Draft version November 19, 2025
Typeset using LaTeX **modern** style in AASTeX631

The Eddington Ratio Distribution of Narrow Line Active Galactic Nuclei

Michael R. Blanton[1], Arjun Suresh[1], Kyle B. Westfall[2], Dou Liu[3], and John Moustakas[4]

[1] *Center for Cosmology and Particle Physics, Department of Physics, New York University, 726 Broadway Rm. 1005, New York, NY 10003, USA*
[2] *University of California Observatories, University of California, Santa Cruz, 1156 High St., Santa Cruz, CA 95064, USA*
[3] *Wells Fargo, 550 S Tryon St, Charlotte, NC 28202, USA*
[4] *Department of Physics and Astronomy, Siena College, 515 Loudon Road, Loudonville, NY 12211, USA*

ABSTRACT

We measure the Eddington ratio distribution of local optical narrow line active galactic nuclei (AGN) as a function of host galaxy properties. These measurements are a potential test of galaxy formation theories that predict supermassive black hole growth and feedback. We base our sample on integral field spectroscopy from the Mapping Nearby Galaxies at APO (MaNGA) sample from the Sloan Digital Sky Survey IV Data Release 17 (SDSS-IV DR17). Starting with MaNGA's calibrated row-stacked spectra, we produce new spectroscopic data cubes with minimal covariance between spaxels and higher resolution point spread functions (PSF), run MaNGA's Data Analysis Pipeline, and extract line fluxes for the central PSF. Using the line ratio diagnostic techniques of Ji & Yan (2020), we identify AGN galaxies and determine their Hβ and [O III] line luminosities. For all galaxies not identified as AGN, we determine the threshold line luminosity they would have needed to be identified as AGN. These luminosity thresholds are essential to determine, because many star forming galaxies likely host AGN of significant luminosity that are unidentified because they are outshone by star formation related emission. We show that ignoring these selection effects when measuring the Eddington ratio distribution of optical narrow line AGN would lead to biased results. From the Hβ luminosities and luminosity detection thresholds, accounting for selection effects, we measure the luminosity and Eddington ratio distributions of Seyferts as a function of specific star formation rate (sSFR) and stellar mass. Defining $F_{\rm AGN}$ as the occurrence rate of AGN above a fixed Eddington ratio of 10^{-3}, we find that $F_{\rm AGN}$ is constant or increasing with stellar mass for star forming galaxies and declines strongly with stellar mass for quiescent galaxies. At stellar masses $\log_{10} M_* > 10.25$, the occurrence rate increases monotonically with sSFR. At low statistical significance, in our lowest mass bins $9.25 < \log_{10} M_* < 10.25$, $F_{\rm AGN}$ peaks at intermediate sSFR. These patterns reveal a complicated dependence of AGN activity on galaxy properties for theoretical models to explain.

1. INTRODUCTION

At the centers of essentially all massive galaxies are supermassive black holes, with masses that can exceed 10^9 $M_\odot$. During the growth of these black holes, they can emit jets, winds, and radiation, sometimes outshining the galaxy itself, a condition known as an active galactic nucleus (AGN).

Modern models of galaxy formation rely on this emission to explain basic phenomena of galaxies, like their distribution of luminosities, colors, and, star formation rates (for reviews see Fabian 2012; Somerville & Davé 2015). Essentially, this AGN feedback is thought to modulate and in some cases shut down star formation. Therefore the relationship between the AGN activity and the other properties of the galaxies is a potential observable that could distinguish competing models of galaxy formation.

Virtually all galaxies have supermassive black holes, but only a small fraction have detectable AGN at any given time, even at higher redshifts. The general interpretation of this difference is that quasars have limited periods of activity, so that at any given time only a fraction of galaxies have luminous AGN (Yu & Tremaine 2002; Shankar et al. 2009; Hickox et al. 2014). The fraction of the time a galaxy spends as an AGN is referred to as the "duty cycle." The total amount of time a galaxy spends as an AGN is often called the "lifetime," though sometimes this term is reserved for the length of individual episodes (e.g. Khrykin et al. 2021). The numerical values of these quantities depend on the definition of what an AGN is—e.g. what the minimum luminosity, Eddington ratio, or other characteristic of the emission has to be for us to identify it as an AGN.

One quantification of AGN demographics is the distribution of their luminosities and/or Eddington ratios. In the ansatz we adopt in this paper, every galaxy is an AGN of some luminosity, with a broad distribution of luminosities, but for most galaxies most of the time the AGN luminosity is undetectably low. The occurrence rate of AGN above some threshold of activity is then interpretable as the mean duty cycle of AGN above that threshold. If we want to understand how AGN and galaxy relate to each other and assess theoretical models of galaxy formation, key measurements are how these occurrence rates vary with galaxy mass, star formation rate, and other properties.

This work is part of a larger program to reassess the demographics of AGN with respect to their host galaxies, which began with Suresh & Blanton (2024). There are many manifestations of AGN, as X-ray, UV, optical broad and narrow line, mid-infrared UV, and radio sources, which need to be considered. Furthermore, patterns of AGN activity are known to vary strongly with redshift, peaking at redshifts $z \sim 2$–3. An enormous amount of work has gone into selecting samples of AGN and observing them out to high redshift. However, because the selection process is complex, it is difficult to infer relationships between the AGN samples and their host galaxies.

Aird et al. (2012) argues convincingly that selecting samples by their AGN luminosity, as many X-ray, radio, and optical studies do, means we are catching only galaxies in the rarer moments when they are most strongly active, and makes it difficult to understand the relationship between AGN and the galaxy host population. Hickox et al. (2014) explicitly demonstrates the resulting importance of selecting AGN from statistically complete galaxy samples and measuring the distribution of AGN properties conditioned on galaxy properties, rather than the converse. Among other problems, it is difficult to constrain quantities like the duty cycle from samples selected by AGN luminosity: because only luminous AGN are in the sample, the luminous AGN fraction cannot be directly determined relative to the full distribution of galaxies. In such samples, it is also difficult to account for selection effects associated with finding lower luminosity AGN that might be outshone by their host galaxy. For all these reasons, our approach is to start with a galaxy sample and measure the Eddington ratio distribution of the AGN, accounting for the selection effects.

In this paper we concern ourselves with a single AGN manifestation, optical narrow line emission. We use a low redshift ($z \sim 0.03$) sample that will allow us to have as much understanding of selection effects as we can while still retaining a statistically useful sample. The primary point of this paper is to accurately measure the [O III] λ5007 and Hβ narrow-line luminosity distributions and the corresponding Eddington ratio distributions as a function of galaxy stellar mass and star formation rate. Our methodology follows closely that of Trump et al. (2015), who performed a similar analysis on the Sloan Digital Sky Survey Legacy Survey (York et al. 2000). Suresh & Blanton (2024) has conducted a similar study using the 1.4 GHz radio emission of AGN.

We make use of the Mapping Nearby Galaxies at APO (MaNGA; Bundy et al. 2015) program, part of the Sloan Digital Sky Survey IV (SDSS-IV; Blanton et al. 2017). The sample we use is much smaller than the SDSS Legacy Survey Main Sample (Strauss et al. 2002) or the many millions in the Dark Energy Spectroscopic Instrument (DESI; DESI Collaboration et al. 2022) Bright-time Galaxy Survey (BGS)—about 10^4 galaxies in MaNGA versus 10^6 in SDSS Legacy and 10^7 in DESI BGS. But MaNGA has far better measurements for each galaxy—so if there is any sample for which we can disentangle the AGN selection effects, it is this one.

In Section 2, we describe our MaNGA central spectrum line measurements. In Section 3, we discuss our narrow-line AGN selection, how we determine detection thresholds, and how we measure Eddington ratios. In Section 4, we present the luminosities, Eddington ratios, and AGN detection thresholds for both quantities for the MaNGA sample. In Section 5, we estimate the intrinsic narrow-line luminosity distributions inferred from the sample, and in Section 6, we do the same for the Eddington ratio distributions. We compare to previous results in Section 7 and present our conclusions in Section 8.

The cosmological parameters we use to calculate luminosities are those of Planck Collaboration et al. (2020) (far right column of their Table 2), as implemented by the software in `astropy.cosmology.Planck18`.

2. MANGA DATA

2.1. *MaNGA Overview*

MaNGA is an integral field survey that observed around 10^4 galaxies during dark time at the Sloan Foundation Telescope at Apache Point Observatory, New Mexico, as part of SDSS-IV (2014–2020). It used 17 optical fiber bundles of varying sizes (19, 37, 61, 91, and 127 fibers) plugged into spectroscopic plug-plates and feeding the Baryon Oscillation Spectroscopic Survey (BOSS) spectrographs (Drory et al. 2015; Smee et al. 2013).

The sample spans stellar masses of about 10^9 to 10^{11} $M_\odot$, and includes all galaxy types in that range in its Primary, Secondary, and Color-Enhanced samples. The target selection is not uniform for a volume-limited sample; Wake et al. (2017) provides a description of the selection and the catalogs.

Using a suite of dithered observations, the Data Reduction Pipeline extracted each fiber spectrum and built a data cube for each galaxy (Law et al. 2015, 2016). A Data Analysis Pipeline (DAP) measured stellar and gas velocity and velocity dispersion maps, as well as spectral index and emission line flux maps (Belfiore et al. 2019; Westfall et al. 2019).

Sánchez et al. (2022) ran their Pipe3D software on the data cubes to create a value-added product containing alternative maps and high level derived parameters. These parameters include stellar masses, star formation rates based on Hα emission, and star formation rates based on stellar population synthesis models fit to the stellar continuum. In this paper we use the Pipe3D star formation rate, specifically its Hα-based estimate, excluding spaxels with strong emission line ratios consistent with Seyfert ionization (i.e. their `log_SFR_SF`).

For this work we applied the data cube reconstruction methods of Liu et al. (2020), which achieve slightly higher resolution than those found in DR17 and yield zero off-diagonal covariances between flux values at a given wavelength. We ran the DAP on these cubes, using the same version as used for DR17 with some minor changes (`3.1.0`). The most significant change to the DAP is that we set the off-diagonal covariances to zero in the calculation, which is accurate for our new cubes. We also added a check to mask a very small number of very high inverse variance pixels that appeared due to an error in our cube reconstruction. We have evaluated the results of this paper using both our measurements and the DR17 measurements, and do not find any significant differences.

2.2. *Central Emission Line Fluxes*

To find AGN we search for central ionized gas emission with line ratios consistent with AGN ionization spectra. The lines we consider here are Hα, Hβ, [O III] λ5007, [N II] λ6584, and [S II] $\lambda\lambda$6716, 6731.

We measure central line fluxes from the DAP maps, using a PSF-weighted aperture:

$$f_c = \frac{\sum_i f_i \mathrm{PSF}_i}{\sum_i \mathrm{PSF}_i^2} \tag{1}$$

where i indexes the pixels. We limit the PSF contributions to within $3.5''$ radius of the center, at which point the weighting is small (of order 0.05).

We calculate the uncertainties from the variances reported by the DAP for the individual map pixels (Belfiore et al. 2019, Section 3) added in quadrature using the weights. In calculating uncertainties in the line ratios, we use simple error propagation based on small errors.

We consider a measurement "good" if three conditions are met: if the redshift $z > 0$ for the galaxy; if, for each of the six lines mentioned above, there are three or fewer bad pixels contributing to the central line flux; and if the flux has a positive inverse variance (even if the flux is consistent with zero). Otherwise we do not analyze the galaxy. There are 9,971 MaNGA observations that pass this criterion. For the purposes of the AGN identification below, we consider a line "detected" if its flux is measured at 2σ or higher significance.

2.3. *MaNGA Sample*

For our determination of the distribution of luminosities and Eddington ratios, we use only MaNGA galaxies in its main statistical samples. There are other targets observed by MaNGA, but using these could lead to a statistical bias, since they are selected on known galaxy properties, including in some cases that they are known to be AGN.

Specifically, we use galaxies in the Primary, Secondary, or Color-Enhanced samples as described by Wake et al. (2017). We use the subsampled Secondary sample, which we identify as galaxies with a selection weight `ESRWEIGHT` > 0; this cut eliminates 515 galaxies. The final sample contains 9,029 galaxies. When we need black hole masses we further cut the sample to stellar velocity dispersions $\sigma_e > 60$ km s^{-1}, as measured by DAP within 1 R_e, which leaves 5,678 galaxies.

Figure 1 shows the Pipe3D stellar masses and sSFRs for the sample of 9,029 galaxies. The red symbols show the galaxies with AGN detections (see details on selection below), and the black symbols show the galaxies with AGN detection thresholds.

3. IDENTIFYING NARROW LINE AGN

3.1. *Line Ratio Diagnostics*

To identify AGN, we use the line ratio diagnostics defined by Ji & Yan (2020). These diagnostics involve the same line ratios used by Baldwin et al. (1981) and Veilleux & Osterbrock (1987), namely [O III]/Hβ (using [O III] λ5007), [N II]/Hα (using [N II] λ6584), and [S II]/Hα (using [S II] $\lambda\lambda$6716, 6731). These line ratios are each relatively insensitive to spectrophotometric calibration errors or dust. The line

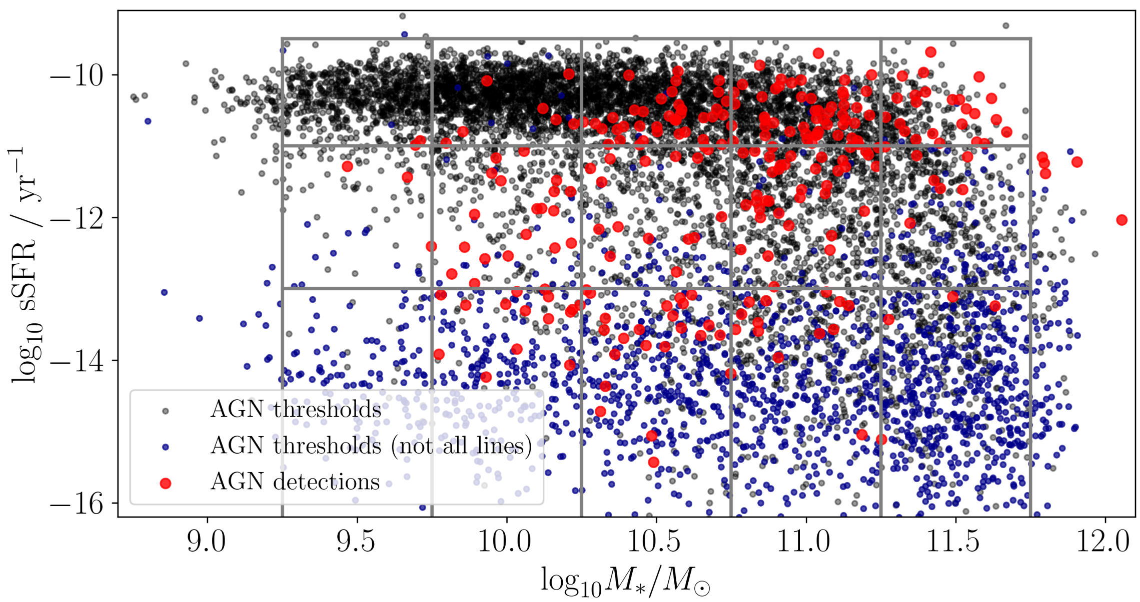

Figure 1. Stellar masses and sSFRs for the MaNGA sample used here. Red symbols show galaxies with AGN detections, black symbols show galaxies with AGN detection thresholds (which have all necessary line ratios detected), and blue symbols show galaxies with AGN detection thresholds (without all necessary line ratios detected). The boxes show divisions between stellar mass and sSFR samples we use to study the AGN luminosity and Eddington ratio distribution dependence on galaxy properties.

ratio diagrams of Baldwin et al. (1981) and Veilleux & Osterbrock (1987) typically are used as two separate two-dimensional classification tools. In contrast, Ji & Yan (2020) use three line ratios as a single three-dimensional classification space.

Using the photoionization model CLOUDY Ji & Yan (2020) found an informative reprojection of the three-dimensional space of these line ratios. We use this reprojection $P1$, $P2$, and $P3$, which allows a clean classification of spectra into those with ionizing spectra consistent with star formation, LINER, and Seyfert sources. It is defined as:

$$P1 = 0.63 \log_{10}\left(\frac{[\text{N II}]}{\text{H}\alpha}\right) + 0.51 \log_{10}\left(\frac{[\text{S II}]}{\text{H}\alpha}\right) + 0.59 \log_{10}\left(\frac{[\text{O III}]}{\text{H}\beta}\right)$$
$$P2 = -0.63 \log_{10}\left(\frac{[\text{N II}]}{\text{H}\alpha}\right) + 0.78 \log_{10}\left(\frac{[\text{S II}]}{\text{H}\alpha}\right)$$
$$P3 = -0.46 \log_{10}\left(\frac{[\text{N II}]}{\text{H}\alpha}\right) - 0.37 \log_{10}\left(\frac{[\text{S II}]}{\text{H}\alpha}\right) + 0.81 \log_{10}\left(\frac{[\text{O III}]}{\text{H}\beta}\right) \quad (2)$$

Ji & Yan (2020) analyze the interpretation of this reprojection thoroughly in terms of ionization models using both star formation related ionization spectra (i.e. stellar populations) and Seyfert/LINER-related ionization spectra (power law spectra). A simplified version of their interpretation is as follows. $P1$ correlates with the hardness of the ionization spectrum, and separates star formation related ionization spectra (low $P1$) from LINERs and Seyferts (high $P1$). $P2$ is negatively correlated with metallicity for all classes. $P3$ is correlated well with the ionization parameter $U =$

$\Phi_i/n_H c$, where Φ_i is the flux of ionizing photons per unit area per unit time, n_H is the hydrogen atom density, and c is the speed of light. For star formation related ionization (low $P1$), $P3$ also anticorrelates with metallicity in the models of Ji & Yan (2020), because metallicity determines the ionizing spectrum. For the LINER and Seyfert models (high $P1$), $P3$ correlates primarily with U.

Thus, we can select Seyfert galaxies as high $P1$, high $P3$ galaxies; in Ji & Yan (2020), the left panel of their Figure 15 illustrates the distribution for SDSS Legacy Main Sample galaxies (Strauss et al. 2002). In this paper, we consider $P1 > -0.3$ and $P3 > 0.55$ to define a Seyfert. We assume that the lower $P3$, LINER-like sources are powered by sources other than AGN (see Sarzi et al. 2010; Yan & Blanton 2012; Belfiore et al. 2016 and references therein).

The left panel of Figure 2 shows the distribution of MaNGA galaxy central emission line ratios in the space of $P1$ and $P3$; only the 7,672 galaxies with all relevant emission lines detected at 2σ are shown. The galaxies with emission line ratios dominated by star formation form the sequence on the left. The LINER-like galaxies are on the bottom right. The Seyfert galaxies are on the upper right, and our selection is shown as the blue box. The right panel of Figure 2 shows the image of one galaxy at each corresponding $P1$–$P3$ value. The r-band image is shown as grey scale, and the Hα map is overlaid in red and the [O III] λ5007 map is overlaid in green. The metallicity sequence on the right (increasing from top to bottom) is visible as the images go from [O III] to Hα dominated. The LINER [O III] emission tends to be subdominant to the galaxy, but for Seyferts it is more visible (cf. the right hand panel of Figure 15 of Ji & Yan 2020 showing the Hα equivalent widths as a function of $P3$).

There are 306 detected Seyfert galaxies in the full MaNGA sample by these definitions, 265 in the statistical sample (Primary, Color-Enhanced, and subsampled Secondary samples), and 210 of those have $\sigma_e > 60$ km s^{-1}. For each detected Seyfert we assume the central [O III] λ5007 luminosity (L([O III])) and Hβ luminosity (L(Hβ)) reflect mostly the light from a central AGN source (we discuss dust corrections in Section 3.3.2).

3.2. *Determining Identification Thresholds*

The galaxies in the sample (the majority) that are not identified as Seyfert are best understood as not having high enough Seyfert emission to be detectable. Many of them could, and most likely do, host lower luminosity Seyfert emission surrounding a central supermassive black hole that does not perturb the central line ratios enough to be detectable. Following the methodology of Trump et al. (2015), we determine detection thresholds for the luminosity of the Seyfert galaxy that could be in each galaxy yet remain undetectable according to our threshold estimates.

For each galaxy, we take its central line fluxes and their errors and we add to them a template set of line fluxes with nominal Seyfert-like line ratios:

$$\log_{10}\,[\text{O III}]/\text{H}\beta = 0.779$$

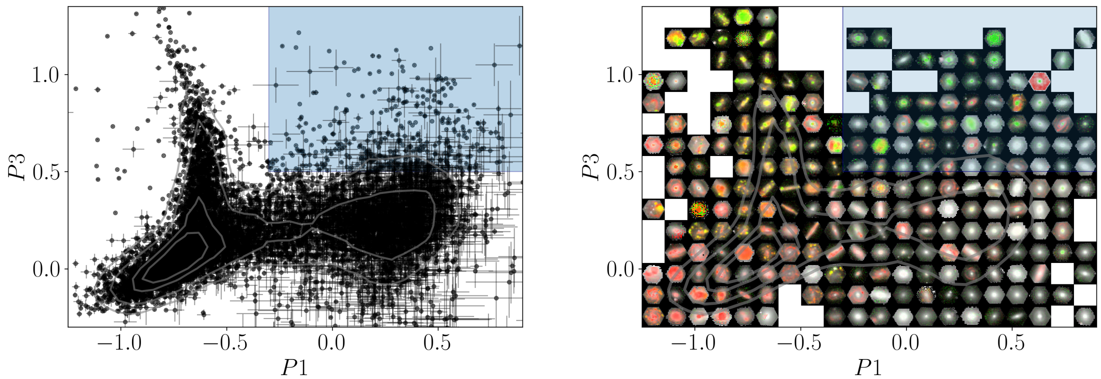

Figure 2. Line ratio diagnostic diagram of Ji & Yan (2020), showing $P1$ and $P3$ for MaNGA galaxy central emission line ratios. The left hand panel shows the measured values and errors, with the distribution overlaid. The right hand panel shows one randomly selected galaxy at each location in the plane. The r-band image is shown as a grey scale, with the Hα map contributing to the red channel and the [O III] λ5007 map contributing to the green channel. In both panels, the blue box shows our Seyfert galaxy classification.

$$\log_{10}[\text{N II}]/\text{H}\alpha = -0.017$$
$$\log_{10}[\text{S II}]/\text{H}\alpha = -0.233 \quad (3)$$

which are equivalent to $P_1 = 0.330$, $P_2 = -0.171$, and $P_3 = 0.725$.

Scaling these according to L([O III]), we then find the AGN luminosity L([O III]) that would perturb the line ratios enough to both exceed the 2σ detection limit in each line and cross into our Seyfert selection. Figure 3 illustrates the technique for a subset of the galaxies. The lines show the path in $P1$–$P3$ space as the AGN luminosity is increased. Note that some lines originate from a large black point on the plot (those galaxies who already have detected lines at 2σ), and some lines do not (those galaxies for which one or more lines are not detected already).

This process yields an AGN luminosity detection threshold in L_{lim}([O III]) and L_{lim}(Hβ) (and similar quantities for the dust-corrected luminosities) for each galaxy in our sample that does not have an actual detection. We use these thresholds below to help constrain the Eddington ratio distributions.

3.3. *Eddington Ratio Determination*

3.3.1. *Eddington Ratio Definition*

The Eddington ratio is defined as the ratio of the bolometric luminosity to the Eddington luminosity, where the latter is the luminosity that would prevent gravitational accretion of hydrogen gas in a spherically symmetric flow, which is $L_{\text{Edd}} = 1.38 \times 10^{38} (M_{\text{BH}}/M_\odot)$ erg s^{-1}. The Eddington ratio is therefore closely related to the observed luminosity divided by the black hole mass. As such, it provides a reasonable characterization of the level of "activity" associated with the black hole accretion. However, as defined it requires converting the observed luminosity in the AGN signature, in our case L([O III]) or L(Hβ), to bolometric luminosity, and it requires an

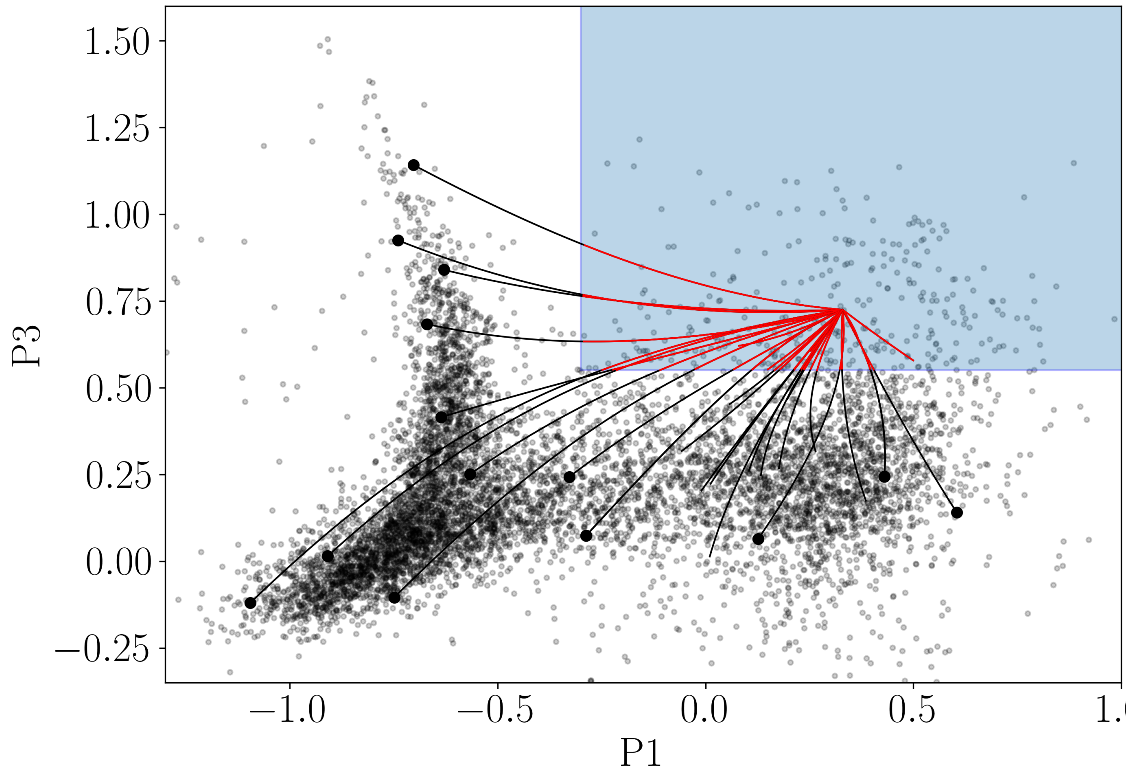

Figure 3. Same line ratio diagnostic diagram as in Figure 2, showing the same galaxies, illustrating our determination of L([O III]) detection thresholds for non-Seyfert galaxies. The blue region shows the line ratios for which we define Seyfert galaxies. The small grey points show each galaxy with all five lines detected at at least 2σ significance (error bars are omitted for clarity). The large black points are a random set of non-Seyfert galaxies chosen to span a range of metallicities and ionization spectral shapes (i.e. $P3$ and $P1$ values). The lines emanating from each large black point show how the line ratios change with an addition of an AGN component of increasing luminosity. Note that some lines exist without a large black point; those lines are examples for which the host galaxy's lines are not all detected at 2σ, and the lines start at AGN luminosities which would be necessary to make the lines detectable. The lines are colored red for AGN luminosities high enough to be identified as AGN by our methods, and black otherwise.

inference of the black hole mass. In this section we describe these two uncertain inferences.

3.3.2. *Bolometric Luminosities*

Determining $L_{\rm bol}$ is fraught with systematic issues. We use several different determinations to quantify the nature of the errors these issues might cause in our analysis. We consider conversions that are observationally based on the relationship between narrow lines and bolometric luminosity for Type 1 AGN, whose panchromatic emission can be measured along with the narrow lines (Heckman et al. 2004; Kauffmann & Heckman 2009), as well as the theoretical model of Netzer (2019).

An initial consideration is dust extinction. For a narrow line AGN, there is dust associated with the gas in the narrow line region itself, and there is dust in the host galaxy's interstellar medium that the light may have to pass through. With the data available for our sample, we cannot disentangle these effects. When we need to correct for dust extinction, we use a Balmer decrement-based approach. We measure the Hα/Hβ flux ratio, which assuming a universal attenuation curve then determines the absolute level of attenuation. For this purpose, we assume the attenuation curve of Whitford (1958) as parametrized by Miller & Mathews (1972). For those galaxies without measured Hα or Hβ lines, even though they are not detected AGN (since they do not have all of the necessary lines detected), we still need to calculate a dust correction to calculate the dust-corrected luminosity detection threshold. For this purpose, we use the median dust correction for the detected AGN ($A_{\rm H\alpha} = 0.949$ mag).

Kauffmann & Heckman (2009) applies a Balmer decrement-based dust correction to $L([\textrm{O III}])$, producing a dust-corrected luminosity $L([\textrm{O III}])_{\rm corr}$. They then apply a factor of 600 to obtain a bolometric luminosity:

$$\log_{10} L_{\rm bol} = 2.778 + \log_{10} L([\textrm{O III}])_{\rm corr} \quad (4)$$

They chose this conversion because it was roughly between the corrections cited by Ho (2008) for lower luminosity AGN and those cited by LaMassa et al. (2009) (referred to as an in preparation paper in Kauffmann & Heckman 2009) for higher luminosity AGN. Ho (2008) quotes an Hα-based bolometric correction, which they state is based on a Ho et al. (in prep) paper, which may be related to Ho & Kim (2009), though the latter does not quote any bolometric corrections. Regardless, Kauffmann & Heckman (2009) converts the Hα-based correction to [O III], and finds that the bolometric correction factor is 300–600. LaMassa et al. (2009) calculates the bolometric correction factor for higher luminosity AGN as 500-900 based on comparisons to mid-infrared luminosities. Kauffmann & Heckman (2009) motivate the choice of their factor of 600 as being approximately between these two ranges.

An alternative conversion from Heckman et al. (2004) uses uncorrected [O III] luminosities:

$$\log_{10} L_{\rm bol} = 3.544 + \log_{10} L([\textrm{O III}]) \quad (5)$$

They derive this conversion from the ratio of optical continuum to [O III] λ5007 luminosity for broad-line quasars (Kauffmann et al. 2003a; Zakamska et al. 2003) and the conversion from optical continuum to bolometric from Marconi et al. (2004). Because in their first step they do not correct either the line or continuum luminosity for dust, the conversion applies to uncorrected luminosities. The difference between Equation 5 and Equation 4 is about 0.2 dex (or about 0.5 mag) larger than the difference we would infer from the median dust correction quoted above.

Whether to base a bolometric luminosity on dust-corrected luminosities or not is unclear. The ratio between the optical [O III] and the mid-infrared [O IV] lines suggest that in Type 1 AGN the narrow lines experience less dust attenuation than do Type 2 sources (Diamond-Stanic et al. 2009), by a factor similar to the difference between the conversions in Equation 4 and 5 (i.e. 1–2 mag). So the two conversions appear roughly consistent, and applying the dust correction should lead to higher accuracy. Yet LaMassa et al. (2010) show that correlations between [O III] and presumably less dust-affected indicators like the mid-infrared luminosity, [O IV], and hard X-rays are better with $L([\text{O III}])$ than with $L([\text{O III}])_{\text{corr}}$.

The bolometric correction may also depend on black hole mass, especially for $L([\text{O III}])$, due to the changing spectrum of the accretion disk predicted by models like those of Shakura & Sunyaev (1973). Cann et al. (2019) explore this dependence and find that at high masses ($M_{\text{BH}} \sim 10^8$) most of the oxygen is singly- or doubly-ionized, that as black hole mass decreases more of the oxygen is more highly ionized, and that at a given accretion rate in their models $L([\text{O III}])$ is correlated with black hole mass.

Whether to account for a luminosity dependence in the bolometric correction is also unclear. Ho (2008), Kauffmann & Heckman (2009), and others claim that the bolometric corrections rise with luminosity. Theoretically, Netzer (2019) and Cann et al. (2019) predict bolometric corrections that rise with luminosity.

Netzer (2019) makes a specific prediction for Hβ; he does not perform the calculations necessary to make a reliable prediction for [O III]. His prediction is:

$$\log_{10} L_{\text{bol}} = 45.661 + 1.18 \left(\log_{10} L(\text{H}\beta)_{\text{corr}} - 42\right), \quad (6)$$

(where we write $L(\text{H}\beta)_{\text{corr}}$ to indicate that this luminosity needs to be dust-corrected to use this theoretical model). Additionally, Netzer (2019) find that black hole mass and spin variations produce anywhere from 0.2–1 dex in peak-to-peak variations of $L(\text{H}\beta)$, depending on the bolometric luminosity. These models therefore predict considerable scatter in the bolometric corrections.

In short, bolometric corrections from narrow line luminosities are highly uncertain, despite extensive efforts to determine them. We derive three different bolometric luminosities meant to span the space of possibilities: (a) $L([\text{O III}])$ uncorrected for internal dust extinction and apply Equation 5 to $L([\text{O III}])$; (b) $L([\text{O III}])_{\text{corr}}$ corrected for internal dust extinction using the Balmer decrement and apply Equation 4; and (c) $L(\text{H}\beta)_{\text{corr}}$ corrected for internal dust exinction using the Balmer decrement and apply Equation 6. Our figures below show the results for the $L(\text{H}\beta)_{\text{corr}}$-based bolometric luminosities, but the differences among these approaches turn out to not affect our qualitative results.

3.3.3. *Black hole masses*

Estimating black hole masses is similarly fraught, and for our sample these masses cannot be obtained directly. Instead we use the relationship between black hole mass and stellar velocity dispersion (M_{BH}–σ) calibrated using high angular resolution stellar or gas dynamics for nearby galaxies (see Kormendy & Ho 2013 and references therein). We use three versions of this relationship to test the sensitivity of our results to this choice.

For the stellar velocity dispersion, we use σ_e within a half-light radius as estimated by the MaNGA DAP. As noted above, we will use $\sigma_e > 60$ km s^{-1}. This quantity is not precisely the quantity used by most $M_{\rm BH}$–σ calibrations, which usually calculate σ^2 from the light-weighted average of $V^2 + \sigma^2$ within the half-light radius, where V is the mean stellar velocity in the galaxy's frame. At the level of accuracy we are working, we expect this difference to be a small effect, and in any case the physical resolution available for MaNGA makes a central estimate of this quantity difficult.

The form of $M_{\rm BH}$–σ is usually taken to be a power law:

$$\log_{10} M_{\rm BH} = \alpha + \beta \log_{10}\left(\frac{\sigma}{200 \text{ km s}^{-1}}\right) \tag{7}$$

with $\beta \sim$ 4–6 depending on the analysis. Kormendy & Ho (2013) find $\beta = 4.4 \pm 0.3$ and $\alpha = 8.50 \pm 0.05$ (specifically, their Equation 5). Graham & Scott (2013) find $\beta = 5.2 \pm 0.3$ and $\alpha = 8.14 \pm 0.04$ (we use their regression given at the end of their Section 3.1 of black hole mass on σ for barred and unbarred galaxies together, which of their methods is the most appropriate one for inferring the $M_{\rm BH}$ from σ). Finally, fits to this relationship determined only for local AGN find a lower zeropoint and shallower slope, for example $\alpha = 7.86 \pm 0.04$ and $\beta = 3.65 \pm 0.13$ (Greene & Ho 2006).

The samples of non-AGN galaxies that lead to different values for α and β samples tend to have overlapping samples, but the values are affected by precisely which sample is used (Park et al. 2012). The flattening and shift of the AGN-based sample may be affected by the inability to measure black hole masses dynamically for the highest mass black holes when there is an AGN (Woo et al. 2013).

In short, the black hole masses are rather uncertain. We use the three relations quoted above (Kormendy & Ho 2013; Graham & Scott 2013; Greene & Ho 2006), and evaluate whether the exact adopted black hole masses matter, which qualitatively they do not.

4. LUMINOSITIES AND EDDINGTON RATIOS

Figure 4 shows the dust-corrected Hβ luminosities and luminosity thresholds for the AGN. We show the dependence on stellar mass, velocity dispersion σ_e, star formation rate, and specific star formation rate. Although we plot the values and detection thresholds for the dust-corrected Hβ luminosity $L({\rm H}\beta)_{\rm corr}$ here, the same analysis produces results for $L({\rm H}\beta)$, $L([\text{O III}])$, and $L([\text{O III}])_{\rm corr}$.

This figure illustrates the difficulty in interpreting the narrow-line AGN population statistics. Only a handful of narrow-line AGN are so luminous that they would be detectable in all galaxies. For a typical identified AGN luminosity of $L({\rm H}\beta)_{\rm corr} \sim 10^{40}$ erg s^{-1}, it would be detectable in about 65% of the galaxies. As Figure 4 demonstrates, the detectability is a strong function of host galaxy properties. Therefore, the shape of the luminosity distribution and its dependence on host galaxy properties must be dramatically affected by these selection effects.

Trump et al. (2015) and Jones et al. (2016) have both explored these detection thresholds in detail before, in the context of the SDSS Legacy Main Sample of galaxies (Strauss et al. 2002). Comparing to Figure 8 of Trump et al. (2015), we find a similar range of detection thresholds when expressed in the units they use, with $\log_{10} L([\mathrm{O\ III}])_{\mathrm{corr}}/M_* \sim 29$–31, expressing the ratio in units of ergs s^{-1} $M_\odot^{-1}$.

Heckman et al. (2004) also analyzed the completeness effects in the same SDSS Legacy Main Sample, claiming completeness above $L([\mathrm{O\ III}]) \approx 10^6\ L_\odot \sim 4 \times 10^{39}$ ergs s^{-1} (not dust corrected) for all potential host galaxies. Figure 8 in Trump et al. (2015) shows that much of the host galaxy parameter space actually has a higher threshold for detection. For example, accounting for a typical dust correction of 0.6 dex, a star forming galaxy with $\log_{10} M_* \sim 10$ has a threshold of $L([\mathrm{O\ III}]) \sim 10^{40}$ ergs s^{-1} and a galaxy at $\log_{10} M_* \sim 11$ has a threshold of $L([\mathrm{O\ III}]) \sim 10^{41}$ ergs s^{-1}, much higher than Heckman et al. (2004) claim. The numbers from our analysis are not directly comparable because we use a different data set and selection method, but our conclusions are similar to those of Trump et al. (2015), with only 20% of potential host galaxies in the MaNGA sample having thresholds as sensitive as Heckman et al. (2004) claims for all galaxies. Our conclusion is that the claimed completeness limits of Heckman et al. (2004) cannot be used to accurately infer AGN host galaxy population properties either in the SDSS Legacy Survey Main Sample or the MaNGA sample.

Figure 5 shows the Eddington ratios and Eddington ratio detection thresholds derived from $L(\mathrm{H}\beta)$, using the Greene & Ho (2006) M_{BH}–σ relation and the Netzer (2019) bolometric corrections. These values look as expected given Figure 4, again illustrating that the selection thresholds must play an important role in shaping the distribution of identified AGN luminosities and of their host galaxy properties.

We point out two features in particular. First, the detection thresholds in Eddington ratio become quite high at low velocity dispersion (and to a lesser extent, at low mass). This feature is caused by the strong assumed dependence of black hole mass on velocity dispersion; although the $L(\mathrm{H}\beta)$ thresholds are not a strong function of σ_v (Figure 4), these thresholds correspond to much higher Eddington ratios at low σ_v.

Second, at high sSFR, the thresholds tend to exceed the luminosities of most detected AGN, and . This effect is mostly due to the greater degree of contamination by star formation related narrow line emission at high sSFR, and partly due to an anti-correlation of sSFR and σ_v, which therefore leads to higher Eddington ratios at higher sSFR.

These sorts of effects complicate the interpretation of the distribution of host galaxy properties of AGN. However, it is useful to consider the distribution AGN detections as a function of stellar mass and sSFR, before any correction for these effects. Figures 6 and 7 show these fractions for luminosity-selected and Eddington ratio-selected AGN, respectively. There is a strong stellar mass dependence for star-forming galaxies. Both intermediate and high sSFR galaxies show an enhancement in AGN relative to quiescent galaxies, with about the same fraction at $M_* \sim 10^{10} M_\odot$, and a more dra-

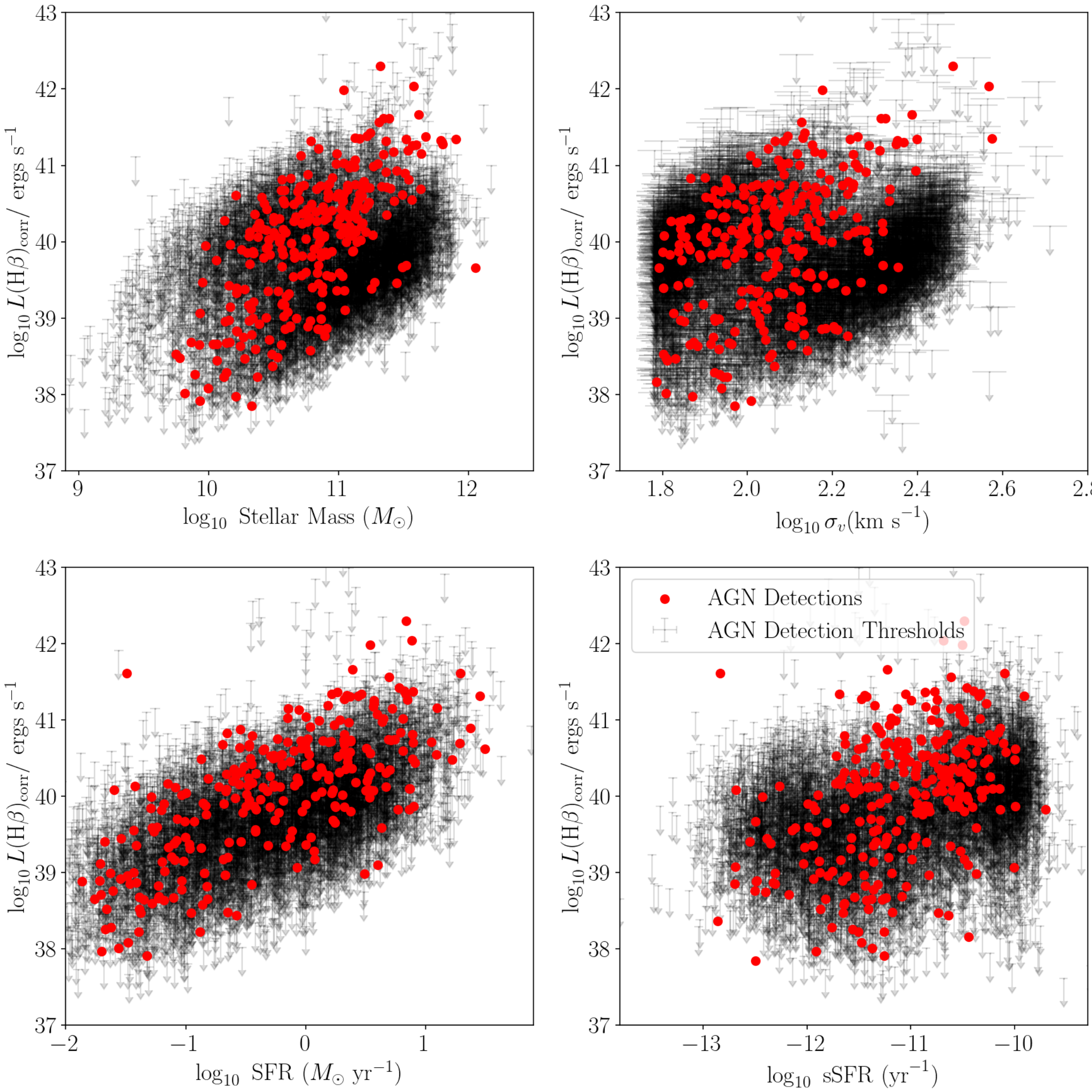

Figure 4. AGN-related dust-corrected Hβ luminosities and detection thresholds for MaNGA galaxies, as a function of various host galaxy properties: stellar mass (upper left), stellar velocity dispersion σ_v (upper right), SFR (lower left), and sSFR (lower right). The red points are detections, and the black upper limits are detection thresholds.

matic correlation with sSFR at higher masses. However, these results do not reflect the total AGN fraction because of the severe selection effects on these detection rates.

5. NARROW-LINE LUMINOSITY DISTRIBUTIONS

5.1. *Methodology*

In this section we estimate the narrow-line luminosity distributions for L([O III]), $L([\mathrm{O\,III}])_{\mathrm{corr}}$, $L(\mathrm{H}\beta)$, and $L(\mathrm{H}\beta)_{\mathrm{corr}}$, as a function of mass and sSFR. We use a parametrized form for these functions, based on the Schechter function:

$$\Phi(L) = \begin{cases} \phi_0 \left(\frac{L}{L_*}\right)^{\alpha} \exp\left(-L/L_*\right) & \text{if } L > L_{\mathrm{min}} \\ 0 & \text{if } L \leq L_{\mathrm{min}} \end{cases} \tag{8}$$

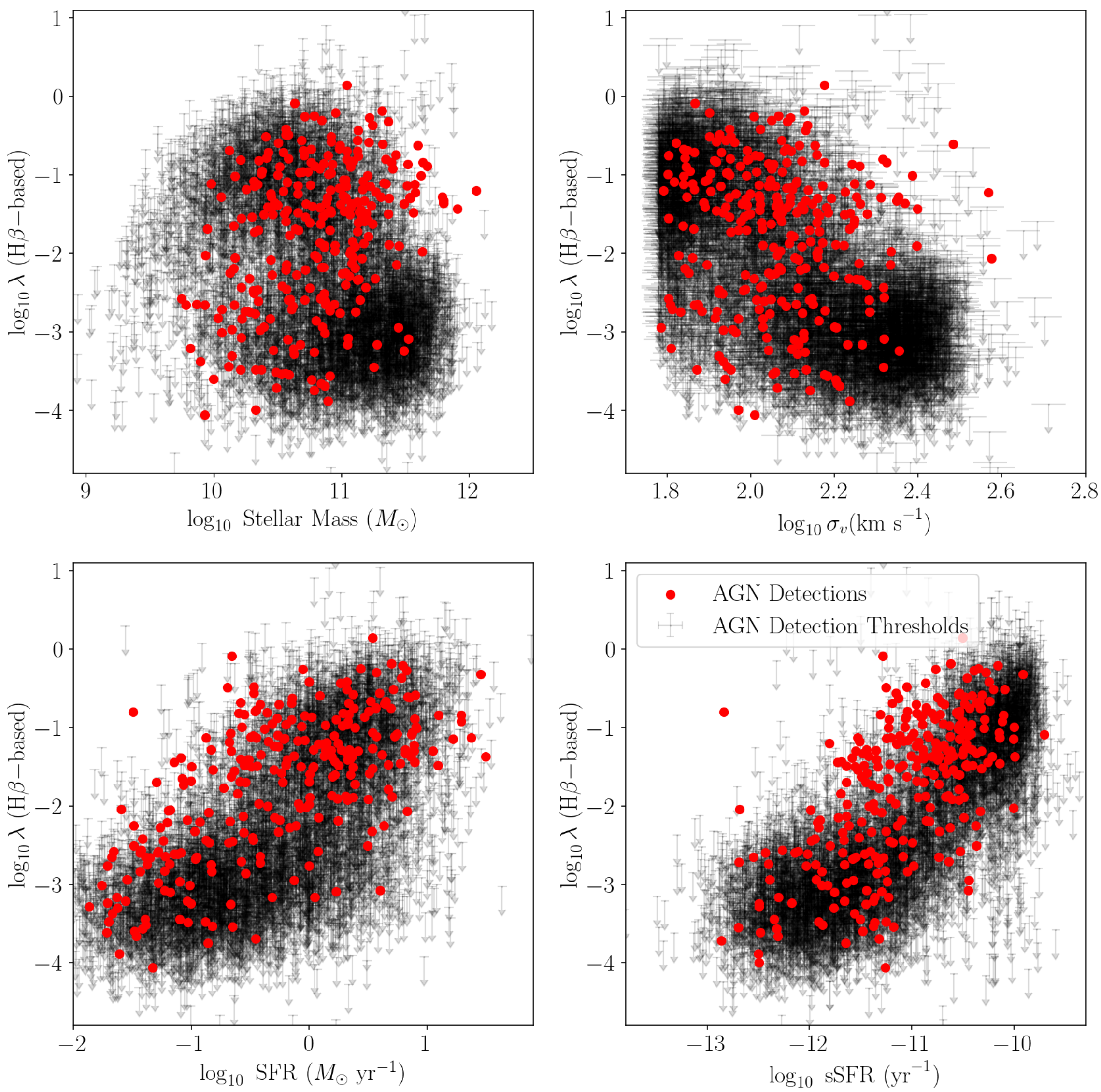

Figure 5. Similar to Figure 4, for Eddington ratios and Eddington ratio detection thresholds. We use the Greene & Ho (2006) $M_{\rm BH}$–σ relation and the Netzer (2019) bolometric corrections to calculate the Eddington ratios; the qualitative trends do not depend on these choices.

where $\Phi(L)$ is per unit luminosity, and there is a minimum luminosity $L_{\rm min}$ below which $\Phi(L) = 0$. In Equation 8, ϕ_0 is defined such that the integral of $\Phi(L)$ is unity, so the free parameters of the model are α, L_*, and $L_{\rm min}$.

As we show below, with our data set the individual parameters of this Schechter function are highly covariant and not particularly well-constrained. However, based on a given set of parameters we can define $F_{\rm AGN,Y}$ as the fraction of galaxies with AGN above a particular luminosity L_Y, where $Y = \log_{10}(L_Y/({\rm erg~s^{-1}}))$. It turns out that for a choice of $Y \sim 39$–40, near the typical luminosity thresholds seen in Figure 4, $F_{\rm AGN,Y}$ is well constrained.

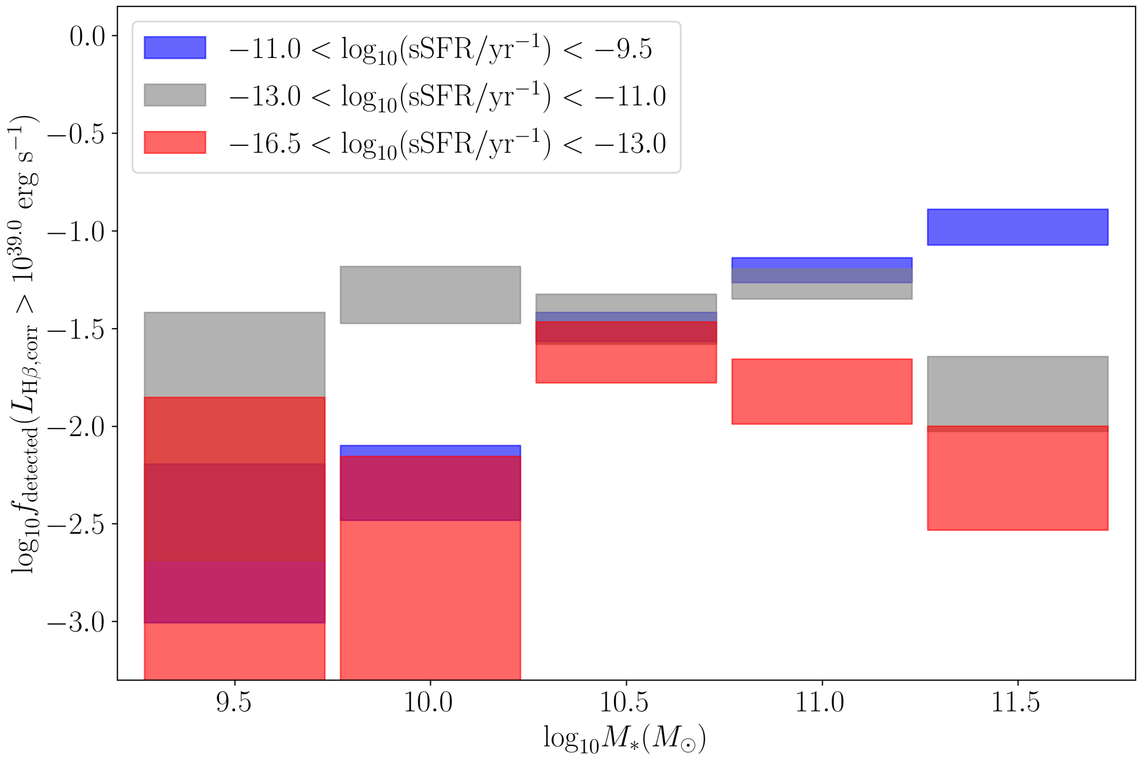

Figure 6. AGN raw detection rate $f_{\mathrm{detected}}(L(\mathrm{H}\beta)_{\mathrm{corr}} > 10^{39}$ erg s^{-1}) as a function of stellar mass and sSFR. These detection rates have no correction for the rather severe selection effects and therefore should not be taken at face value. The bands show 68% confidence intervals (using the Clopper-Pearson method for binomial statistics; Clopper & Pearson 1934).

For any sample or subsample of galaxies with luminosities and upper luminosities we can define the log-likelihood of the data as:

$$\ln P(\{L_i, L_{\mathrm{lim},i}\}|\alpha, L_*) = \left(\sum_i \ln \Phi(L_i)\right) + \left(\sum_i \ln \int_{L_{\mathrm{min}}}^{L_{\mathrm{lim},i}} \mathrm{d}L \Phi(L)\right), \qquad (9)$$

To constrain α and L_* for any sample or subsample, we take a Bayesian approach and perform Markov Chain Monte Carlo using `emcee` (Foreman-Mackey et al. 2013). We assume a uniform prior on α, $\log_{10} L_*$, and $\log_{10} L_{\mathrm{min}}$ over a finite range in each value. We limit the slope to $-2 < \alpha < 2.5$, and for $\log_{10} L_*$ and $\log_{10} L_{\mathrm{min}}$ we limit to ranges that depend on the exact sample (designed to be less than the minimum observed luminosity or luminosity threshold).

5.2. *Full sample analysis of the Hβ luminosity distribution*

To illustrate the results, we first consider the full sample, undifferentiated by stellar mass or sSFR. We fit the distribution of the dust-corrected Hβ luminosity, $L(\mathrm{H}\beta)_{\mathrm{corr}}$.

The left panel of Figure 8 illustrates the bivariate posterior distributions among α, L_*, L_{min}, and $F_{\mathrm{AGN},39}$ for this case. The Schechter parameters are reasonably well-constrained for the full sample, but with a strong degeneracy between α and

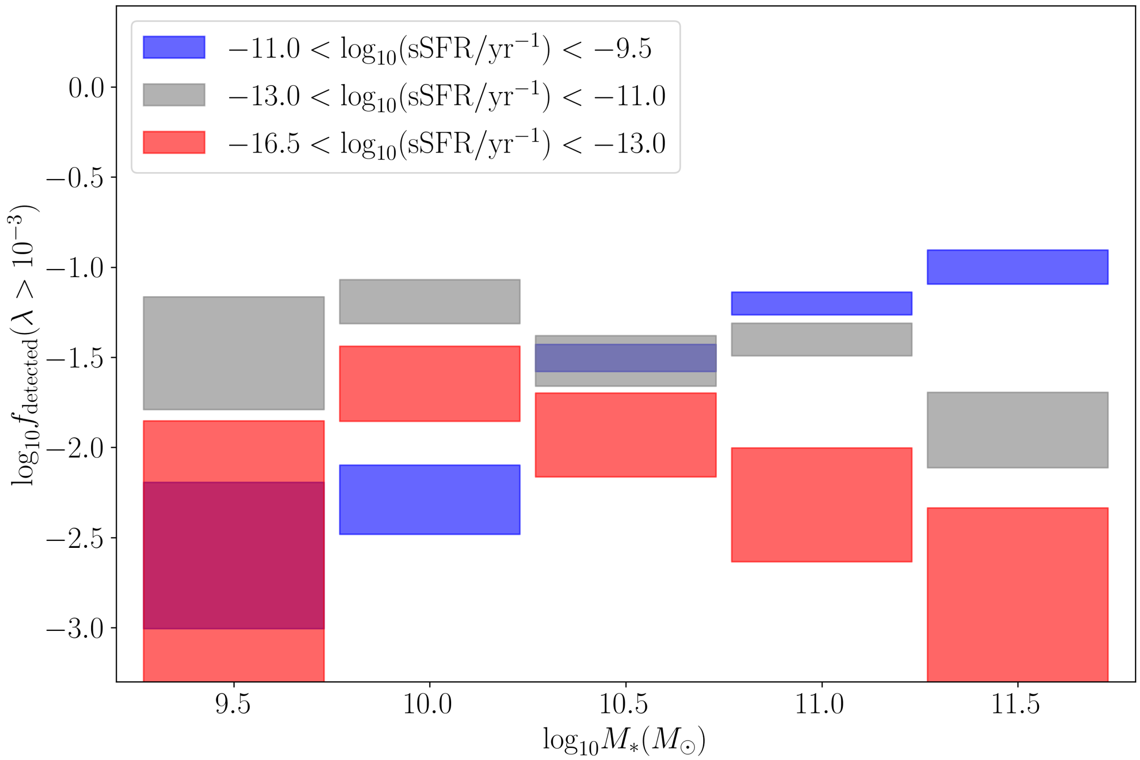

Figure 7. Similar to Figure 6, for Eddington ratio selected AGN with $f_{\mathrm{detected}}(\lambda > 10^{-3})$. These detection rates have no correction for the rather severe selection effects and therefore should not be taken at face value.

L_{min}. It is straightforward to understand the origin of this degeneracy; for example, to increase the number of non-detections, the model can either increase α or decrease L_{min}. Whereas these model parameters are hard to interpret in the context of such degeneracies, $F_{\mathrm{AGN,39}}$ is relatively well-defined, and it is does not share a degeneracy with the model parameters (i.e. the model parameter degeneracies are along lines of constant $F_{\mathrm{AGN,39}}$).

The right panel shows $F_{\mathrm{AGN,Y}}$ as a function of $Y = \log_{10} L_c$ in erg s^{-1}. The thick line is the mean, the thin lines are 1,000 samples of the posterior, and the pink band shows the standard deviation of $\log_{10} F_{\mathrm{AGN}}$ around the mean. The standard deviation is narrowest around $\log_{10} F_{\mathrm{AGN}} \sim 39$, which is why we picked that threshold.

We can also check whether this model does a good job at explaining the distribution of the luminosities of detected AGN. Using a Kolmogorov-Smirnov (KS) statistic based on this distribution, we find that it is worse than the observed sample for a fraction $p = 0.43$ of Monte Carlo samples generated under our model; that is, the observed distribution is consistent with the expectation under the model. This result suggests that this model is flexible enough to explain the observations.

5.3. *Stellar mass and sSFR dependence*

To study the stellar mass and sSFR dependence of the luminosity, we have divided the sample into twelve bins, four in $\log_{10} M_*/M_\odot$ with bin

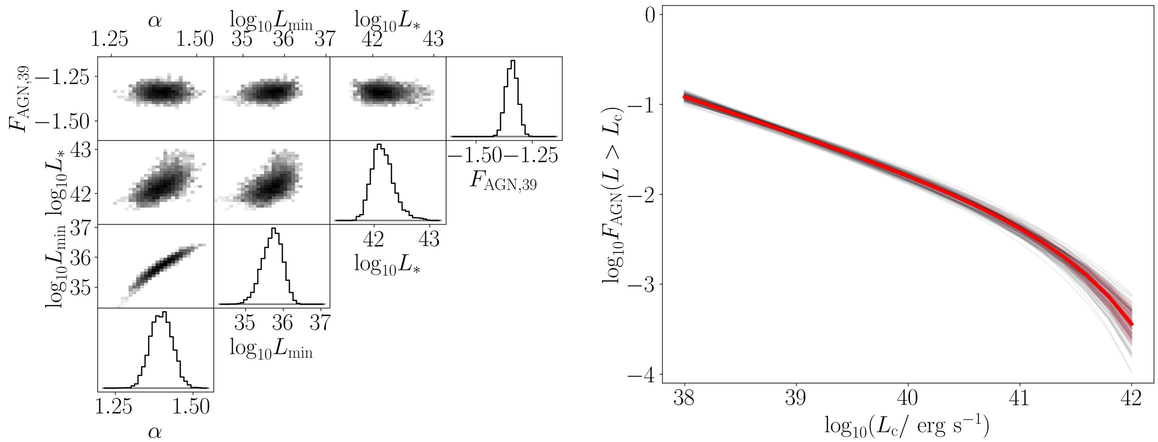

Figure 8. *Left panel:* Bivariate posterior distributions among α, L_*, $L_{\rm min}$, and $F_{\rm AGN,39}$, for the analysis of the entire sample of AGN and AGN non-detections, using $L(\mathrm{H}\beta)_{\rm corr}$. This analysis include all stellar masses and sSFRs, and we show later that the best fit parameters depend strongly on those galaxy properties. *Right panel:* $F_{\rm AGN,Y}$ as a function of $Y = \log_{10} L_c$ in erg s^{-1}. The thick line is the mean, the thin lines are 1,000 samples of the posterior, and the pink band shows the standard deviation of $\log_{10} F_{\rm AGN}$ around the mean.

edges $[9.25, 10.25, 10.75, 11.25, 11.75]$, and three in sSFR with bin edges $[-16.5, -13.0, -11.0, -9.5]$. Figure 1 depicts these bins.

In each bin, we perform the analysis described in Section 5.2 to constrain the distribution of $L(\mathrm{H}\beta)_{\rm corr}$. The individual posterior samples for each bin behave qualitatively similarly to that for the full sample, with strong degeneracies between α and $L_{\rm min}$ but with $F_{\rm AGN,39}$ better constrained.

Figure 9 shows the resulting dependence of $F_{\rm AGN,39}$ as a function of stellar mass and sSFR. This plot contains the first important result of this paper: the dependence of AGN occurrence rate as a function of galaxy properties, corrected for selection effects, with "AGN" defined as having $L(\mathrm{H}\beta)_{\rm corr} > 10^{39}$ erg s^{-1}. For star-forming galaxies, there is a clear and steady mass dependence. Quiescent galaxies have a much lower occurrence rate of AGN and less dependence on mass. The occurrence rate for the central sSFR bin ("green valley" galaxies) sits in between that of the highest and lowest sSFR bins, except for at the lowest masses where the green valley galaxies have the highest occurrence rates. There is also some evidence of a peak occurrence rate at $M_* \sim 10^{11} M_\odot$ from green valley galaxies.

There is little dependence of the results on the detection significance we require on emission lines, if we raise the threshold of detection from 2σ to 3σ. If we change the luminosity threshold on $L(\mathrm{H}\beta)_{\rm corr}$ from 10^{39} erg s^{-1} to 10^{40} erg s^{-1}, $F_{\rm AGN}$ drops in all bins by about 0.2–1.0 dex, with larger drops for the quiescent galaxies than star-forming galaxies, but the qualitative trends remain; a difference is that for the higher luminosity thresholds, there are no mass bins in which green valley galaxies have the highest occurrence rates.

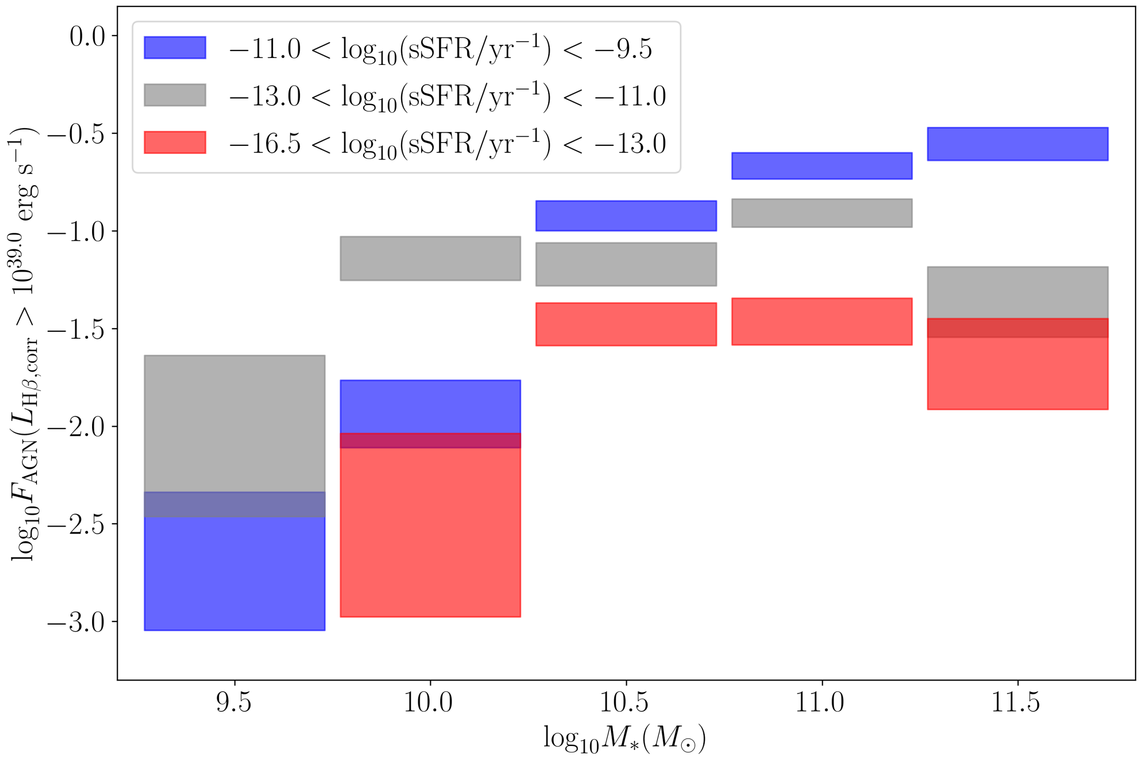

Figure 9. AGN occurrence rate $F_{\mathrm{AGN}}(L(\mathrm{H}\beta)_{\mathrm{corr}} > 10^{39}\ \mathrm{erg\ s^{-1}})$, based on model fits (i.e. corrected for selection effects), as a function of stellar mass and sSFR. The bands are $\pm 1\sigma$ uncertainties based on the posterior distribution.

We have also examined the luminosity distribution of $L([\mathrm{O\ III}])_{\mathrm{corr}}$, finding very similar trends, though with $L([\mathrm{O\ III}])_{\mathrm{corr}}$ about ~ 0.2 dex median difference in the luminosities. There is a spread of about 1 dex in the differences between $L([\mathrm{O\ III}])_{\mathrm{corr}}$ and $L(\mathrm{H}\beta)_{\mathrm{corr}}$, but these differences only have a minor effect on the trends in the luminosity distributions.

The thresholds depend on our choice for the nominal template AGN (Equation 3). We have therefore chosen an alternative set of line ratios that are more extreme in the Seyfert direction, equivalent to $P1 = 0.544$, $P2 = -0.156$, $P3 = 0.907$. Using these line ratios for the nominal AGN changes the results in minor ways, similar to or less than the uncertainties.

Finally, we have explored other choices of the $P1$–$P3$ selection criteria. If we make them more inclusive, we do not find consistent results—both the detection rates and the occurrence rates increase. This result indicates that the distribution of line ratios is not well described by the model AGN luminosity distribution beyond these limits. If we make them less inclusive, the detection rates decrease but the occurrence rates remain roughly constant, though the uncertainties rise. This result suggests that the AGN model distribution does explain the distribution of observed line ratios in the domain of our standard selection criteria.

6. EDDINGTON RATIO DISTRIBUTIONS

6.1. *Methodology*

In this section we estimate the narrow-line Eddington ratio distributions. In this section, we estimate Eddington ratio estimates based on the bolometric luminosity estimated from the dust-corrected Hβ luminosity with Netzer (2019) and the black hole mass estimated from the velocity dispersion within $1R_e$ and the $M_{\rm BH}$–σ relation of Kormendy & Ho (2013) (see Section 3.3.1).

We use the same Schechter function form and analysis techniques that we used for the luminosity distribution in Section 5.1.

6.2. *Full sample analysis of the Eddington ratio distribution*

Again as an illustration of the method, we consider the Eddington ratio distribution from the full sample undifferentiated by stellar mass or sSFR. Figure 10 shows the resulting constraints on the parameters of the distribution. In this case, we define $F_{{\rm AGN},-3}$ as the fraction of AGN in the model with $\lambda > 10^{-3}$; as we found in the case of the luminosity distribution, this quantity is relatively stable and not degenerate with the model parameters. The Schechter fit proves a reasonably good fit to the data in this case, with a KS statistic of $p = 0.10$; the deviation from the model driving the KS statistic low manifests as a $\sim 30\%$ dip in the observed distribution around $\lambda_c \sim 10^{-3}$.

6.3. *Stellar mass and sSFR dependence*

To investigate the stellar mass and sSFR dependence of the Eddington ratio distribution, we use the same stellar mass and sSFR bins as we did for the luminosity distribution in Section 5.3, as illustrated in Figure 1.

Figure 11 shows the results for $F_{{\rm AGN},-3}$ as a function of stellar mass and sSFR. For this figure we use bolometric luminosities based on the dust-corrected $L({\rm H}\beta)_{\rm corr}$, and the relation of Netzer (2019) (Equation 6), and the $M_{\rm BH}$–σ relation of Kormendy & Ho (2013). The results are close to as expected based on Figure 9. Using a Eddington ratio threshold instead of a luminosity threshold is similar to (though not exactly the same as!) dividing by stellar mass. Therefore it tilts the stellar mass dependence downwards. The result is that at the higher masses, the dependence on stellar mass is flat for star-forming galaxies, declining strongly for the quiescent galaxies, and somewhere in between for the green valley galaxies. The lowest mass bin is again somewhat anomalous—for galaxies with $9.25 < \log_{10} M_*/M_\odot < 10.25$, there is a weak detection that green valley galaxies have the highest AGN occurrence rates.

We have explored the use of other $M_{\rm BH}$–σ calibrations and bolometric luminosity determinations. The $M_{\rm BH}$–σ from Kormendy & Ho (2013) yields somewhat lower $F_{{\rm AGN},-3}$, by about 0.1–0.3 dex depending on the bin, relative to the calibrations of Greene & Ho (2006), Gültekin et al. (2009), or Graham & Scott (2013), which all yield results similar to each other. Otherwise the results are qualitatively similar. The offset for Kormendy & Ho (2013) arises because their value of α (in Equation 7) is high, leading to higher black hole masses and consequently lower Eddington ratios.

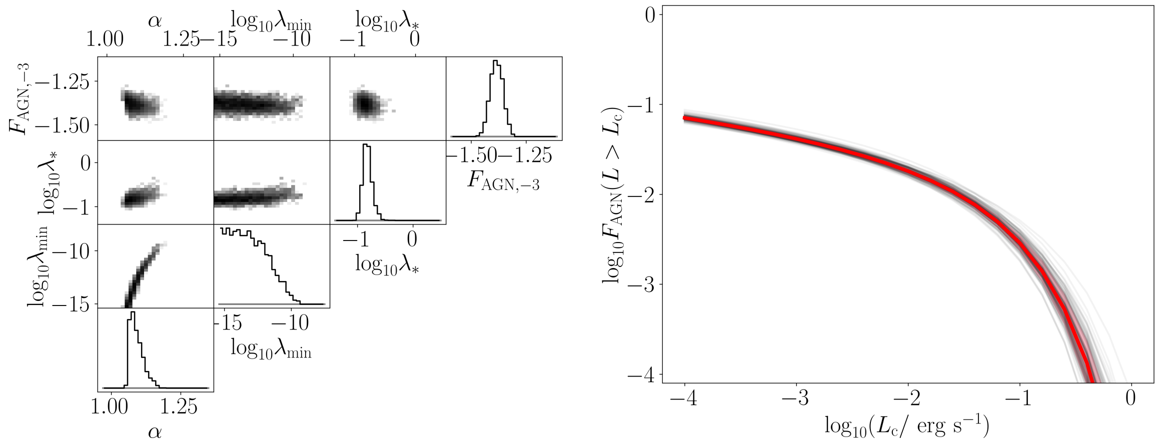

Figure 10. Similar to Figure 8, but for the Eddington ratio distribution instead of the luminosity distribution. The Eddington ratio is calculated using the calibration of Netzer (2019) for the bolometric luminosity and the $M_{\rm BH}$–σ relationship of Kormendy & Ho (2013). *Left panel:* Bivariate posterior distributions among α, λ_*, $\lambda_{\rm min}$, and $F_{\rm AGN,-3}$, for the analysis of the entire sample of AGN and AGN non-detections. *Right panel:* $F_{\rm AGN,Y}$ as a function of $Y = \log_{10}\lambda$. The thick line is the mean, the thin lines are 1,000 samples of the posterior, and the pink band shows the standard deviation of $\log_{10} F_{\rm AGN}$ around the mean.

The bolometric luminosities based on dust-corrected $L({\rm H}\beta)_{\rm corr}$ (Netzer 2019; Equation 6) yield results very similar to those based on the dust-corrected $L([{\rm O\ III}])_{\rm corr}$ (Kauffmann & Heckman 2009; Equation 4). The bolometric luminosities based on the uncorrected $L([{\rm O\ III}])$ (Heckman et al. 2004; Equation 5) are fairly similar but yield somewhat weaker dependence of $F_{\rm AGN,-3}$ on sSFR. This difference may be related to higher dust corrections in high sSFR galaxies, which in the case of the dust-corrected luminosities widens the difference in bolometric luminosities in these classes of galaxy, though the difference in the median dust corrections is only of order 0.2 dex between the low and high sSFR AGN samples.

As for the luminosity dependence of $F_{\rm AGN}$, there is little dependence on the detection significance (raising the threshold of line detection from 2σ to 3σ). If we change the threshold λ_c from 10^{-3} to 10^{-2}, $F_{\rm AGN}$ drops in all bins by about 0.2–0.8 dex, again with larger drops for the quiescent galaxies than star-forming galaxies. If we change the choice of the nominal AGN line ratios, the results alter only slightly. Otherwise the results are qualitatively similar.

7. COMPARISON WITH PREVIOUS RESULTS

Our analysis agrees qualitatively with the one other study that we have found of narrow-line AGN dependence on stellar mass and sSFR that accounts for selection effects in assessing their demographics, that of Trump et al. (2015). They study the SDSS Legacy Survey Main Sample galaxies (York et al. 2000; Strauss et al. 2002), which is a far larger sample than MaNGA, though with lower signal-to-noise ratio spectra and poorer measurements of stellar mass and sSFR. Using a narrow-line selection criteria based on Baldwin et al. (1981), their Figure 13 shows the detected

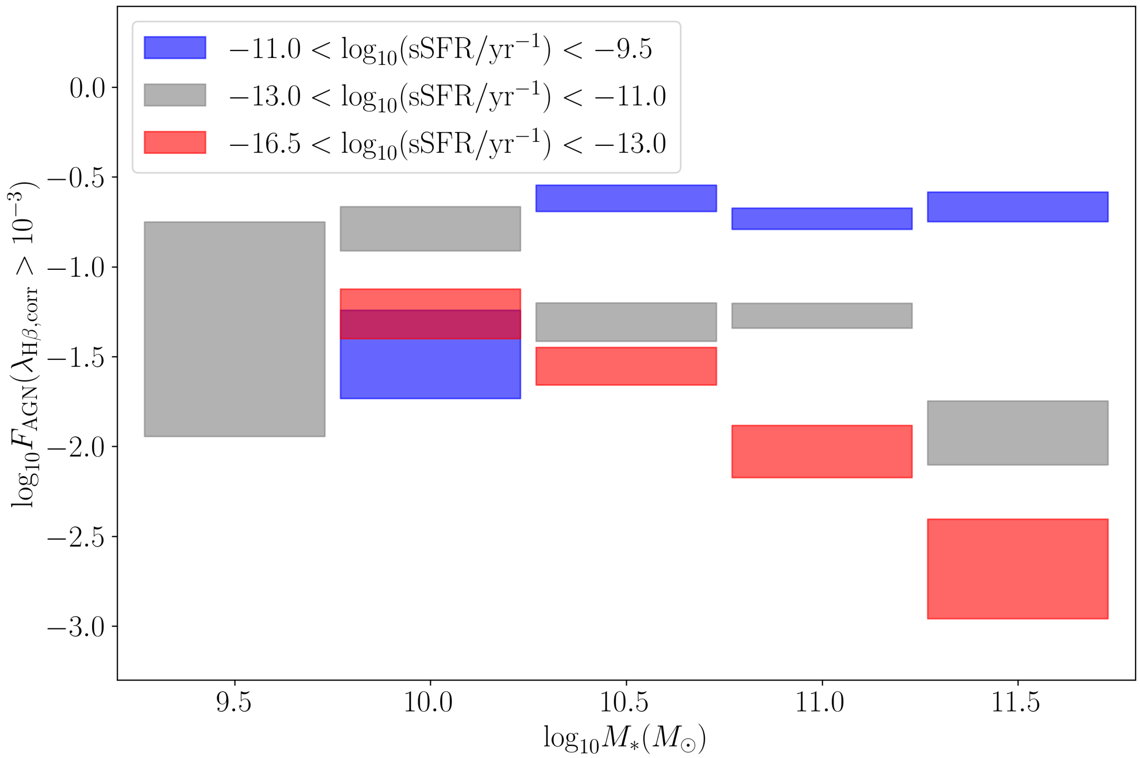

Figure 11. AGN occurrence rate $F_{\mathrm{AGN}}(\lambda > 10^{-3})$, based on model fits (i.e. corrected for selection effects), as a function of stellar mass and sSFR. The bands are $\pm 1\sigma$ uncertainties based on the posterior distribution. For this plot we use the M–σ relation of Kormendy & Ho (2013) and bolometric luminosities based on dust-corrected Hβ and the conversion of Netzer (2019).

AGN fraction as a function of stellar mass and sSFR, compared to a best-fit model with a uniform Eddington ratio distribution. The residuals are qualitatively similar to our results: relative to a constant Eddington ratio distribution, AGN occurrence increases with stellar mass for star-forming galaxies (though perhaps more strongly than we find at high mass), and decreases with stellar mass for quiescent galaxies.

Our results are not appropriate for comparison with most other investigations of the stellar mass and sSFR dependence of narrow-line AGN occurrence, because all other such investigations we found in the literature do not account for the selection effects. MaNGA has the highest signal-to-noise ratios available for a statistically interesting sample, i.e. considerably higher than the SDSS Legacy or DESI samples for typical galaxies. The MaNGA sample is also at lower redshifts, allowing the central spectrum to be better isolated from star-formation related contamination. Yet even for MaNGA we have shown that the detected AGN demographics are primarily shaped by the selection effects. Any analysis of narrow-line AGN host galaxy demographics that omits a consideration of selection effects is therefore likely to be inaccurate except for the most luminous AGN (e.g. Kauffmann et al. 2003b; Heckman et al. 2004; Martin et al. 2007; Salim et al. 2007; Kauffmann & Heckman 2009; Schawinski et al. 2010; Ellison et al. 2016; Sánchez et al. 2018; Man et al. 2019; Greene et al. 2020; McDonald et al. 2021; Woodrum et al. 2024; Pucha et al. 2025). Many of these studies conclude that narrow-line AGN are most frequent for galaxies with intermediate sSFRs, i.e. in the green valley, without investigating whether this effect is caused by the difficulty of detecting narrow line AGN in galaxies with high sSFR.

8. SUMMARY AND DISCUSSION

We have estimated AGN narrow-line luminosities and luminosity detection thresholds for the MaNGA sample, as well as the Eddington ratios and Eddington ratio detection thresholds. The detection thresholds in our sample are important—for most AGN non-detections these thresholds are similar to the luminosities found for the AGN detections. Clearly many AGN are hiding in the sample that cannot be detected, primarily due to contamination by star-formation-related line emission. By including the thresholds in our analysis, we estimate the actual AGN occurrence rates $F_{\rm AGN}$, meaning for this purpose the fraction of galaxies with AGN above some fixed luminosity or Eddington ratio. We find a fairly complicated dependence of $F_{\rm AGN}$ on stellar mass and sSFR, with the fraction of galaxies with Eddington ratios $\lambda > 10^{-3}$ ($F_{\rm AGN,-3}$) increasing with stellar mass for star-forming galaxies and decreasing with stellar mass for quiescent galaxies. At higher masses ($M > 10^{10.25} M_\odot$) the occurrence rate $F_{\rm AGN}$ increases with sSFR, whereas in our lowest mass bin $F_{\rm AGN}$ peaks for green valley galaxies (at low significance).

Over the next few years, DESI will release a public catalog of moderate resolution spectra of low redshift galaxies that dwarfs the SDSS Legacy sample and the MaNGA sample. However, this catalog will only be usable for statistical studies of the host galaxy demographics of narrow-line AGN if its selection effects are characterized. Trump et al. (2015) presented a method to determine selection thresholds. Here we have extended this method to more cleanly use non-detections of individual lines and have shown how to use the thresholds directly in Bayesian or maximum-likelihood fits to luminosity or Eddington ratio distributions. Although we have applied this method to the narrow-line AGN classification method of Ji & Yan (2020), it is applicable to the more traditional classification methods as well.

There nevertheless are still a number of flaws in our analysis that could be better tested or addressed. First, any analysis that depends on the bolometric luminosity, as our investigation of the Eddington ratio distribution does, depends on accurate bolometric corrections. Particularly in the realm of low mass galaxies (and thus intermediate mass black holes), it is possible that [O III]/Hβ depends on luminosity due to the expected change in hardness of the ionizing spectrum (e.g. Cann et al. 2019), which would require us to adjust our estimates of the selection effects. Similar concerns surround the dust corrections.

Second, our analysis has excluded the broad-line AGN, many of which have narrow-line emission as well. This effect would be minor for our analysis if the obscuration fraction were independent of stellar mass and sSFR; it would just mean we were missing a relatively small fraction of narrow line AGN. But we have not shown that this is the case.

Third, our detection threshold analysis assumes that all AGN would have the same set of strong line ratios if they were observed uncontaminated. This uniformity is certainly not the case, and we have not ruled out that our inferred trends are sensitive to this assumption. As a related point, our method does not model the distribution of non-AGN line ratios and relies on sharp detection criteria for AGN; therefore the results are still somewhat sensitive to these criteria (Section 5.3).

In short, although our analysis has advanced the state-of-the-art in accounting for selection criteria in studying AGN demographics, there are many new samples to investigate and many improvements to consider to our techniques. We encourage future investigators to apply and improve upon these methods and apply them to the massive new samples becoming available.

We thank Renbin Yan and Xihan Ji for useful discussions about narrow-line AGN selection methods, Julia Comerford for useful discussions about selecting AGN in MaNGA, and Shobita Satyapal for pointing us to the work on the potential dependence of [O III]/Hβ on mass. We thank Janelle Sy for her contributions to an early investigation of this data set.

Funding for the Sloan Digital Sky Survey IV has been provided by the Alfred P. Sloan Foundation, the U.S. Department of Energy Office of Science, and the Participating Institutions.

SDSS-IV acknowledges support and resources from the Center for High Performance Computing at the University of Utah. The SDSS website is www.sdss4.org.

SDSS-IV is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS Collaboration including the Brazilian Participation Group, the Carnegie Institution for Science, Carnegie Mellon University, Center for Astrophysics — Harvard & Smithsonian, the Chilean Participation Group, the French Participation Group, Instituto de Astrofísica de Canarias, The Johns Hopkins University, Kavli Institute for the Physics and Mathematics of the Universe (IPMU) / University of Tokyo, the Korean Participation Group, Lawrence Berkeley National Laboratory, Leibniz Institut für Astrophysik Potsdam (AIP), Max-Planck-Institut für Astronomie (MPIA Heidelberg), Max-Planck-Institut für Astrophysik (MPA Garching), Max-Planck-Institut für Extraterrestrische Physik (MPE), National Astronomical Observatories of China, New Mexico State University, New York University, University of Notre Dame, Observatário Nacional / MCTI, The Ohio State University, Pennsylvania State University, Shanghai Astronomical Observatory, United Kingdom Participation Group, Universidad Nacional Autónoma de México, University of Arizona, University of Colorado Boulder, University of Oxford, University of Portsmouth, University of Utah, University of Virginia, University of Washington, University of Wisconsin, Vanderbilt University, and Yale University.

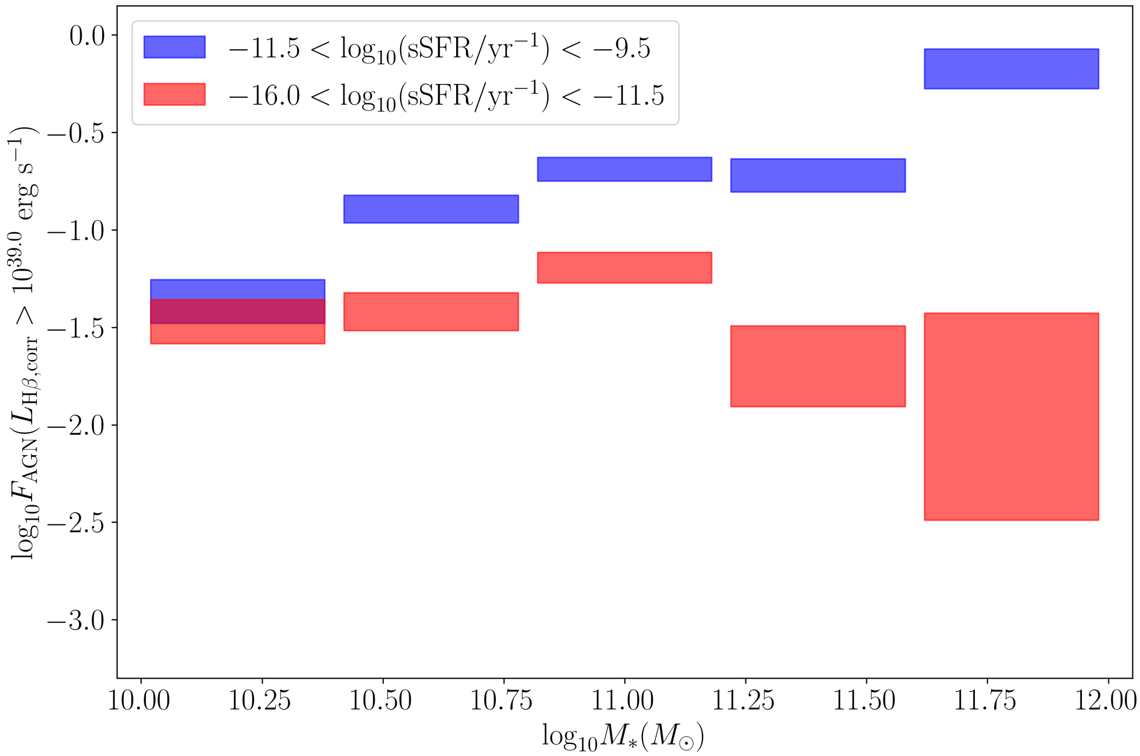

Figure 12. Similar to Figure 9, with a different division of samples according to sSFR.

APPENDIX

A. ANALYSIS OF QUIESCENT AND STAR-FORMING GALAXY HOSTS

We have performed the same analysis shown in Figures 9 and 11 for different choices of the sSFR bins. In particular, for the purposes of a companion effort comparing our results to numerical simulations of galaxy formation, we have divided the sSFRs into two bins, one star-forming and one quiescent, corresponding to $\log_{10}$ sSFR greater or less than -11.5. Figures 12 and 13 show F_{AGN} as a function of stellar mass for each category. These results show the same general trends as Figures 9 and 11.

REFERENCES

Aird, J., Coil, A. L., Moustakas, J., et al. 2012, ApJ, 746, 90, doi: 10.1088/0004-637X/746/1/90

Baldwin, J. A., Phillips, M. M., & Terlevich, R. 1981, PASP, 93, 5

Belfiore, F., Maiolino, R., Maraston, C., et al. 2016, MNRAS, 461, 3111, doi: 10.1093/mnras/stw1234

Belfiore, F., Westfall, K. B., Schaefer, A., et al. 2019, AJ, 158, 160, doi: 10.3847/1538-3881/ab3e4e

Blanton, M. R., Bershady, M. A., Abolfathi, B., et al. 2017, AJ, 154, 28, doi: 10.3847/1538-3881/aa7567

Bundy, K., Bershady, M. A., Law, D. R., et al. 2015, ApJ, 798, 7, doi: 10.1088/0004-637X/798/1/7

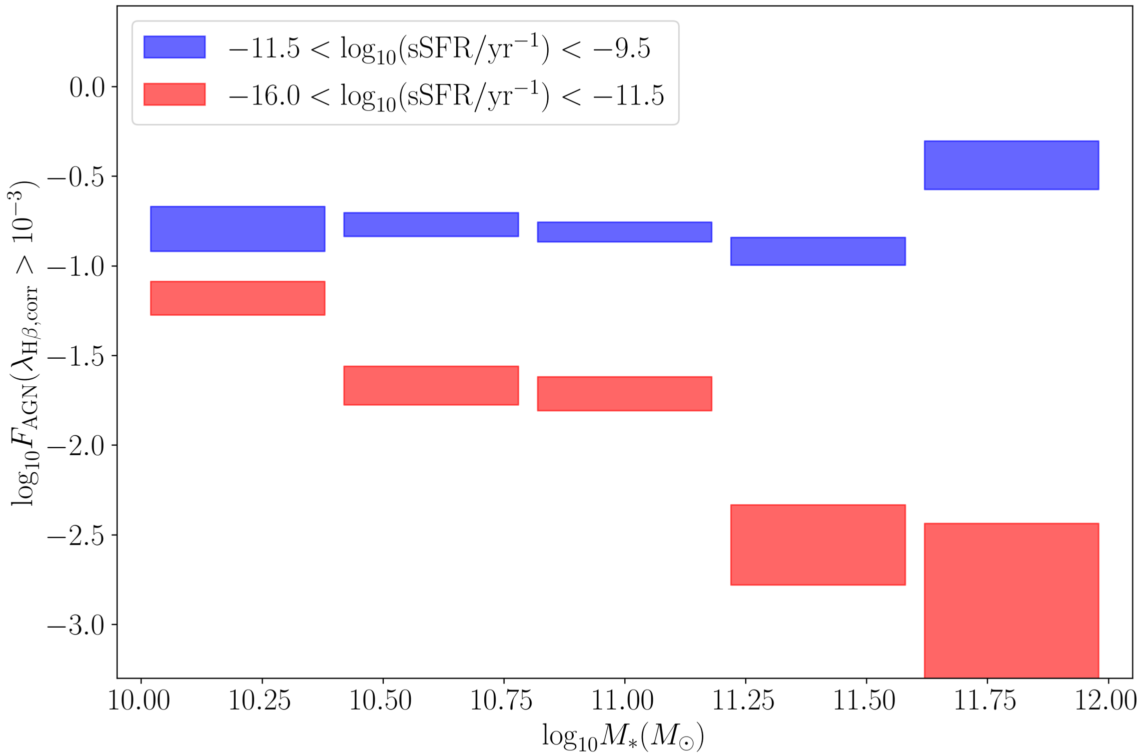

Figure 13. Similar to Figure 11, with a different division of samples according to sSFR.

Cann, J. M., Satyapal, S., Abel, N. P., et al. 2019, ApJL, 870, L2, doi: 10.3847/2041-8213/aaf88d

Clopper, C. J., & Pearson, E. S. 1934, Biometrika, 26, 404. http://www.jstor.org/stable/2331986

DESI Collaboration, Abareshi, B., Aguilar, J., et al. 2022, AJ, 164, 207, doi: 10.3847/1538-3881/ac882b

Diamond-Stanic, A. M., Rieke, G. H., & Rigby, J. R. 2009, ApJ, 698, 623, doi: 10.1088/0004-637X/698/1/623

Drory, N., MacDonald, N., Bershady, M. A., et al. 2015, AJ, 149, 77, doi: 10.1088/0004-6256/149/2/77

Ellison, S. L., Teimoorinia, H., Rosario, D. J., & Mendel, J. T. 2016, MNRAS, 458, L34, doi: 10.1093/mnrasl/slw012

Fabian, A. C. 2012, Annual Review of Astronomy and Astrophysics, 50, 455, doi: 10.1146/annurev-astro-081811-125521

Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306, doi: 10.1086/670067

Graham, A. W., & Scott, N. 2013, ApJ, 764, 151, doi: 10.1088/0004-637X/764/2/151

Greene, J. E., & Ho, L. C. 2006, ApJL, 641, L21, doi: 10.1086/500507

Greene, J. E., Setton, D., Bezanson, R., et al. 2020, ApJL, 899, L9, doi: 10.3847/2041-8213/aba534

Gültekin, K., Richstone, D. O., Gebhardt, K., et al. 2009, ApJ, 698, 198, doi: 10.1088/0004-637X/698/1/198

Heckman, T. M., Kauffmann, G., Brinchmann, J., et al. 2004, ApJ, 613, 109, doi: 10.1086/422872

Hickox, R. C., Mullaney, J. R., Alexander, D. M., et al. 2014, ApJ, 782, 9, doi: 10.1088/0004-637X/782/1/9

Ho, L. C. 2008, ARA&A, 46, 475, doi: 10.1146/annurev.astro.45.051806.110546

Ho, L. C., & Kim, M. 2009, ApJS, 184, 398, doi: 10.1088/0067-0049/184/2/398

Ji, X., & Yan, R. 2020, MNRAS, 499, 5749. https://arxiv.org/abs/2007.09159

Jones, M. L., Hickox, R. C., Black, C. S., et al. 2016, ApJ, 826, 12, doi: 10.3847/0004-637X/826/1/12
Kauffmann, G., & Heckman, T. M. 2009, MNRAS, 397, 135, doi: 10.1111/j.1365-2966.2009.14960.x
Kauffmann, G., et al. 2003a, MNRAS, 346, 1055, doi: 10.1111/j.1365-2966.2003.07154.x
—. 2003b, MNRAS, 341, 33
Khrykin, I. S., Hennawi, J. F., Worseck, G., & Davies, F. B. 2021, MNRAS, 505, 649, doi: 10.1093/mnras/stab1288
Kormendy, J., & Ho, L. C. 2013, ARA&A, 51, 511, doi: 10.1146/annurev-astro-082708-101811
LaMassa, S. M., Heckman, T. M., Ptak, A., et al. 2009, ApJ, 705, 568, doi: 10.1088/0004-637X/705/1/568
—. 2010, ApJ, 720, 786, doi: 10.1088/0004-637X/720/1/786
Law, D. R., Yan, R., Bershady, M. A., et al. 2015, AJ, 150, 19, doi: 10.1088/0004-6256/150/1/19
Law, D. R., Cherinka, B., Yan, R., et al. 2016, AJ, 152, 83, doi: 10.3847/0004-6256/152/4/83
Liu, D., Blanton, M. R., & Law, D. R. 2020, AJ, 159, 22, doi: 10.3847/1538-3881/ab5aeb
Man, Z.-y., Peng, Y.-j., Kong, X., et al. 2019, MNRAS, 488, 89, doi: 10.1093/mnras/stz1706
Marconi, A., Risaliti, G., Gilli, R., et al. 2004, MNRAS, 351, 169, doi: 10.1111/j.1365-2966.2004.07765.x
Martin, D. C., Wyder, T. K., Schiminovich, D., et al. 2007, ApJS, 173, 342, doi: 10.1086/516639
McDonald, M., McNamara, B. R., Calzadilla, M. S., et al. 2021, ApJ, 908, 85, doi: 10.3847/1538-4357/abd47f
Miller, J. S., & Mathews, W. G. 1972, ApJ, 172, 593, doi: 10.1086/151378
Netzer, H. 2019, MNRAS, 488, 5185, doi: 10.1093/mnras/stz2016
Park, D., Kelly, B. C., Woo, J.-H., & Treu, T. 2012, ApJS, 203, 6, doi: 10.1088/0067-0049/203/1/6
Planck Collaboration, Aghanim, N., Akrami, Y., et al. 2020, A&A, 641, A6, doi: 10.1051/0004-6361/201833910
Pucha, R., Juneau, S., Dey, A., et al. 2025, ApJ, 982, 10, doi: 10.3847/1538-4357/adb1dd
Salim, S., Rich, R. M., Charlot, S., et al. 2007, ApJS, 173, 267, doi: 10.1086/519218
Sánchez, S. F., Avila-Reese, V., Hernandez-Toledo, H., et al. 2018, RMAA, 54, 217
Sánchez, S. F., Barrera-Ballesteros, J. K., Lacerda, E., et al. 2022, ApJS, 262, 36. https://arxiv.org/abs/2206.07062
Sarzi, M., Shields, J. C., Schawinski, K., et al. 2010, MNRAS, 402, 2187, doi: 10.1111/j.1365-2966.2009.16039.x
Schawinski, K., Urry, C. M., Virani, S., et al. 2010, ApJ, 711, 284, doi: 10.1088/0004-637X/711/1/284
Shakura, N. I., & Sunyaev, R. A. 1973, A&A, 24, 337
Shankar, F., Weinberg, D. H., & Miralda-Escudé, J. 2009, ApJ, 690, 20, doi: 10.1088/0004-637X/690/1/20
Smee, S. A., Gunn, J. E., Uomoto, A., et al. 2013, The Astronomical Journal, 146, 32
Somerville, R. S., & Davé, R. 2015, ARA&A, 53, 51. https://arxiv.org/abs/1412.2712
Strauss, M. A., Weinberg, D. H., Lupton, R. H., et al. 2002, AJ, 124, 1810, doi: 10.1086/342343
Suresh, A., & Blanton, M. R. 2024, arXiv e-prints, arXiv:2404.04780, doi: 10.48550/arXiv.2404.04780
Trump, J. R., Sun, M., Zeimann, G. R., et al. 2015, ApJ, 811, 26. https://arxiv.org/abs/1501.02801
Veilleux, S., & Osterbrock, D. E. 1987, ApJS, 63, 295, doi: 10.1086/191166
Wake, D. A., Bundy, K., Diamond-Stanic, A. M., et al. 2017, AJ, 154, 86, doi: 10.3847/1538-3881/aa7ecc
Westfall, K. B., Cappellari, M., Bershady, M. A., et al. 2019, AJ, 158, 231, doi: 10.3847/1538-3881/ab44a2

Whitford, A. E. 1958, AJ, 63, 201, doi: 10.1086/107725
Woo, J.-H., Schulze, A., Park, D., et al. 2013, ApJ, 772, 49, doi: 10.1088/0004-637X/772/1/49
Woodrum, C., Williams, C. C., Rieke, M., et al. 2024, ApJ, 974, 305, doi: 10.3847/1538-4357/ad74f1
Yan, R., & Blanton, M. R. 2012, ApJ, 747, 61, doi: 10.1088/0004-637X/747/1/61
York, D. G., et al. 2000, AJ, 120, 1579, doi: 10.1086/301513
Yu, Q., & Tremaine, S. 2002, MNRAS, 335, 965, doi: 10.1046/j.1365-8711.2002.05532.x
Zakamska, N. L., Strauss, M. A., Krolik, J. H., et al. 2003, AJ, 126, 2125, doi: 10.1086/378610