationships that we found between the nebular parameters is beyond the scope of this paper.

6.4 Application of the scaling relations

Lastly, we mention possible application of the scaling relations we derived in this paper. Our results can be implemented into a subgrid recipe of galaxy evolution simulations to compute the conditions of ionized nebula from the properties of host galaxies, and predict the fluxes of emission lines of interest. Predicting strong-line fluxes is useful to prepare mock catalogs for future spectroscopic surveys as well as to test the models by comparing observations. Our results would be also implemented into SED fitting codes to include the contribution from emission lines for given galaxy properties.

Here we present an application of our scaling relations for a semi-analytic model galaxy catalog. In semi-analytic models, the evolution of galaxies embedded in collapsed dark halos, is traced following the halo merger trees while solving a simplified model that controls the star formation involving relevant physical phenomena such as supernova and/or AGN feedback. As a result, some fundamental quantities of the galaxies, such as $M_\ast$, $M_\text{gas}$, SFR, metal abundance and the stellar population, are obtained as functions of cosmic time. Once $M_\ast$ and $M_\text{gas}$ are given, our scaling relations can be utilized to compute the conditions of the H$\text{II}$ regions, $n_e$ and $U$, from Equations (16) and (12), respectively. With the gas-phase $Z$ traced in the simulation, it is then possible to compute the emission line fluxes by interpolating a prepared table of the line fluxes for grids of $(n_e, Z, U)$.

For a demonstration, we use the New Numerical Galaxy Catalog ($^2$GC; Ishiyama et al. 2015; Makita et al. 2016)$^3$. From the public light-cone catalog, we extracted a sample of 4855 galaxies limiting to those at $z < 0.3$ and having sSFR $\geq 10^{-11}$ yr$^{-1}$. We define the stellar mass as the sum of disk (M$\text{star}$) and bulge (M$\text{starb}$) masses, and take the gas mass ($M_\text{gas} = M_\text{cool}$) and metal mass ($M_Z = M_Z\odot$) in the cold phase. The metallicity is defined as $12 + \log(O/H) = 8.69 + \log(M_\text{Z}/M_\text{gas}/Z_\odot)$ where $Z_\odot = 0.014$. Figure 9, we show the correlations among $M_\ast$, sSFR, $Z_\ast$, given by the semi-analytic model, and $n_e$ and $U$ computed with our scaling relations. In the $^2$GC sample, galaxies with sSFR $\geq 10^{-8.5}$ yr$^{-1}$ are missing, while a space of low $M_\ast < 10^9 M_\odot$ and low sSFR ($< 10^{-9.5}$ yr$^{-1}$) is filled up, where we have no SDSS stacks (see Figure 4). Moreover, a population of very metal-rich galaxies ($12 + \log(O/H) > 9$) exists. Regardless of these discrepancies, the trends of the SDSS sample seen in Figure 4 are globally reproduced here, manifesting the effectiveness of the application to semi-analytic model galaxy samples. Detailed comparison of the $^2$GC catalog with observations and/or predictions of any concrete observables are beyond the scope of this paper.

7 CONCLUSIONS

We have investigated the connection between the physical parameters characterizing the conditions of star-forming ionized gas and the global properties of the host galaxies, especially focusing on five key parameters, i.e., stellar mass, sSFR, oxygen abundance, ionization parameter, and electron density of the ionized gas. In order to obtain accurate estimates of the oxygen abundance, we have utilized the direct method based on the detection of auroral lines ([O iii]λλ4363, and [O ii]λλ3720,7330) in the set of stacked spectra of the SDSS galaxies at $z \sim 0.1$, which are binned into the $M_\ast$-sSFR grids. Our key conclusions can be summarized as follows.

(i) The gas-phase oxygen abundance was successfully measured for a wide range of $M_\ast$ and sSFR by using stacked spectra. The resultant values range between $7.6 \lesssim 12 + \log(O/H) \lesssim 8.9$, tightly correlating with sSFR and $M_\ast$.

(ii) The electron density $n_e$ was measured from the [S ii] doublet ratio, accounting for the dependence on the electron temperature, and found to increase with increasing sSFR at fixed $M_\ast$ as well as with increasing $M_\ast$ at fixed sSFR.

(iii) The ionization parameter $U$ was measured from the [O iii] / [O ii] ratio, accounting for the dependencies on $Z$ and $n_e$. The resultant values of $U$ were found to be tightly correlated with the sSFR, and also to depend on $Z$ and $n_e$.

(iv) The partial anticorrelation between $U$ and $Z$ ($U \propto Z^{-0.36}$ at fixed sSFR and $n_e$) was found to be weaker than the apparent correlation ($U \propto Z^{-1.32}$). The observed moderate (i.e., not as strong as its apparent correlation) anti-

3 http://hpc.imit.chiba-u.jp/~ngen/index.html
correlation between $U$ and $Z$ is qualitatively consistent with theoretical predictions of the intrinsic $U-Z$ relation.

(v) The ionization parameter is also found to negatively correlate with $n_e$ at fixed sSFR and Z as $U \propto n_e^{-0.29}$. This is qualitatively consistent with the definition of the ionization parameter, though the dependence on $n_e$ is weaker than the inverse proportion. This may indicates that the electron density probed by the [S\text{ii}] doublet scales with the inner surface density more sensitively than proportionality.

Our results presented in this paper have firmly provided the global picture on the behavior of the nebulae at the galaxy-wide scales. The scaling relationships that we found between the fundamental properties of galaxies and the nebulae provide a set of boundary conditions which are to be reproduced by high-resolution simulations. On the other hand, our results provide a convenient tool to compute the nebular conditions and to predict fluxes of strong emission lines with little computational cost in semi-analytic models and/or SED fitting codes. Such application is thus useful to construct mock catalogs for forthcoming large spectroscopic surveys with the next-generation spectrographs, such as the the Subaru/Prime Focus Spectrograph (PFS), Dark Energy Spectroscopic Instrument (DESI), Multi-Object Optical Near-infrared Spectrograph (MOONS), and Euclid and WFIRST in space. These next-generation surveys will then enable us to extend our analysis beyond the low redshift Universe in the near future, and to understand the physical processes behind the connections of galaxy properties in more detail.

ACKNOWLEDGEMENTS

We are grateful Brett Andrews for kindly providing us with their catalog and data. We also thank K. Yabe and T. Kojima for useful discussions and comments. This work has been supported by KAKENHI (17H01114) and by NAOJ ALMA Scientific Research Grant number 2016-01 A.

Funding for the SDSS and SDSS-II has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, the U.S. Department of Energy, the National Aeronautics and Space Administration, the Japanese Monbukagakusho, the Max Planck Society, and the Higher Education Funding Council for England. The SDSS Web Site is http://www.sdss.org/.
O

of flux these measurements, we defined the fraction of the absorbed subtracted using pPXF (Section 2.3) with those based on the composite spectra in which the stellar continuum is 

for the reduction due to the stellar atmospheric absorption

It would be useful to see the relation between the strong-

DIAGRAMS

APPENDIX B: NEBULAR DIAGNOSTIC

Here we present empirically-calibrated formulae for correct-

line ratio indices are defined as follows:

\[
\frac{\lambda 3727}{\lambda 5007} \quad \frac{\lambda 4959}{\lambda 5007} \quad \frac{\lambda 4959}{\lambda 5007} / \frac{\lambda 5007}{\lambda 5007} / \frac{\lambda 3727}{\lambda 5007} / \frac{\lambda 3727}{\lambda 5007}
\]

\[
O_2 = \frac{\lambda 3727}{\lambda 5007} / \frac{\lambda 3727}{\lambda 5007} \quad N_2 = [N \text{II}] / [H \alpha] \quad S_2 = [S \text{II}] / [H \alpha]
\]

The contours show the distribution of the individual SDSS galaxies in the parent catalog. To draw the contours, we additionally impose on the sample the detection of [O III] at 3σ (89 percent of the sample), although the original selection does not care about the detection of [O III] . The line ratios of the individual sources are corrected for extinction using the Cardelli et al. (1989) reddening curve and assuming the intrinsic Hα/Hβ = 2.86. The SDSS stacks with the successful metallicity estimation are shown by circles, color-coded by M∗, SFR, sSFR, 12 + log(O/H), U, and n∞ in each panel.

We can see clear trends of the location in the diagrams against all the average physical properties considered here. The trends with M∗ are consistent with those of the individual galaxies (Masters et al. 2016, see Figure 2). At a fixed SFR, the stacks cover a wide range of the emission-line ratios, almost from the one to the other tail-end in all four diagrams (second row from top in Figure B1), forming a narrow sequence parallel to the ridge line of the contours. The nebular conditions are also tightly correlated with the location in the diagram. The metallicity varies along the sequence, from the lowest Z at the high end-tail of the y-axis values to the highest Z at the opposite side (forth row in Figure B1). The trend with U is very similar to sSFR (fifth row). This is corresponding to the tight correlation between U and sSFR which we have seen in Figure 4. The electron density n∞ is correlated with the location in a similar way to SFR: stacks of fixed n∞ covers a wide range of these strong-line ratios and form a narrow sequence parallel to the ridge line of the contours. The trend in the N2 vs. O3 diagram was reported by Brinchmann et al. (2008).

This paper has been typeset from a TeX/LATEX file prepared by the author.

APPENDIX A: BALMER ABSORPTION

Here we present empirically-calibrated formulae for correcting the observed fluxes of the Balmer lines (Hα and Hβ) for the reduction due to the stellar atmospherical absorption as a function of M∗ and sSFR. To evaluate the effects of the Balmer absorption, we compare the flux measurements on the composite spectra in which the stellar continuum is subtracted using pPXF (Section 2.3) with those based on the continuum subtraction using a simple linear fitting, as usually done for continuum-faint or low-S/N spectra. From these measurements, we defined the fraction of the absorbed flux fabs for each stack, 

\[
fabs = \frac{F_{\text{int}} - F_{\text{obs}}}{F_{\text{int}}}
\]

where Fobs and Fint are the flux measured, respectively, with the linear-fit based and pPXF-based continuum subtraction.

In Figure A1, we show the absorbed fraction of the Hβ line in the M∗-sSFR plane. Stacks are limited to those with S/N(Hβ) ≥ 40. Each box represents a single stack, color-coded by fabs(Hβ). It is clear that a strong trend exists between fobs and a quantity, which can be defined as a combination of M∗ and SFR. The absorption fraction of the Hα flux exhibits a similar trend to Hβ, but the absolute values are much smaller. Considering a linear combination as

\[
x_{\alpha} = \log \text{SFR} - \alpha \log M^*
\]

we found that, when α = 1.32, the correlation becomes the tightest for both Hα and Hβ. The relation xα=1.32 = 0 is shown by a red solid line. In Figure A2, we show the absorption fraction of the Hα and Hβ fluxes as a function of xα=1.32. To express these trends, we fit an empirical functional form to data, and obtained the forms as follows,

\[
fobs(H\alpha) = \frac{1}{2} \left[ \text{erf}(-0.373(x + 2.331)) + 1 \right]. \quad (A1)
\]

\[
fobs(2H\beta) = \frac{1}{2} \left[ \text{erf}(-0.720(x + 0.107)) + 1 \right]. \quad (A2)
\]

APPENDIX B: NEBULAR DIAGNOSTIC DIAGRAMS

It would be useful to see the relation between the strong-line flux ratios and the galactic and nebular properties. In Figure B1, we show four routinely-used diagrams – R23 vs. O32, O2 vs. O3, N2 vs. O3, and S2 vs. O1 diagrams (the so-called BPT diagrams; Veilleux & Osterbrock 1987). Here the line ratio indices are defined as follows:

\[
R_{23} = [\text{O }\text{II}] / [\text{H}\alpha] \quad R_{32} = [\text{O }\text{III}] / [\text{H}\beta]
\]

\[
O_2 = [\text{O }\text{II}] / [\text{H}\beta] \quad N_2 = [\text{N }\text{II}] / [\text{H}\alpha] \quad S_2 = [\text{S }\text{II}] / [\text{H}\alpha]
\]
Figure A1. The fraction of the absorbed H$\beta$ flux relative to the intrinsic value is shown by colors in the $M_*$ vs. sSFR plane. Each box represents a single stack. For reference, we overplot the main-sequence of local galaxies (thick dotted line; Elbaz et al. 2007) and a relation corresponding to SFR = 1 $M_\odot$ yr$^{-1}$ (thin dotted line). Red solid line indicates a relation expressed as log SFR/(M$_\odot$ yr$^{-1}$) = 1.32 × (log $M_*/M_\odot$ − 10).

Figure A2. The fraction of the absorbed flux of H$\alpha$ (red) and H$\beta$ (blue) each as a function of $x = \log$ SFR/(M$_\odot$ yr$^{-1}$) − 1.32(log $M_*/M_\odot$ − 10). Solid curves indicate the best-fit empirical relations for each line (Equations A1 and A2).
Figure B1. $R_2$ vs. $O_2$, $O_1$ vs. $O_3$, $N_2$ vs. $O_1$, and $S_2$ vs. $O_3$ diagrams from left to right. The SDSS stacks are shown by circles, color-coded by different properties of host galaxies ($M_*$, SFR, and sSFR; upper 12 panels) and nebular parameters ($Z$, $U$, $n_e$; lower 12 panels). Symbol size corresponds to the number of galaxies in the stack (see Figure 4). Contours show the distribution of the individual galaxies in the parent sample.