

Bayesian AGN Decomposition Analysis for SDSS spectra: a correlation analysis of [O III] $\lambda 5007$ outflow kinematics with AGN and host galaxy properties

Remington O. Sexton¹,¹★ William Matzko,² Nicholas Darden,¹ Gabriela Canalizo¹ and Varoujan Gorjian³

¹Department of Physics and Astronomy, University of California, Riverside, 900 University Avenue, Riverside, CA 92521, USA

²Department of Physics and Astronomy, George Mason University, MS3F3, 4400 University Drive, Fairfax, VA 22030, USA

³Jet Propulsion Laboratory, M/S 169-327, 4800 Oak Grove Drive, Pasadena, CA 91109, USA

Accepted 2020 October 18. Received 2020 October 2; in original form 2020 July 13

ABSTRACT

We present Bayesian active galactic nucleus (AGN) Decomposition Analysis for Sloan Digital Sky Survey (SDSS) Spectra, an open source spectral analysis code designed for automatic detailed deconvolution of AGN and host galaxy spectra, implemented in PYTHON, and designed for the next generation of large-scale surveys. The code simultaneously fits all spectral components, including power-law continuum, stellar line-of-sight velocity distribution, Fe II emission, as well as forbidden (narrow), permitted (broad), and outflow emission line features, all performed using Markov chain Monte Carlo to obtain robust uncertainties and autocorrelation analysis to assess parameter convergence. Our code also utilizes multiprocessing for batch fitting large samples of spectra while efficiently managing memory and computation resources and is currently being used in a cluster environment to fit thousands of SDSS spectra. We use our code to perform a correlation analysis of 63 SDSS type 1 AGNs with evidence of strong non-gravitational outflow kinematics in the [O III] $\lambda 5007$ emission feature. We confirm findings from previous studies that show the core of the [O III] profile is a suitable surrogate for stellar velocity dispersion σ_* , however there is evidence that the core experiences broadening that scales with outflow velocity. We find sufficient evidence that σ_* , [O III] core dispersion, and the non-gravitational outflow dispersion of the [O III] profile form a plane whose fit results in a scatter of ~ 0.1 dex. Finally, we discuss the implications, caveats, and recommendations when using the [O III] dispersion as a surrogate for σ_* for the $M_{\text{BH}}-\sigma_*$ relation.

Key words: methods: data analysis – galaxies: active – quasars: absorption lines – quasars: emission lines.

1 INTRODUCTION

Data analysis codes, recipes, and software distributions for spectroscopic data analysis have become commonplace in the astronomy and astrophysics community, especially in the advent of large all-sky surveys, such as the Sloan Digital Sky Survey (SDSS; York et al. 2000) and the highly anticipated Large Synoptic Survey Telescope (Ivezić et al. 2019). Despite their widespread use, many of the computational methods used to perform these analyses are either (1) not shared by authors for various reasons or (2) not open source and cannot be accessed without the purchasing of proprietary software. In addition to this, many data analysis pipelines designed for large-scale surveys are written with the intent of fitting as many objects as possible in the shortest amount of time, while other analysis recipes may be suited for more-detailed analyses. Finally, many software packages are suited for fitting for specific types of objects, usually either galaxies or active galactic nuclei (AGNs), with no general means of fitting for both or other types of objects. As astronomy advances through the 21st century, sacrificing quality for speed will

no longer be necessary given the increasingly widespread use and availability of supercomputing resources in astronomy. Likewise, software designed for fitting specific astronomical objects will yield to more general fitting algorithms which can fit a diverse set of objects autonomously and in great detail.

Some notable existing software packages and codes have attempted to address the aforementioned issues. The Gas AND Absorption Line Fitting (GANDALF; Sarzi et al. 2006) code was one of the first large-scale algorithms to fully decompose gas emission from stellar absorption features, using penalized pixel-fitting (PPXF; Cappellari & Emsellem 2004; Cappellari 2017) to measure the stellar line-of-sight velocity distribution (LOSVD) with stellar templates. In the context of AGN studies, GANDALF was ill-suited for the complexities of fitting type 1 AGNs, which contain additional features such as broad lines, Fe II emission, power-law continuum, and possible ‘blue-wing’ components indicative of outflowing narrow-line gas. The Quasar Spectral Fitting package (QSFit; Calderone et al. 2017) allowed for fitting of type 1 AGNs with a variety of optional features, making it ideal for large-scale surveys of many thousands of objects. More recently, the release of PYQSOFIT (Guo, Shen & Wang 2018) includes many similar features of QSFit with the added functionality of Python. However, since QSFit and PYQSOFIT use a library of galaxy

* E-mail: rsext001@ucr.edu

templates to model the host galaxy component instead of attempting to model the LOSVD using stellar templates, AGNs with a strong stellar continuum component, such as type 2 AGNs, suffer from poor continuum modelling. Furthermore, while both GANDALF and QSFIT are technically open source, they are implemented using proprietary software and language (namely IDL). While certain licensed software may have once been prevalent in the astronomical community, there is a growing push towards open source software that can be easily shared, modified, and used among the research community. Among these open source languages is PYTHON, which is one of the fastest growing programming languages for data analysis today, and its widespread use makes it ideal for research in the astronomical community.

We address the limitations of current spectral fitting codes with a comprehensive fitting package implemented in Python and utilizing a Markov chain Monte Carlo (MCMC) fitting approach for accurate estimation of parameters and uncertainties, which we call *Bayesian AGN Decomposition Analysis for SDSS Spectra*. In its current version, our code is written for the SDSS spectra data model in the optical (specifically 3460–9463 Å, based on choice of stellar template library), however, because it is written in PYTHON and is open source,¹ it can be easily modified to accommodate other instruments, wavelength ranges, stellar libraries, templates, and has already been used successfully to perform decomposition on 22 type 1 AGNs observed with the Keck-I LRIS instrument (Sexton et al. 2019).

Our software attempts to address some notable and relevant problems with spectral fitting software available today. Because our code was designed for detailed decomposition of type 1 AGNs, which contain various components such as forbidden ‘narrow’ (typical $\text{FWHM} < 500 \text{ km s}^{-1}$) and permitted ‘broad’ (typical $\text{FWHM} > 500 \text{ km s}^{-1}$) emission lines, broad and narrow Fe II emission, AGN power-law continuum, ‘blue-wing’ outflow components, and the host galaxy stellar continuum, these components can be optionally turned on or off to fit less-complex objects such as type 2 AGNs or non-AGN host galaxies altogether, with all of these options easily configured through the Jupyter Notebook (Kluyver et al. 2016) interface. As a result, our code can be deployed for fitting a diverse range of astronomical objects and customized to the user’s needs. Additionally, the choice of PYTHON as the programming language of our code follows suit with a number of other software packages, such as ASTROPY (Astropy Collaboration 2013), which aim to replace antiquated software such as IRAF (Valdes 1984) or proprietary languages such as IDL, for astronomers now entering the field and/or adopting the PYTHON programming language for their analyses. If anything, the open source nature of our software will serve as a template for developing various implementations of the software for individual specific needs.

To our knowledge, our algorithm is the first of its kind to address a number of issues specific to the fitting of AGN spectra that other algorithms have yet to implement. First, our code was initially designed to fit all spectral components simultaneously, as opposed to masking regions of spectrum and fitting components separately. This is specifically advantageous for the decomposition of the stellar continuum and Fe II emission from other components for studies of AGN and host galaxy relations such as the $M_{\text{BH}} - \sigma_*$ relation. As noted in Sexton et al. (2019), stellar kinematics remain the single-most difficult quantity to measure in type 1 AGN, and obtaining reliable values and uncertainties for stellar quantities is a non-trivial

effort that includes a number caveats and systematics which can be difficult to account for (Greene & Ho 2006). Simultaneous fitting with Fe II emission templates also allows for detailed study of Fe II emission properties of type 1 AGN while taking into account the underlying stellar continuum. Finally, our code is the first software of its kind to use specific criteria for the automated detection and decomposition of outflow components in forbidden emission lines, which have recently become a topic of much study in the context of AGN and host galaxy evolution (see Section 3.1 and references therein).

The Bayesian MCMC approach used by our code for fitting spectral parameters is unique in that it provides an easily extensible framework for the user to modify the fitting model, free parameters, and convergence criteria. Many fitting software packages typically utilize a simpler least-squares minimization approach, however, it is recommended (almost universally) to perform Monte Carlo resampling of the data and re-fitting (also known as ‘bootstrapping’) to ensure accurate estimation of uncertainties. While the least-squares approach is typically faster, an MCMC approach allows the user to estimate robust uncertainties, visualize possible degeneracies, and assess how well individual parameters are constrained or if they have properly converged on a solution. While fitting algorithms that utilize random-sampling techniques admittedly suffer from slower runtimes, modern personal computers capable of multiprocessing to decrease runtimes are becoming commonplace. Since our code utilizes the affine invariant MCMC sampler EMCEE (Foreman-Mackey et al. 2013), multiprocessing is also an available option for fitting large samples of objects. The use of powerful Bayesian and computational techniques, open source framework, and diverse fitting options together make help achieve the ultimate goal of our code, which is to provide the most detailed and versatile fitting software for optical spectra in future sky surveys.

We describe the model construction, fitting procedure, and autocorrelation analysis used by our code to assess parameter convergence in Section 2. In Section 3, we discuss the significant correlations between stellar velocity dispersion σ_* , the decomposed [O III] core and outflow dispersions of the [O III] $\lambda 5007$ emission line found using our code.

Throughout this work, we assume a standard cosmology of $\Omega_m = 0.27$, $\Omega_\Lambda = 0.73$, and $H_0 = 71 \text{ km s}^{-1} \text{ Mpc}^{-1}$.

2 THE ALGORITHM

In the following subsections, we discuss the algorithm model construction, spectral components, fitting procedure, and autocorrelation analysis used to assess parameter convergence. We also discuss the results of benchmarking tests for the recovery of stellar velocity dispersion using our code.

2.1 Model construction

Our algorithm constructs a model for each spectrum under the assumption that all AGN components (i.e. the AGN power-law continuum, emission lines, possible outflows, and Fe II emission) reside atop a host galaxy component whose stellar contribution we wish to measure. A general overview of the model construction is as follows.

First, each spectral component is initialized using reasonable assumptions from the data. For example, the stellar continuum and power-law continuum are initialized at amplitudes that are each half of the total galaxy continuum level, which is estimated using the median flux of the fitting region. As another example, emission line

¹<https://github.com/remingtonsexton/BADASS3>

amplitudes are initialized at the maximum flux value within fixed wavelength regions centred at the expected rest-frame locations of the emission line. These initial parameter values need not be exact, as our algorithm will iteratively improve on their estimated values with each fitting iteration.

Models of each spectral component are then sequentially subtracted from the original data, and any remaining continuum is assumed to be the stellar continuum contribution, which is then fit with a predefined host galaxy template or empirical stellar templates to estimate the LOSVD. Once all parameters have been estimated and all model components have been constructed, their sum-total is used to assess the quality of the fit to the original data. This process is repeated for each iteration of the algorithm until a best fit is achieved.

We describe each of the spectral components used for constructing the model below.

2.1.1 AGN power-law continuum

In the simplest construction, the non-stellar thermal continuum in type 1 AGNs can be modelled as the sum of different temperature blackbodies at various radii within the AGN accretion disc (Malkan 1983). This manifests itself in the UV and optical as a ‘big blue bump’, which flattens out at longer wavelengths towards the near-IR, resembling a power-law continuum. We adopt the QSFIT simple power-law implementation from Calderone et al. (2017) given by

$$p(\lambda) = A \left(\frac{\lambda}{\lambda_b} \right)^{\alpha_\lambda}, \quad (1)$$

where A is the power-law amplitude, α_λ is the power-law index (or spectral slope), and λ_b that is a reference wavelength chosen to be the central wavelength value of the fitting region and determines the break in the power-law model. The power-law amplitude A and slope index α_λ are free parameters throughout the fitting process. The flat priors we set on these parameters dictate that A must be non-negative and no greater than the maximum flux density value of the data, and α_λ can vary in the range $[-4, 2]$. As in QSFIT, the reference break wavelength λ_b is fixed by default to be the centre wavelength value of the fitting region (i.e. $(\lambda_{\max} - \lambda_{\min})/2$), since the power-law slope is poorly constrained at optical wavelengths, however, this constraint can be relaxed if there is sufficient wavelength coverage in the near-UV. We show different values of the power-law slope in Fig. 1.

We find that the simple power-law model adequately describes the AGN continuum in the optical, especially if the object fitting region is limited to rest-frame $\lambda_{\text{rest}} > 3460 \text{ \AA}$, which is the lower limit of the wavelength range of the Indo-US Stellar Library (Valdes et al. 2004) used for fitting the stellar LOSVD. To better model the true shape of the power-law continuum, a large fitting region at $\lambda_{\text{rest}} < 3500 \text{ \AA}$ is necessary to better constrain the power-law index α_λ . For fitting regions $\lambda_{\text{rest}} > 3500 \text{ \AA}$, the AGN continuum can become highly degenerate with the host galaxy stellar continuum, especially if the shape of the power-law continuum is relatively flat. We nevertheless include a power-law continuum in our fitting model because its inclusion does not affect the overall fitting process.

2.1.2 Broad and narrow emission lines

All broad and narrow emission line features are, by default, modelled as a simple Gaussian function given by

$$g(\lambda) = A \exp \left[-\frac{1}{2} \left(\frac{(\lambda - v)^2}{\sigma^2} \right) \right], \quad (2)$$

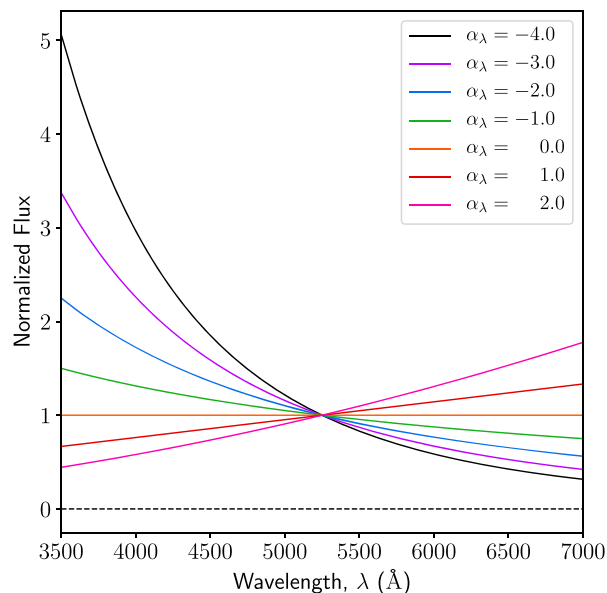


Figure 1. The AGN simple power-law model adopted from Calderone et al. (2017). Colours represent different values of α_λ in the range $[-4, 2]$ for the wavelength range $[3500, 7000]$. The reference wavelength for this wavelength range, $\lambda_b = 5250 \text{ \AA}$, is the locus of different models for α_λ and is held fixed by default to be the centre of the fitting range.

where A is the line amplitude, σ is the Gaussian dispersion, and v is the velocity offset of the Gaussian profile from the rest-frame wavelength of the line. Some types of objects, such as NLS1s, exhibit broad lines with extended wings (Moran, Halpern & Helfand 1996; Leighly 1999; Véron-Cetty, Véron & Gonçalves 2001; Berton et al. 2020), for which our algorithm can optionally model the emission line with a Lorentzian function given by

$$\ell(\lambda) = \frac{A\gamma^2}{\gamma^2 + (\lambda - v)^2}, \quad (3)$$

where $\gamma = \text{FWHM}/2$. A comparison of the two emission line models is shown in Fig. 2.

Because SDSS spectra are logarithmically rebinned, each pixel represents a constant velocity scale measured in km s^{-1} . This convenience allows us to initialize width and velocity parameters in units of km s^{-1} and fitting performed in units of pixels without the conversion from $\text{\AA}$, which is wavelength dependent. All narrow and broad line widths are corrected for the wavelength-dependent instrumental dispersion of the SDSS spectrograph during the fitting process so that final reported widths do not have to be corrected by the user. However, we still place a minimum value for all measured emission line widths to be the velocity scale of our spectra (in units of $\text{km s}^{-1} \text{ pixel}^{-1}$), to ensure that emission line widths are at least greater than a single pixel in width to avoid the fitting the noise spikes.

Given the modest resolution ($\sigma \sim 69 \text{ km s}^{-1}$) and signal to noise (S/N) of SDSS spectra, we find that in most cases a simple Gaussian function is sufficient to model the full shape emission lines. Other fitting algorithms attempt to model emission lines in higher detail, using Gauss–Hermite polynomials or additional higher order moments, in order to account of line asymmetries, which are especially obvious in broad line emission. Sexton et al. (2019) however showed that some line asymmetries can be attributed to strong absorption near Balmer features, and that emission line

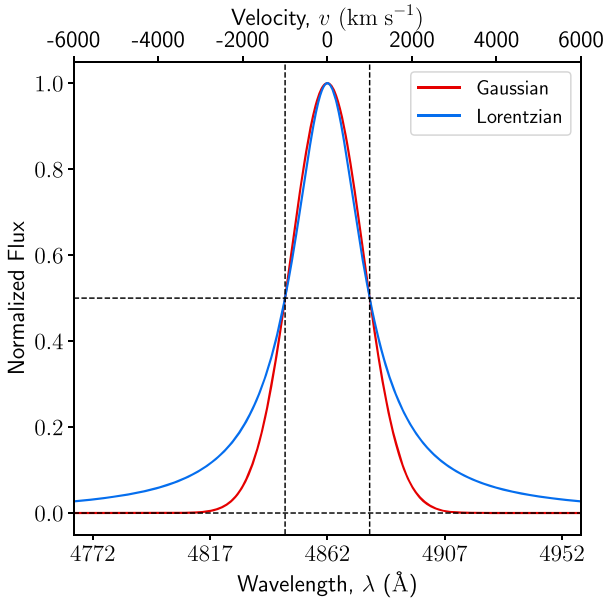


Figure 2. A comparison of the Gaussian and Lorentzian emission line models centred on the location rest-frame $H\beta$ with a FWHM of 2000 km s^{-1} . Dashed lines indicate the location and width of the FWHM.

asymmetries are generally resolved with simple Gaussian models as long as the underlying stellar population is modelled. Narrow lines that exhibit a ‘blue wing’ outflow component, as typically seen in the $[O III]$ emission lines, can be fit as an additional Gaussian component and is a standard feature of our code (see Section 2.2.2). Higher resolution spectra of nearby objects with strong emission lines can exhibit further complex non-Gaussian profiles even after outflow components are accounted for. These non-Gaussian profiles are best modelled iteratively using multiple Gaussian components until an optimal fit is achieved, and adding additional components can be easily achieved by modification of the code.

Many line fitting algorithms tie the widths of narrow lines to be the same across the entire fitting region, however, we leave this as an optional constraint in our code. The advantage of tying the widths only decreases the number of free parameters, while the disadvantage of tying all narrow line widths to each other can lead to a worse fit. Instead, our algorithm ties widths of lines that are nearest to each other in groups. For example, the narrow $[N II]/H\alpha/[S II]$ line group’s widths are tied, and the narrow $H\beta/[O III]$ line group’s widths are tied and fit separately. The $H\alpha/[N II]/[S II]$ line widths can be biased due to line blending, and/or the presence of outflows and broad lines, and since these lines tend to have larger fluxes than most other lines, they carry greater statistical weight in determining widths if they are tied to other lines in the spectrum. A similar argument can be made for the $H\beta/[O III]$ line group. Ideally, one would model each individual line separate from the rest, however, tying widths is still required for narrow forbidden lines obscured by broad permitted lines. Thus, tying widths of groups of lines both reduces fitting bias within each group while also reducing the number of free fitting parameters.

2.1.3 Fe II templates

To account for Fe II emission typically present in the spectra type 1 AGNs, our algorithm uses the broad and narrow Fe II templates from Véron-Cetty, Joly & Véron (2004), which are optimal for subtraction since they include emission features that are commonly found in

many Seyfert 1 galaxies, as opposed to a single template based solely on I Zw 1 (Barth et al. 2013). Fig. 3 shows the narrow the broad Fe II emission features of the Véron-Cetty et al. (2004) template.

All Fe II lines from Véron-Cetty et al. (2004) are modelled as Gaussian functions using equation (2) and are summed together into two separate broad and narrow templates, each of which can be scaled by a multiplicative free-parameter amplitude A during the fit. Following QSFIT, the default full width at half-maximum (FWHM) of broad and narrow Fe II lines are fixed at 3000 and 500 km s^{-1} , respectively, which are adequate given the resolution and typical S/N of SDSS data. The velocity offset of each line is also fixed by default. We justify holding the FWHM and velocity offsets fixed due to the fact that broad and narrow emission are blended together and superimposed atop one another, usually at varying amplitudes. This leads to a strong degeneracy in both the FWHM and velocity offsets in these features. However, the FWHM and velocity offset constraints can be optionally turned off or adjusted to particular values for each the broad and narrow templates for more-detailed fitting.

In addition to the template from Véron-Cetty et al. (2004), our code can alternatively use the temperature-dependent Fe II model from Kovačević, Popović & Dimitrijević (2010), shown in Fig. 4, which independently models each of the F , S , and G atomic transitions of Fe II, as well as some strong lines from I Zw 1, in the region between 4400 and 5500 Å , with amplitude, FWHM, velocity offset, and temperature as free parameters. The template from Kovačević et al. (2010), while slightly smaller in wavelength coverage compared to the Véron-Cetty et al. (2004) template, can more accurately model objects with particularly strong Fe II such as NLS1s (Véron-Cetty et al. 2001; Xu et al. 2012; Rakshit et al. 2017). Individual transition amplitudes, widths, and velocity offsets can be optionally fixed during the fitting process as well.

2.1.4 Host galaxy and stellar absorption features

The original purpose of our code was to extract the host galaxy contribution in type 1 AGN spectra, and in particular, estimate the stellar LOSVD to obtain stellar velocity v_* and stellar velocity dispersion σ_* . In this regard, our code serves as a wrapper for the stellar-template fitting code PPXF (Cappellari & Emsellem 2004; Cappellari 2017), allowing the user to optionally fit the underlying stellar population to extract stellar kinematics. After subtracting off all the aforementioned components from the original data, our algorithm models the stellar population using 50 empirical stellar templates from the Indo-US Library of Coudé Feed Stellar Spectra (Valdes et al. 2004). The Indo-US Library was chosen for its high resolution (FWHM resolution of 1.35 Å ; Beifiori et al. 2011) as well as its wide wavelength coverage from 3460 to 9464 Å . The 50 chosen templates include the full range of spectral types from O to M, and were specifically chosen for minimal gaps in coverage. We find that using 40–50 stellar templates strikes an optimal balance between reducing the chances of template mismatch and large computation times, since the non-negative least-squares routine used by PPXF for choosing templates and calculating weights carries the largest computational overhead for our algorithm and scales with the number of templates used in the fit.

By default, only v_* and σ_* are fit for SDSS spectra, however, if given higher resolution spectra, it is still possible for PPXF to estimate the higher order Gauss–Hermite moments of the LOSVD. One caveat to fitting the LOSVD via stellar template fitting is the limited wavelength range of the stellar template library chosen for fitting the LOSVD. In this regard, the Indo-US Library is provides the largest optical range and highest resolution for currently

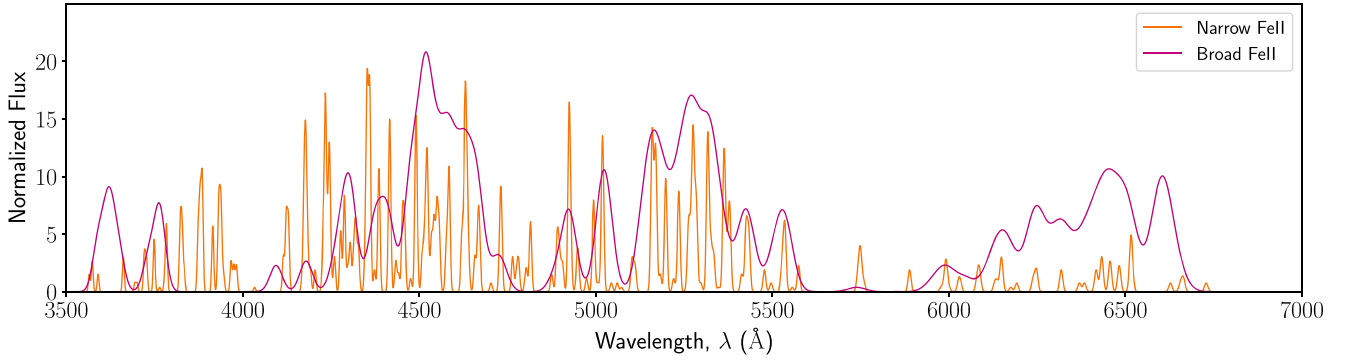


Figure 3. Broad and narrow Fe II templates from Véron-Cetty et al. (2004), which (by default) have fixed zero velocity offset and fixed width in our code. Broad Fe II is initialized with an FWHM of 3000 km s^{-1} , and narrow Fe II with an FWHM of 500 km s^{-1} , and held constant throughout the fitting process, following the implementation of QSFit (Calderone et al. 2017).

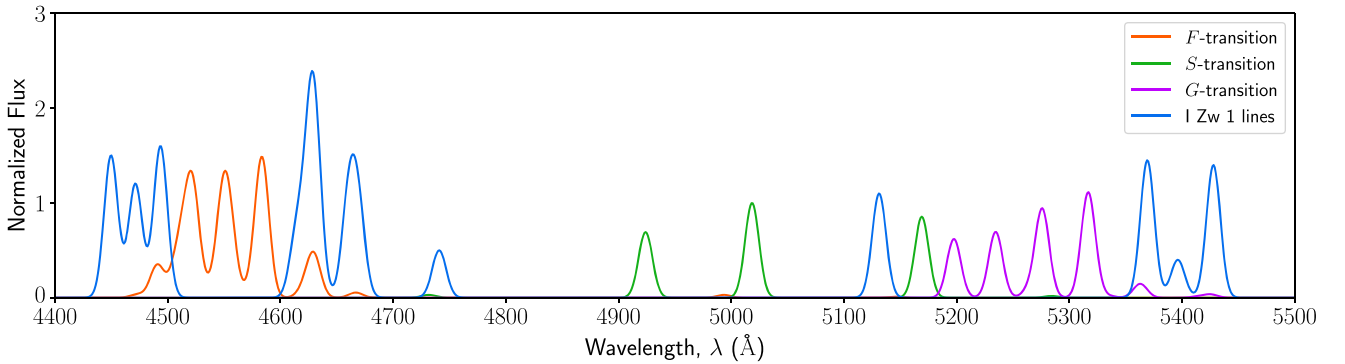


Figure 4. The temperature-dependent Fe II template from Kovačević et al. (2010) for a more-detailed analysis of optical Fe II emission. Our algorithm independently models each of the *F*, *S*, and *G* atomic transitions of Fe II, as well as some strong lines from I Zw 1, in the region between 4400 and 5500 Å , with amplitude, FWHM, velocity offset, and temperature as a free parameters.

available empirical stellar libraries. However, if one chooses to use a different library of stellar templates, the LOSVD fitting range can be extended.

In general, the quality of the fit to the LOSVD is S/N dependent. We find that if the continuum S/N < 10 , estimates of v_* and σ_* can have uncertainties > 50 per cent therefore we allow the user to optionally disable fitting of the LOSVD and instead fit the stellar continuum with a single stellar population (SSP) model generated using the MILES Tune Stellar Libraries Webtool (Vazdekis et al. 2010), which is initialized with a metallicity $[M/H] = 0.0$, age of 10.0 Gyr , and dispersion of 100 km s^{-1} to match the depth of stellar absorption features typically seen in SDSS galaxies. The SSP template is normalized at 5500 Å and is scaled by a multiplicative factor that is a free-parameter during the fitting process. We also include alternative MILES SSP models with ages ranging from 0.1 to 14.0 Gyr which can be used as optional substitutes. We show a range of MILES SSP models which can be used by our code in Fig. 5.

Due to the limited range of the MILES stellar library, if the fitting range is outside ($3525 \text{ Å} \leq \lambda \leq 7500 \text{ Å}$), SSP models from Maraston et al. (2009) are used instead, which have a coverage of ($1150 \text{ Å} \leq \lambda \leq 25000 \text{ Å}$), however have much larger dispersions that do not match strong stellar absorption features in SDSS spectra.

We note that dispersions measured with PPXF already take into account the instrumental dispersion of the SDSS, since it first convolves input templates to the resolution of the SDSS before the fitting process. We none the less place a lower limit on the allowed

values for σ_* to be the velocity scale (in units of $\text{km s}^{-1} \text{ pix}^{-1}$) of the input spectra.

2.2 Fitting procedure

2.2.1 Determination of initial parameter values

As is true for all fitting algorithms, the number of required MCMC iterations required for convergence on a solution is sensitive to the initial parameter values. Ideally, one should initialize parameters as close as possible to their actual posterior values in order to minimize the number of iterations used for searching parameter spaces and maximize the number of posterior sampling iterations. To do this, we employ maximum-likelihood estimation of all parameters using the SCIPY (Virtanen et al. 2020) `scipy.optimize.minimize` function to find the *negative* maximum (minimum) of the log-likelihood function, which we derive as follows.

We assume that each datum of the spectrum can be approximated as a normally distributed random variable of mean $y_{\text{data},i}$ and standard deviation σ_i . The likelihood of the data given the model y_{model} is given by

$$L = \prod_{i=1}^N \frac{1}{(2\pi\sigma_i^2)^{1/2}} \exp \left[-\frac{(y_{\text{data},i} - y_{\text{model},i})^2}{\sigma_i^2} \right] \quad (4)$$

Since equation (4) can result in very large values, which often times exceeds the numerical precision of most computing machines, it is

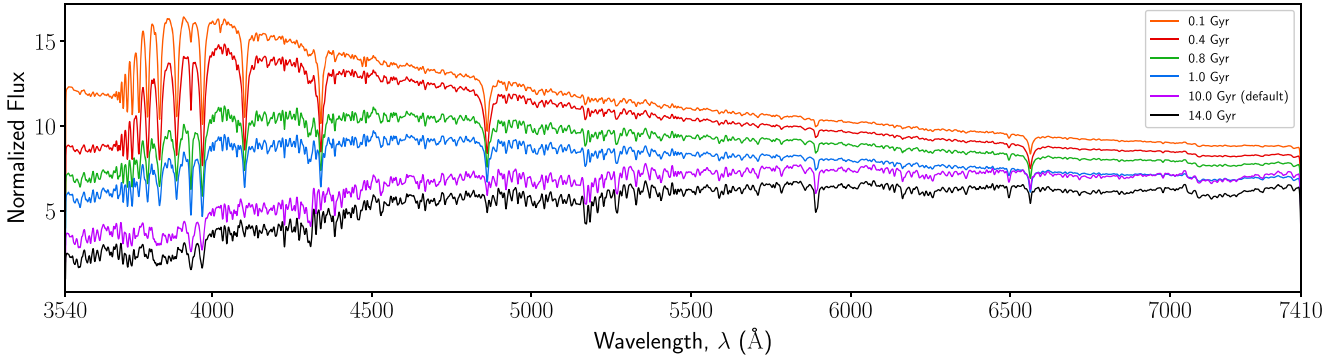


Figure 5. Single stellar population models from Vazdekis et al. (2010) generated using the MILES Tune Stellar Libraries webtool used for fitting the host galaxy contribution if the LOSVD (stellar template) fitting is not performed.

easier to use the natural log of the likelihood, i.e. the *log likelihood*:

$$\mathcal{L} = \log(L) = -\frac{1}{2} \sum_{i=1}^N \log(2\pi\sigma_i^2) + \frac{(y_{\text{data},i} - y_{\text{model},i})^2}{\sigma_i^2}. \quad (5)$$

The constant terms in this sum that do not change from one iteration to the next, including $\log(2\pi\sigma_i^2)$, can be dropped, as they do not play a role in determination of the minimum. The log likelihood is therefore given by

$$\mathcal{L} = \sum_{i=1}^N \frac{(y_{\text{data},i} - y_{\text{model},i})^2}{\sigma_i^2}, \quad (6)$$

where the sum is performed over each spectral channel i for each datum $y_{\text{data},i}$, $y_{\text{model},i}$ is the value of the model at each i , and σ_i is the 1σ uncertainty at each i determined from the SDSS inverse variance of the spectrum.

The *scipy.optimize.minimize* function, which employs the built-in Sequential Least Squares Programming (SLSQP; Kraft 1988) method, is used to include bounds and constraints on all parameters. Parameter bounds, which are the minimum and maximum values for each parameter, are determined from the data. For example, the emission line amplitudes must be non-negative, and can have a maximum value of the data in the fitting region. These simple boundary conditions, which are later used by EMCEE, are effective in limiting the parameter space for timely convergence. Constraints on parameters are used for the testing for possible blueshifted wing components in [O III] (see Section 2.2.2).

By default, our algorithm only performs one maximum-likelihood fit to obtain initial parameter values in the interest of time and a single fit is sufficient for initializing parameters for MCMC fitting. However, if one chooses not to perform MCMC fitting, or if one desires more robust initial parameter values, our code can optionally perform multiple iterations of maximum-likelihood fitting by resampling the spectra with random normally distributed noise from the spectral variance and re-fitting the spectrum, i.e. Monte Carlo ‘bootstrapping’.

2.2.2 Testing for presence of outflows in narrow emission lines

Additional ‘blue wing’ components in narrow emission lines, indicative of possible outflowing gas from the central BH, are known to be commonplace in AGN-host galaxies (Nelson & Whittle 1996; Mullaney et al. 2013; Zakamska & Greene 2014; Woo et al. 2016; DiPompeo et al. 2018; Rakshit & Woo 2018; Davies et al. 2020). If present, failure to account for the blue excess in narrow

line emission can lead to significant difference in measured line quantities, especially if one uses narrow line width as a proxy for σ_* (Woo et al. 2006; Komossa & Xu 2007; Bennert et al. 2018; Sexton et al. 2019).

Blue wings are most visibly obvious in the narrow [O III] emission line in type 1 AGNs because [O III] is not significantly contaminated by nearby broad lines. To determine if blue wings are present, our code can optionally perform preliminary single-Gaussian and double-Gaussian fits to the $H\beta$ /[O III] or $H\alpha$ /[N II]/[S II] narrow line complexes to test if an additional Gaussian component in the model is justified. The test for outflows is identical to the process used for fitting initial parameter values using maximum-likelihood estimation (i.e. Monte Carlo bootstrapping). The double-Gaussian fit makes the assumption that outflows are present by including a narrower ‘core’ and a broader ‘outflow’ component for the narrow emission lines. Monte Carlo bootstrapping for a user-defined set of iterations is then used to obtain uncertainties on core and outflow components to assess the quality of the fit. During the fitting process the FWHM of the outflow component is constrained to be greater than the core component, and the amplitude of the outflow component is constrained to be less than the core component, following what is typically seen in the literature. However, since outflows are not necessarily always blueshifted, but sometimes at equal or redshifted velocities with respect to the core component (albeit in rare occurrences), we do not constrain the velocity offset of either component during the fitting process. This feature is useful for detecting outflow components found in star-forming galaxies, which are known to be less offset from the core component causing a more symmetric [O III] line profile (Cicone, Maiolino & Marconi 2016; Davies et al. 2019; Manzano-King, Canalizo & Sales 2019).

To quantify the presence of outflows in [O III], we visually identify 63 objects from a sample of 173 known type 1 AGN that both (1) exhibit the characteristic line profile asymmetry commonly seen in the literature, and (2) have a measurable non-gravitational component in [O III] relative to the systemic (stellar) velocity dispersion (this is discussed in detail in Section 3.2) and derive empirical relationships between measurable parameters that recover these objects. After our algorithm performs fits for both the single-Gaussian (no-outflow) and double-Gaussian (outflow) models, the following empirical diagnostics are used to determine if a secondary outflow component is justified in the model:

$$\text{Amplitude metric: } \frac{A_{\text{outflow}}}{(\sigma_{\text{noise}}^2 + \delta A_{\text{outflow}}^2)^{1/2}} > 3.0, \quad (7)$$

$$\text{Width metric: } \frac{\sigma_{\text{outflow}} - \sigma_{\text{core}}}{(\delta\sigma_{\text{outflow}}^2 + \delta\sigma_{\text{core}}^2)^{1/2}} > 1.0, \quad (8)$$

$$\text{Velocity metric: } \frac{v_{\text{core}} - v_{\text{outflow}}}{(\delta v_{\text{core}}^2 + \delta v_{\text{outflow}}^2)^{1/2}} > 1.0, \quad (9)$$

$$\text{Statistic: } \frac{\left(\frac{\text{RSS}_{\text{no outflow}} - \text{RSS}_{\text{outflow}}}{k_2 - k_1} \right)}{\left(\frac{\text{RSS}_{\text{outflow}}}{N - k_2} \right)}, \quad (10)$$

where A is the line amplitude (in units of $10^{-17} \text{ erg cm}^{-2} \text{ s}^{-1} \text{ \AA}^{-1}$), σ is the Gaussian dispersion (FWHM/2.355; in units of km s^{-1}), and v is the velocity offset of the line relative to the rest frame of the overall spectra (in units of km s^{-1}). The quantity RSS is the sum-of-squares of the residuals within $\pm 3\sigma$ of the full (core + outflow) [O III] line profile, $k_1 = 3$ is the number of degrees of freedom in the single-Gaussian model, $k_2 = 6$ is the number of degrees of freedom in the double-Gaussian model, and N is the size of the sample used to calculate RSS. If parameters of the core or outflow models do not adhere to their bounds or approach to the limits of their constraints, which in turn violates the number of degrees of freedom for each model, our code flags the relevant parameters and defaults to a single-Gaussian (no-outflow) model.

Equation (7) is a measure of the amplitude of outflow component above the noise, while equations (8) and (9) are a measure of how much we can significantly detect measurable differences between the core and outflow FWHM and velocity offsets, respectively. Another way of interpreting equations (7)–(9) is the uncertainty overlap between parameter values, which signifies how well our algorithm can separate core and outflow components in parameter space. For example, a value of 2 for the width metric indicates that there is 2σ separation between the best-fitting values of the core and outflow FWHM.

Equation (10) is statistical F -test for model comparison between the single- and double-Gaussian models. The F -statistic in this context calculates unexplained variance between the outflow and no-outflow models as a fraction of the unexplained variance in the outflow model alone. The F -statistic is then used to calculate a p -value, which if less than a critical value (by default, $\alpha = 0.05$), indicates that we can reject the null hypothesis that there is no significant difference between the single- and double-Gaussian models, and that the difference is greater than that which could be attributed to random chance. We express our confidence in the outflow model by calculating $1 - \alpha$. For example, $\alpha = 0.05$ indicates a 95 per cent confidence that a double-Gaussian model explains the variance in fitting the [O III] profile significantly better than a single-Gaussian model.

All of these criteria can be toggled on or off, and the significance and confidence thresholds for each can be changed to meet the user's specific needs. We find that the above criteria provides a satisfactory method in finding objects with strong blueshifted excess in [O III] with a success rate of >90 per cent compared to visual identification, and recommend this method to determine if strong outflows are present or if the [O III] core dispersion needed as a surrogate for σ_* when it cannot be otherwise measured. We note that while these empirical criteria are successful in describing the types of objects in our sample, more sophisticated statistical modelling and cross-validation techniques with a larger sample will be required to improve the capability of our algorithm to identify objects with outflows.

If the above criteria are met, the final set of parameters that are fit with EMCEE will include a second Gaussian component for all narrow lines in the fitting region, otherwise, the final model will only fit a single Gaussian component to each narrow line. It is important

to note that any emission line, whether it exhibits blue excess or not, can be fit with more than one component and produce a better fit. However, if one is using the core component of [O III] as a surrogate for stellar velocity dispersion, using a two-component fit when only one component is justified by the data can significantly underestimate the stellar velocity dispersion. Likewise, not correcting for a strong and clearly visible outflow component can significantly overestimate the stellar velocity dispersion if the core component is used as a surrogate.

While our code can perform tests for outflows on either the H β or H α region, it is recommended that if one wants to fit outflows in both the H β and H α regions that it be done simultaneously, in which case our algorithm uses the [O III] $\lambda 5007$ outflow component to constrain the properties of the outflow components for H α , [N II], and [S II]. This constraint is activated because even if a broad line is not present in H α , narrow H α and [N II] can still be severely blended due to the resolution of the SDSS. However, if one chooses to fit outflows in H α independently from H β , it is still possible at the user's discretion.

As an aside, there is no single definition or quantification of what constitutes an 'outflow' in regards to emission lines. Our algorithm, admittedly, can only detect significantly broad and offset [O III] wings that are commonly seen in the literature, but outflows can generally produce a wide range of emission line profiles. For example, Bae & Woo (2016) modelled biconical outflows in 3D and found that the emission line profile of outflows is strongly dependent on orientation and dust extinction, resulting in sometimes redshifted and non-Gaussian profiles. Some models also indicate that a high-velocity outflow can be present without exhibiting a broad blue excess due simply to bicone orientation.

The criteria we present here constitute a *preliminary* means of filtering out objects with strong outflows which may cause significant residuals if not taken into account, and can make no claims as to whether an outflow is present if the secondary outflow component is close to the velocity offset or width of the primary core component of the emission line. The best method of determining if a secondary component is necessary is to perform a full fit with the double-Gaussian model with EMCEE, as well as fitting the LOSVD to estimate the gravitational influence of the narrow-line region (NLR), and examining the individual parameter chains to ensure that they are not degenerate and have converged on a stable solution.

2.2.3 Final parameter fitting

Final parameter fitting performed using MCMC begins by initializing each parameter at its maximum-likelihood value obtained from the initial fit. We use equation (6) as the likelihood probability and initialize each parameter with a flat prior with lower and upper bounds determined by the data. If the model contains outflow components, constraints on outflow parameters (see Section 2.2.2) are included in their respective flat priors. We place an additional constraint on broad line components, whose widths must be greater than narrow line widths, if broad lines are included in the fit. Fitting is then performed iteratively via MCMC until each component's parameters have converged. Each parameter space is randomly sampled using the affine invariant MCMC sampler EMCEE until a user-defined number of iterations is reached or if autocorrelation analysis (recommended) has determined that parameter convergence has been sufficiently reached (see Section 2.2.4). Fitting using the EMCEE package is advantageous since the use of multiple simultaneous 'walkers' efficiently explores each parameter space in parallel, all of which form MCMC parameter 'chains' from which the final

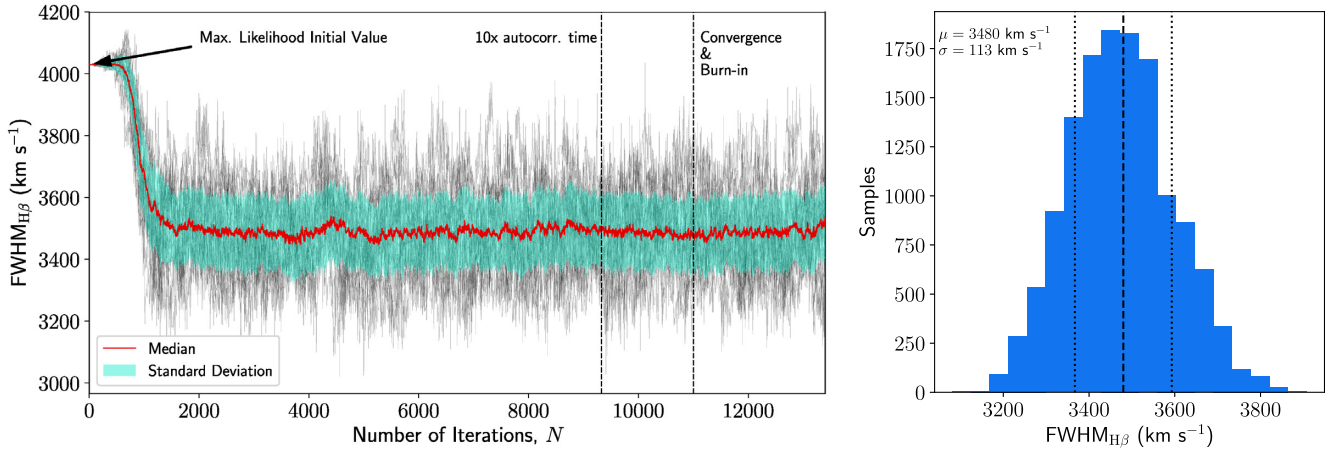


Figure 6. *Left:* The FWHM parameter MCMC chain for the $H\beta$ FWHM from the above example performed using 100 walkers. The initial starting position, estimated using maximum-likelihood fitting, overestimates the final width of the line by more than 500 km s^{-1} . As other parameters are fit, the value $H\beta$ FWHM decreases and settles into a stable solution by ~ 2000 iterations. Convergence is reached at $\sim 11\,000$ iterations (for this example, when 10 times the autocorrelation time per parameter at 10 per cent tolerance per parameter is achieved), and the burn-in is chosen to be the final 2500 iterations after all other parameters has been achieved. *Right:* A histogram of the last 2500 iterations of all 100 walkers.

posterior distribution is estimated for each parameter, as shown in Fig. 6.

The values of parameters estimated using the initial maximum-likelihood routine (*scipy.optimize.minimize*) can differ substantially from parameters estimated using EMCEE, as shown in Fig. 6. This large difference is attributed to the stellar continuum model used in each fit. The initial fit uses only a single SSP template to estimate the contribution to the stellar continuum. The *scipy.optimize.minimize* algorithm, while relatively fast, is not sensitive enough to fit the LOSVD with stellar templates in addition to many other component parameters. Fitting the LOSVD requires many iterations due to the fact that even moderate changes in stellar velocity or stellar velocity dispersion need not drastically change the shape of the resulting stellar continuum model, or drastically change the value of the calculated likelihood, which is partly due to the number and diversity of template stars used by our code. Instead, the *scipy.optimize.minimize* algorithm prioritizes the fitting of components that have the greatest effect on the calculated likelihood, such as emission lines and continuum amplitudes. In short, the limitations of the *scipy.optimize.minimize* routine are due to a combination of (1) degeneracies inherent to the stellar template fitting process and (2) the likelihood threshold requirements needed for *scipy.optimize.minimize* to achieve a solution. As a result, the initial fitting routine in our algorithm only fits a single SSP template for the stellar continuum.

The advantage of MCMC fitting with EMCEE allows for prolonged simultaneous fitting of parameters even after a likelihood threshold has been achieved. Components that are fit very easily (such as emission lines and continuum amplitudes) converge on solutions very quickly, while components that are less sensitive to even moderate changes (such as stellar templates) can continue to converge on a solution, even if the calculated maximum likelihood does not vary considerably. The advantage of MCMC sampling in this context allows our algorithm to explore the LOSVD parameter space even after a maximum likelihood has been reached, which allows for greater variation in template stars used for achieving a best fit. In other words, parameter convergence is more heavily dependent on parameter variation (autocorrelation; see Section 2.2.4) rather than achieving some maximum-likelihood threshold. In the case of

the $\text{FWHM}_{H\beta}$ parameter chain shown in Fig. 6, the initial value of $\text{FWHM}_{H\beta}$ is determined by a SSP of fixed LOSVD and fit with *scipy.optimize.minimize* algorithm. The MCMC fit performed by EMCEE then randomly samples the LOSVD around the initial likelihood values, which in turn adjusts the number and weights of stellar templates, varies the stellar $H\beta$ absorption, and adjusts the $\text{FWHM}_{H\beta}$ to obtain a better fit. Because MCMC randomly samples about the maximum-likelihood value, the algorithm need not stop until the LOSVD parameters have settled on a stable solution.

The MCMC fitting process uses the same likelihood function, priors, and constraints used in the initial fitting procedure, but instead produces parameter distributions from which best-fitting values and uncertainties can be reliably estimated. The best-fitting values are then used to construct a final model, an example of which is shown in Fig. 7.

2.2.4 Autocorrelation analysis

Over a number of iterations, the values of a parameter can fluctuate significantly as each parameter space is explored. Since some parameters can be strongly correlated (for instance, the AGN power-law continuum component and the stellar continuum component amplitudes), there can exist parameter degeneracies that are not taken into account when using conventional fitting techniques. Parameter fluctuations and degeneracies can also vary as a function of S/N. The nature and diversity of galaxy and AGN data makes determination of parameter convergence a non-trivial issue, especially if the goal is to fit a large number of objects that vary intrinsically or in data quality. Fitting algorithms typically address this by setting the number of fitting iterations to an arbitrarily high number or setting a minimum tolerance in the change in the likelihood value. While these methods of convergence are generally good, they only guarantee convergence in the overall fit to the data, and not on the convergence of individual parameters.

To address this, our code employs autocorrelation analysis to assess parameter convergence. Autocorrelation analysis functions are built into EMCEE (see Foreman-Mackey et al. 2013), however, we tailor these functions for the purposes of spectral fitting. The

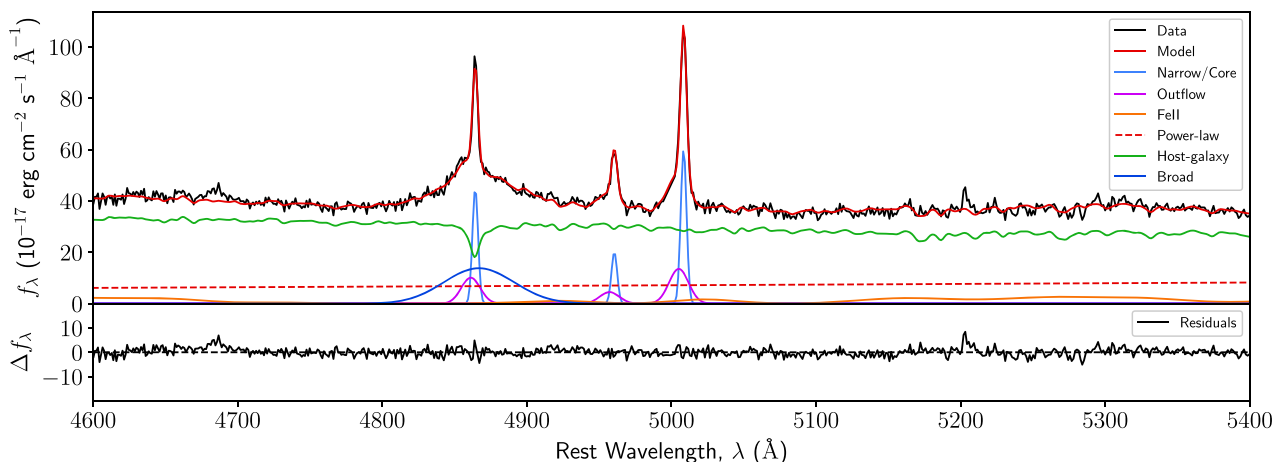


Figure 7. An example of the best-fitting final model output of a spectrum fit in the $H\beta/[O\text{ III}]$ region with individual fitting components designated by different colours. This particular object exhibits the blueshifted $[O\text{ III}]$ outflow components (magenta) which our algorithm is designed to detect and fit.

integrated autocorrelation time, which is the number of iterations required for a parameter chain to produce an independent sample, is calculated for all parameters at incremental fitting iterations (by default, every 100 fitting iterations), and then the convergence criteria are checked to see if they are satisfied by the most current fit parameters. There are two criteria that must be satisfied that define convergence: (1) if calculated autocorrelation times, which are multiplied by some multiplicative factor (by default, 10.0), exceeds the number of performed fitting iterations, and (2) if the difference between the current and previous calculated autocorrelation times is less than a specified percent change (by default, 10 per cent).

In practice, some parameters never reach adequate convergence, usually due to strong degeneracies with other model components (such as Fe II). To accommodate these instances, our code offers the user four modes of convergence in terms of the integrated autocorrelation time: (1) mean, (2) median, (3) user-specified parameters, and (4) all parameters. The first two options calculate the mean or median autocorrelation time of *all* free parameters to determine when an overall solution is reached, however it does not guarantee that all parameters reach convergence. To guarantee the convergence of specific parameters of interest, option (3) allows the user to indicate which specific parameters are considered for autocorrelation analysis, which is useful for ignoring components which have high autocorrelation times (poorly constrained or highly degenerate components) or parameters of low importance. Finally, option (4) allows the user to specify that all parameters must converge on a solution, for which our code runs until all parameters satisfy the autocorrelation conditions or the algorithm reaches the user-defined maximum number of iterations.

Fig. 8 shows an example of different autocorrelation modes and the required number of iterations for a 17-parameter model. We recommend that the user either select the specific parameters of interest for convergence or simply choose the ‘mean’ criteria. The ‘median’ criteria is less sensitive to outlier parameters (parameters with large autocorrelation times) and will generally perform fewer iterations for convergence. The ‘all’ convergence mode is the most strict type of convergence, however, there is no guarantee that convergence can be reached for all parameters within the set maximum number of iterations, especially if the data has low S/N or there are degenerate parameters.

Once the convergence criteria are met, our algorithm continues to fit for a set number of iterations (by default, 2500), which is ultimately used for the posterior distribution to determine the best-fitting parameter values and uncertainties. The iteration at which convergence is achieved defines the ‘burn-in’ for the parameter chains, after which all following iterations contribute to the final parameter estimation (i.e. the iterations that contribute to the histogram in Fig. 6). Note that this is only true if autocorrelation analysis is used to assess convergence (`auto_stop=True`), otherwise, the algorithm runs for the maximum number of iterations using the burn-in defined by the user. If for any reason convergence criteria are met and then subsequently violated, our algorithm resets the burn-in and continues to sample until convergence is met again. This ensures that convergence is maintained at all times after the burn-in and that a best fit is not achieved prematurely.

2.3 Performance tests

Since one of the primary goals of our code is the recovery of the LOSVD, and specifically the stellar velocity dispersion σ_* in AGN host galaxies, we perform a series of performance tests as a function of different components and parameters which may affect measurements of σ_* in the optical Mg II/Fe II region from 4400 Å to 5500 Å. We note that the results of these tests are not strictly limited to our algorithm, but also general fitting techniques concerned with fitting σ_* in AGN host galaxies.

2.3.1 Recovery of σ_* as a function of S/N

To investigate the effects of S/N level on the recovery of σ_* , we generate a series SSP model using the MILES Tune SSP model spectra webtool (Vazdekis et al. 2010). We limit the stellar population ages to 0.1, 1, 5, and 10 Gyr and metallicities $[M/H] = -0.35, 0.15$, and 0.40, and initialize the simulated spectrum at a stellar velocity dispersion $\sigma_* = 90 \text{ km s}^{-1}$, taking into account the wavelength-dependent dispersion of the SDSS. We then artificially add normally distributed random noise at various S/N ratios to simulate real observations, and 10 mock spectra are generated per S/N level. The S/N is measured relative to the value of the data in each spectral channel. No other spectral components are added to these SSP models, and only v_* and σ_* are fit.

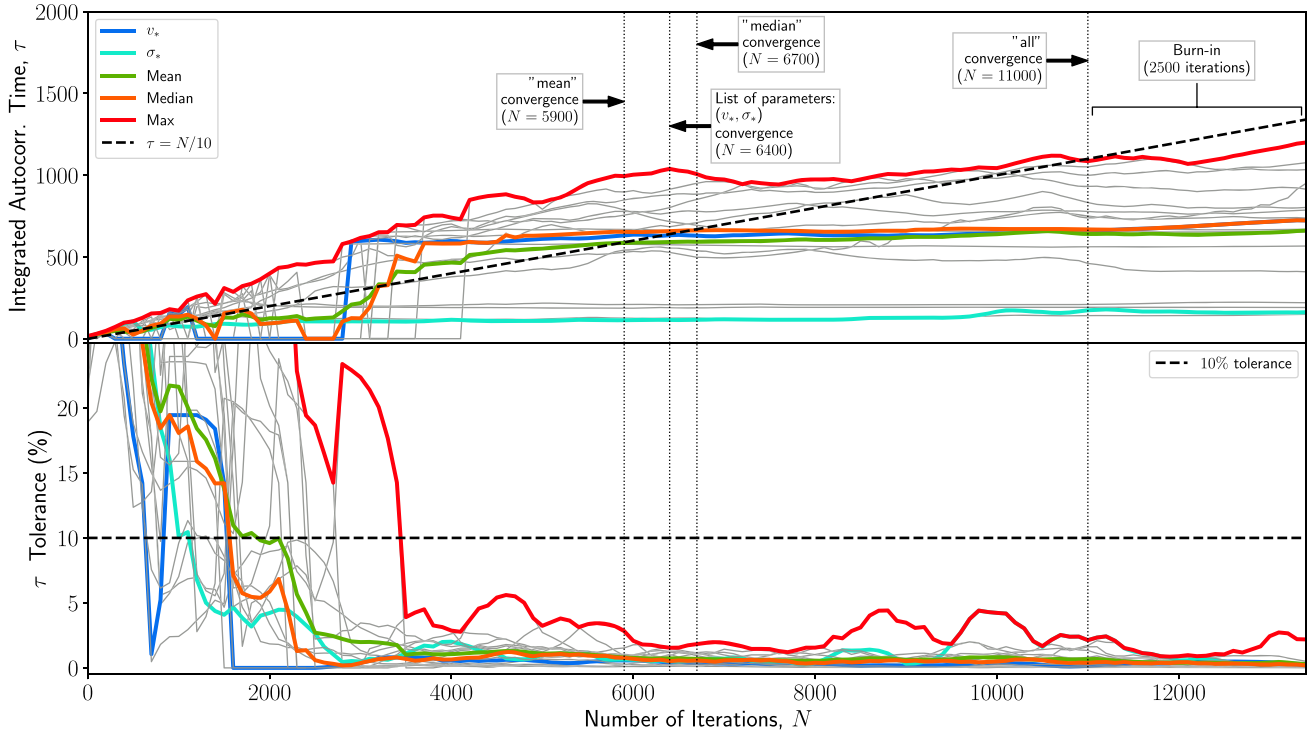


Figure 8. An example of different modes of autocorrelation convergence using a model spectrum with 17 free parameters. The ‘median’ mode of convergence typically converges the fastest because it tends to omit outlier τ values, which are typically values for which there are too few iterations to determine an accurate integrated autocorrelation time. By specifying the LOSVD parameters (v_* and σ_*), convergence is reached in 6400 iterations, however many other parameters have not yet converged. The ‘mean’ convergence, is a more strict convergence criteria because all τ estimates are weighted equally. Requiring ‘all’ parameters to converge requires the highest number of iterations and does not guarantee convergence if some parameters cannot converge within the maximum number of iterations.

Fig. 9 shows the results of S/N tests of various SSP models, where we calculate the percent error (deviation of the best-fitting value from the actual value) and the percent uncertainty in σ_* . The results of these tests provide a lower limit to the S/N of SDSS spectra for which σ_* can be reliably measured. Below a S/N of ~ 15 , the best-fitting measurement of σ_* begins to exceed the actual value by more than 10 per cent, and becomes increasingly unreliable at lower S/N. We find a similar result for the average uncertainties of σ_* . We therefore do not recommend measuring the LOSVD at S/N < 20 if the scientific goal is to report accurate stellar kinematics.

There is also a clear offset for the youngest (0.1 Gyr) SSP models, which can be explained by the lack of younger stellar templates included with our code, since these stellar types are considerably more rare. In cases where our code is used for fitting active star-forming galaxies, we recommend one include more O- and B-type templates for fitting the LOSVD and/or disabling the power-law component.

2.3.2 Recovery of σ_* as a function of Fe II emission

To test the effects of Fe II emission on the measurement of σ_* , we hold the amplitude of the 10 Gyr, [M/H] = 0.15 MILES SPP model constant and incrementally add broad and narrow Fe II at an increasing amplitudes, and fit σ_* at S/N levels of 10, 25, 50, 75, and 100. We generate and fit 10 mock spectra per Fe II amplitude and S/N level. We define the Fe II fraction as a function of stellar continuum amplitude, such that when Fe II amplitude is equal to the stellar continuum amplitude, the Fe II fraction is 100 per cent. We assume

that stellar absorption features are at the same velocity (redshift) as Fe II features, and note that subtle differences in velocity may make recovery of σ_* more difficult, however, since narrow and broad Fe II features are typically blended due to the resolution of SDSS, we can only reliably measure the relative amplitude of Fe II emission. At the very least, recovery of σ_* as a function of Fe II fraction gives insight as to how stellar template fitting can recover the underlying stellar continuum when broad and narrow emission line features are superimposed on them.

The results of our tests in the recovery of σ_* as a function of Fe II fraction are shown in Fig. 10. There is a weak dependence of the best-fitting value of σ_* as Fe II fraction increases, and the variance in σ_* increases with decreasing S/N. Even in the most extreme cases, where Fe II fraction exceeds 50 per cent, σ_* can be reliably recovered. Similarly for the uncertainty in σ_* , we find that Fe II fraction has no discernible effect on the measured uncertainties, and are more dependent on S/N. While the effects of differing velocities between stellar absorption and Fe II features are not taken into consideration, these tests indicate that even extreme fractions of Fe II will not significantly affect stellar template matching of the underlying stellar continuum, as long as absorption features are not significantly diluted by the AGN continuum.

2.3.3 Recovery of σ_* as a function of AGN continuum dilution

To simulate the effects of AGN continuum dilution, we construct a model spectrum using the MILES SSP 10 Gyr [M/H] = 0.15 template held at a constant amplitude, followed by a flat ($\alpha_\lambda = 0.0$)

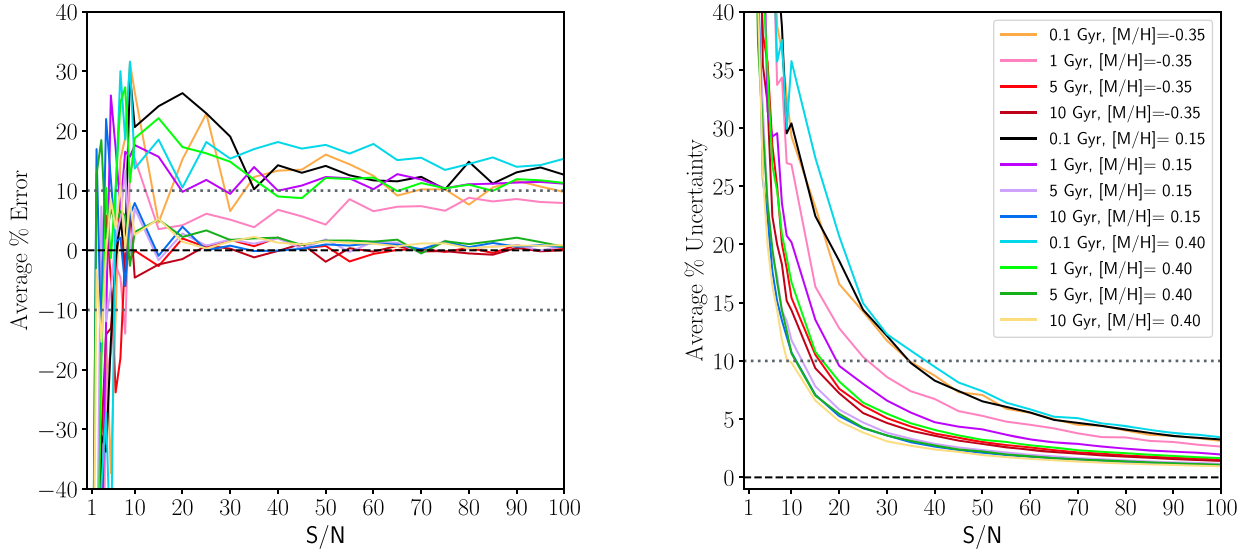


Figure 9. Recovery of σ_* as a function of S/N for different fitted MILES SSP models. The figure on the left shows the percent error relative to the actual value of σ_* , which exceeds 10 per cent (green dotted line) below S/N ~ 15 . The figure on the right shows the percent uncertainty, which increases with decreasing S/N. This provides a minimum S/N for which LOSVD measurements can be recovered.

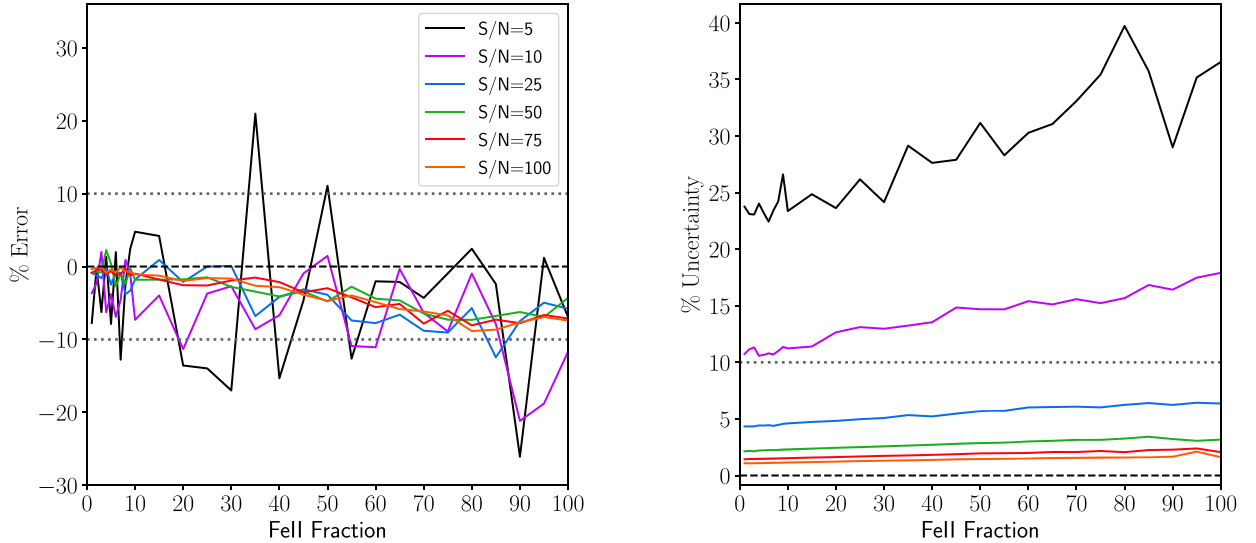


Figure 10. Recovery of σ_* as a function of Fe II fraction. There is a weak dependence on the value of σ_* due to Fe II fraction, and a stronger dependence on S/N. Even in the most extreme cases in which the Fe II fraction exceeds 50 per cent, Fe II emission does not significantly affect stellar template fitting.

continuum of increasing amplitude before re-normalizing the model. Because young stellar types have a similar continuum shape as an AGN power-law continuum, we do not want the effects of template mismatch to skew measurements of σ_* as a function of dilution of stellar absorption features, therefore we initialize the continuum to be flat.

We define the percent continuum dilution as the ratio of the amplitude of the AGN continuum to the constant amplitude of the stellar continuum, and generate model spectra of percent continuum dilution ranging from 0 per cent to 140 per cent. We then fit σ_* at S/N levels of 5, 10, 25, 50, 75, and 100. For each S/N level and continuum dilution level, 10 mock spectra are generated to determine

a mean percent error from the true value of σ_* and mean uncertainty. For the fit, we only fit the LOSVD and do not include a power-law continuum model component, as we wish to determine the effect of dilution on the measurement of σ_* .

Fig. 11 shows the effects of AGN continuum dilution on the recovery of σ_* . We find that continuum dilution can have significant effects at both low levels of dilution and is independent of S/N. Because absorption features are nearly Gaussian (best approximated using Gauss-Hermite polynomials for high-resolution spectra), the correlation between amplitude and width causes the fitting algorithm to fit broader widths to absorption features with shallower depths, caused by the inclusion of the AGN continuum. The effect of this

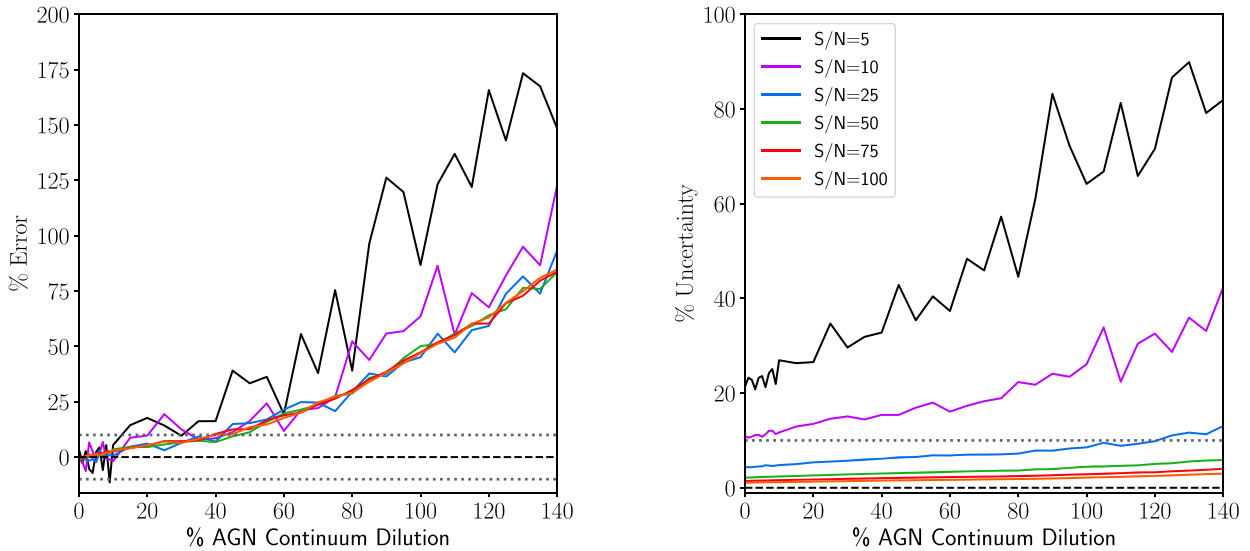


Figure 11. The effects of AGN continuum dilution on the recovery of σ_* when a power-law continuum model component is not included in the fit. At all S/N levels, the measured σ_* can be biased to larger values at even low levels of continuum dilution. Inclusion of a power-law component in the fit to the LOSVD completely resolves the percent error attributed to continuum dilution, and reduces the error to that attributed to S/N alone.

dilution can be seen as the positive trend in per cent error with increasing per cent dilution in Fig. 11. The error in measured σ_* will exceed 10 per cent with only 40 per cent continuum dilution, and exceeds 50 per cent at 100 per cent dilution (when AGN continuum and stellar continuum amplitudes are equal). We find similar trend for the uncertainties in σ_* , although the rate of increase is not as dramatic.

Fortunately, the inclusion of a power-law continuum model to the fit completely solves the problem of continuum dilution, assuming that the power-law model accurately represents the observed continuum of observed type 1 AGNs. The power-law spectrum allows PPXF to accurately recover the true value of σ_* , at even the highest levels of continuum dilution, and reduces the error to that attributed to S/N alone. Therefore, we highly recommend including the power-law model if fitting the LOSVD, for both type 1 and type 2 AGNs, since dilution need not be accompanied by a strong power-law slope typically observed in type 1 AGNs.

At extremely high levels of continuum dilution, the amplitude of absorption features becomes consistent with the amplitude of noise and other variations in the stellar continuum, and recovery of σ_* becomes impossible. This is observed in type 1 AGNs with strong power-law continuum component, whose absorption features are typically significantly diluted or not observable.

Additionally, we tested the recovery of σ_* as a function of the power-law continuum slope to investigate the effects of possible stellar template mismatch, for example, a steep AGN power-law slope resembling the steep stellar continuum of an O- or B-type star, however, we did not find any significant difficulties in the recovery of σ_* .

For objects that exhibit strong Fe II, such as NLS1 (Véron-Cetty et al. 2001; Xu et al. 2012) or broad absorption line (BAL) objects (Boroson & Meyers 1992; Zhang et al. 2010), we find that strong Fe II is usually accompanied by a steep power-law continuum, indicating very strong AGN continuum fraction and thus dilution. Our tests indicate that it isn't necessarily the presence of Fe II or a steep power-law slope, but the presence of strong continuum dilution that makes it nearly impossible to recover the LOSVD in NLS1 or BAL objects,

and extra caution should be used when interpreting LOSVD fitting results from these types of objects.

3 APPLICATION: CORRELATION ANALYSIS OF [O III] OUTFLOWS

3.1 Motivation

The emergence of scaling relations between supermassive black holes (henceforth, BHs) and their respective host galaxies implies that there is a fundamental mechanism that regulates their co-evolution (Kormendy & Ho 2013; DeGraf et al. 2015), however the source of this mechanism remains poorly understood. Large statistical studies of galaxies have since established that galactic-scale outflows are commonplace in galaxies that harbour AGNs, hinting that AGN-driven outflows are strong candidates as the feedback messengers between BHs and their host galaxies (Woo et al. 2016; DiPompeo et al. 2018; Rakshit & Woo 2018; Wang, Xu & Wei 2018). There is some observational evidence and theoretical arguments that point to the AGN as the central engine powering galactic-scale outflows (Fabian 2012; King & Pounds 2015). Additionally, some numerical simulations indicate that AGN feedback can act to disrupt gas cooling and subsequent star formation on galactic scales (Croton et al. 2006; Dubois et al. 2013; Costa, Pakmor & Springel 2020), which could give rise to the scaling relations we observe today.

Evidence of such feedback is believed to manifest itself at optical wavelengths as a broad flux-excess in the base or wings of ionized gas emission lines. The flux-excess, which is most easily identified as extended emission in [O III] $\lambda 5007$, is typically found to be blueshifted with respect to the core component of the line, resulting in a significantly asymmetric line profile (Woo et al. 2016; Komossa, Xu & Wagner 2018). These so-called blue wing outflow components, which have widths ranging from a few hundred to a few thousand kilometers per second (Harrison et al. 2014; Zakamska et al. 2016; Manzano-King et al. 2019), can be interpreted as outflowing ionized gas that is no longer gravitationally bound to the NLR of the galaxy. The absence of a 'red wing' in the profile could also indicate

significant dust attenuation of outflowing gas moving radially along the line of sight, possibly due to the presence of a galactic disc or AGN torus structure (Bae & Woo 2016).

Ionized outflows in narrow forbidden emission lines were first identified in early studies of individual radio sources (Grandi 1977; Afanasev et al. 1980) and it was soon found that signatures of blueshifted outflows were common in larger samples of Seyfert and radio galaxies (Heckman et al. 1980; De Robertis & Osterbrock 1984; Whittle 1985). It was Heckman et al. (1980) that first suggested that because the source of radio emission is due to a compact non-thermal central radio source, the outflow emission must originate along the line of sight between the observer and nuclear region of the galaxy. Later, Heckman, Miley & Green (1984) confirmed a relatively strong correlation between radio emission and the presence of outflows which holds true to this day (Jackson & Browne 1991; Veilleux 1991; Brotherton 1996; Mullaney et al. 2013; Zakamska & Greene 2014). The interest in ionized gas outflows has accelerated within recent years to include extensive IFU observations (Müller-Sánchez et al. 2011; Bae et al. 2017; Freitas et al. 2018; Wylezalek et al. 2020) and hydrodynamical simulations (Melioli & de Gouveia Dal Pino 2015; Costa et al. 2020).

The kinematic properties of ionized gas outflows in relation to other galaxy and AGN properties have also been explored in detail since their discovery. Nelson & Whittle (1996) first studied the relationship between the bulge and NLR stellar and gas kinematics, showing that the stellar velocity dispersion is relatively correlated with the [O III] gas dispersion, largely due to the gravitational potential of the bulge. However, they noted that [O III] lines with blue wings do not correlate as well with stellar velocity dispersion, indicating the presence of a strong non-gravitational component (Nelson & Whittle 1996; Mullaney et al. 2013; Woo et al. 2016; DiPompeo et al. 2018; Rakshit & Woo 2018; Wang et al. 2018). When the blue wing outflow component is properly removed from [O III] line profile, and if any possible Fe II contamination is accounted for, there is better agreement with stellar velocity dispersion (Boroson 2003; Greene & Ho 2005). Although the correlation cannot be used on an object-to-object basis, correcting [O III] for outflow components Fe II emission provides a means to estimate stellar velocity dispersion for scaling relations such as the $M_{\text{BH}}-\sigma_*$ relation for larger, higher redshift statistical samples for which stellar absorption features cannot easily be measured (Wang & Lu 2001; Boroson 2003; Woo et al. 2006; Komossa & Xu 2007; Bennert et al. 2018; Sexton et al. 2019).

Correlations between the kinematics of the [O III] profile and properties of the AGN also exist. For instance, by studying the combined (core + outflow) [O III] profile of large samples of type 2 and type 1 SDSS AGNs, Woo et al. (2016) and Rakshit & Woo (2018) found that the launching velocity of outflows increases with AGN luminosity. Although these studies examined the combined flux-weighted kinematics of the [O III] profile, the results of Bennert et al. (2018) imply that the core component of the [O III] profile can be independently used to estimate stellar velocity dispersion once the outflow component has been removed. This invites inquiry as to whether or not the core or outflow components *independently* may exhibit other relationships with each other or host galaxy properties.

Since detection and fitting of outflow components in [O III] is a feature specifically implemented in our code, it presents an opportunity to examine both the core and outflow component kinematics of the [O III] line to investigate all correlations related to ionized gas outflows and the host galaxy to further understand the physical interpretation of the emission line profile. Furthermore, while the use of powerful techniques such as integral field spectroscopy are

becoming mainstream for studying the spatially resolved kinematics of AGN host galaxies, these studies are limited to nearby objects. Our objective here is to study the emission line profile of [O III] in local AGNs with outflows to better understand the relationships between outflows, AGNs, and their host galaxies, and apply our knowledge in future studies to objects in the non-local universe for which spatially resolved observations are not possible.

Throughout the following sections, we refer to individual objects in our sample using their truncated object ID, for example, J001335. We commonly refer to different components of the double-Gaussian ‘outflow’ model of the [O III] profile as ‘core’ and ‘outflow’ when referring to specific quantities of each, such as σ_{core} for the core component velocity dispersion.

3.2 Sample selection

Since we wish to investigate the relationships between emission line outflow properties of [O III] and both the AGN and host galaxy, we select previously studied nearby type 1 AGNs with SDSS spectroscopy for which BH mass can be measured from the broad H β emission and fit these objects using our code. For this we included 81 type 1 AGNs studied by Bennert et al. (2018) (henceforth B18) originally selected from SDSS Data Release 6 (DR6), which selected BH masses ($6.6 \leq \log_{10}(M_{\text{BH}}/M_{\odot}) \leq 8.7$) at redshifts ($0.02 \leq z \leq 0.10$). The study from B18 performed detailed follow-up observations with Keck/LRIS and performed a detailed decomposition of the [O III] profile to study how different line decompositions affect the $M_{\text{BH}}-\sigma_*$ relation, however we fit only the SDSS spectra from B18 here for the purposes of benchmarking the capabilities of our algorithm.

To extend the sample to lower mass BHs we include NLS1 objects from Woo et al. (2015) (henceforth W15), which have a BH mass range of ($5.6 \leq \log_{10}(M_{\text{BH}}/M_{\odot}) \leq 7.4$) and redshift range ($0.01 \leq z \leq 0.10$). The W15 sample was selected from SDSS DR7 by sequentially selecting objects with ($500 \text{ km s}^{-1} < \text{FWHM}_{\text{Br-H}\beta} \leq 2000 \text{ km s}^{-1}$), ($800 \text{ km s}^{-1} < \text{FWHM}_{\text{Br-H}\alpha} \leq 2200 \text{ km s}^{-1}$), and a line flux ratio of $[\text{O III}]/\text{H}\beta < 3$, resulting in a final sample of 93 NLS1s.

Finally, we include five objects from Sexton et al. (2019) (henceforth S19) for which there is sufficient S/N to adequately measure σ_* in the SDSS spectra and have previously determined outflow signatures in the [O III] profile. The S19 sample consists of 22 type 1 AGNs observed with Keck-I LRIS comprised of both BLS1s and NLS1s. The S19 sample has a BH mass range of ($6.3 \leq \log_{10}(M_{\text{BH}}/M_{\odot}) \leq 8.3$) and is comprised of objects in a broad range of redshifts ($0.03 \leq z \leq 0.57$) used to study evolution in the $M_{\text{BH}}-\sigma_*$ relation in the non-local universe. The 5 objects we include here consists of 4 BLS1 objects and one NLS1 object, which have a BH mass range of ($6.9 \leq \log_{10}(M_{\text{BH}}/M_{\odot}) \leq 8.2$) and range in redshift from ($0.09 \leq z \leq 0.43$) as reported by S19.

We removed 16 objects (8 from B18, 8 from W15) for which we could not fit a broad H β line, i.e. are type 2 AGNs or have significant host galaxy absorption that makes fitting the broad line highly uncertain. We also removed one object (J112229) listed twice in table 1 of W15 after confirming there were no nearby neighbours.

The final sample of 162 objects span a BH mass range of ($5.6 \leq \log_{10}(M_{\text{BH}}/M_{\odot}) \leq 8.7$) and a average redshift of $z = 0.06$. Of these 162 objects, 76 contain measurable outflows in the [O III] line profile as determined using the outflow criteria given in Section 2.2.2. We also included an additional six objects which have some visually identifiable asymmetry in the [O III] line profile, which may be attributable to outflows, bringing the total number of objects with outflows to 82. We plot the distribution of BH mass for all 162 objects in Fig. 12.

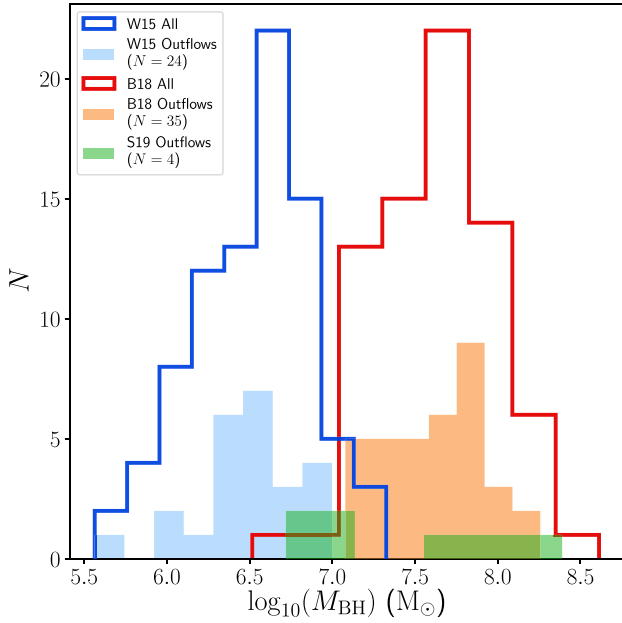


Figure 12. Histogram of BH mass for the individual samples from B18, W15, and S19 used for our sample of 162 type 1 AGN. The shaded regions indicate the 63 objects for which we have detected significant non-gravitational outflow signatures in the [O III] profile.

3.3 Methods

3.3.1 Spectral fitting

All 162 objects are re-fit with our code in two different ways. The first fit is forced to include outflow components in [O III] even if they do not satisfy the outflow criteria used by our algorithm. The second fit is forced to not include outflow components in [O III]. Because we wish to investigate how the non-gravitational component of the [O III] profile correlates with other galaxy properties, both single- and double-Gaussian profiles must be fit to determine which decomposition produces better agreement with σ_* . In both fits, we include all other model components, i.e. broad line H β , narrow and broad Fe II, power-law continuum, and the LOSVD, in the wavelength range ($4400 \leq \lambda \leq 5800$) following the methods from S19. This fitting region is ideal, not only because it contains the emission lines we want to study, but is also large enough to adequately constrain the amplitude of Fe II emission such that it can be distinguished from the stellar absorption features near Mg IB used to estimate the LOSVD.

We allow our algorithm to fit for a minimum of 2500 iterations with 100 walkers until the LOSVD parameters (stellar velocity and velocity dispersion), and emission line parameters (amplitude, width, and velocity offset) for the broad H β , [O III] core, and [O III] outflow components have achieved convergence at a minimum of 10 times the autocorrelation time and within a 10 per cent autocorrelation tolerance, with a post-convergence burn-in of 2500 iterations. We set a maximum iteration ceiling of 50 000 iterations, however, objects with high S/N and clearly visible outflow profiles in [O III] typically converge by $\sim 12\,000$ iterations, which is actually ~ 5 – 10 times the autocorrelation time for the parameters we consider for convergence.

Using both the outflow criteria given in Section 2.2.2 and by visually inspecting the fits of both the outflow and no-outflow models, we determine that 82 of the 162 objects have outflow

components with significant width and offset differences from their core components.

Finally, since we wish to examine the effects of the non-gravitational outflow component of the [O III] profile and compare them to the stronger gravitational component σ_* , we remove 19 objects for which we do not see an improvement in agreement between the σ_{core} and σ_* , which is necessary to remove objects for which our algorithm potentially overfit with a double-Gaussian profile and thus no strong non-gravitational outflow component is present.

The final sample includes 63 objects with strong non-gravitational kinematics in the [O III] $\lambda 5007$ emission line, of which 55 outflow components are blueshifted and 8 outflow components are redshifted relative to their core components. The 63 objects with strong outflows are shown in the shaded regions of 12 and listed in Table 1 with their relevant measurements.

To visualize the diversity of the 63 strong [O III] outflows in our sample, we align their full profiles by shifting them to the rest-frame velocity of the core component and normalize them by the amplitude of the full profile, as shown in Fig. 13, with the luminosity-weighted average shown by the red profile.

3.3.2 Correcting σ_* for disc inclination

The relatively large 3 arcsec diameter SDSS fibre can cover a significant fraction of the host galaxy, introducing contamination from non-bulge components. This is of particular concern for σ_* measurements on the $M_{\text{BH}} - \sigma_*$ relation since BH mass does not correlate with the stellar velocity dispersion of discs (see Kormendy & Ho 2013 for a review of all BH mass correlations).

As such, a significant number of objects in our sample contain discs, which at high inclinations, can artificially increase the measured stellar velocity dispersion, and overestimate values by as much as 25 per cent (Hartmann et al. 2014). Using N -body simulations, Bellovary et al. (2014) derived a prescription to correct measured σ_* to face-on ($i = 0$) values using common observables, which depend significantly on the inclination i and rotational velocity v_{rot} of the disc.

To correct the measured velocity dispersions in our sample, we first obtain disc inclinations and disc scale lengths from (Simard et al. 2011), who performed bulge + disc decompositions of over 1.1 million SDSS galaxies, from which we obtain measurements for 57 objects from our sample of 63. We then estimate the disc rotational velocities from scale lengths using the SDSS RV Relation from Hall et al. (2012). The prescription from Bellovary et al. (2014) depends on the ratio $(v/\sigma)_{\text{spec}}$, for which we assume a value of 0.6 for a fast-rotating late-type galaxy (Falc3n-Barroso et al. 2017) following the same procedure from S19. We note that varying values of $(v/\sigma)_{\text{spec}}$ do not significantly change the correction factor as much as values for i and v_{rot} . We propagate all uncertainties in quadrature and assume an additional 10 per cent uncertainty in correction prescription. The average change in σ_* due to this correction for all of our objects is only 12 km s^{-1} , but can be as high as 39 km s^{-1} for the highest of inclinations. We determine that despite this correction, the overall scatter for σ_* in our sample does not change, and that this correction will have a negligible effect on our results.

3.4 Results

The relevant measurements obtained from spectral fitting with our algorithm for the 63 objects in our sample are presented in Table 1. Calculated AGN luminosities at $5100 \text{ \AA}$ are obtained via the

Table 1. Measurements of relevant quantities for outflow, host galaxy, and AGN properties. Column 1: SDSS object designation. Column 2: Reference; (1) B18, (2) W15, (3) S19. Column 3: Systemic redshift determined using stellar kinematics. Column 4: Stellar velocity dispersion. Column 5: [O III] core component systemic velocity. Column 6: [O III] core component velocity dispersion. Column 7: [O III] core component luminosity. Column 8: [O III] outflow component systemic velocity. Column 9: [O III] outflow component velocity dispersion. Column 10: [O III] outflow component luminosity. Column 11: Maximal outflow velocity measured using W_{80} . Column 12: FWHM of the broad $H\beta$ emission line. Column 13: AGN luminosity at 5100 Å measured using the relation from Greene & Ho (2005). Column 14: BH mass estimated using relation from S19.

Object (SDSS)	Ref.	z	σ_* (km s ⁻¹)	v_{core} (km s ⁻¹)	σ_{core} (km s ⁻¹)	L_{core} (erg s ⁻¹)	v_{outflow} (km s ⁻¹)	σ_{outflow} (km s ⁻¹)	L_{outflow} (erg s ⁻¹)	v_{max} (km s ⁻¹)	FWHM _{Hβ} (km s ⁻¹)	$\lambda L_{5100 \text{ \AA}}$ (erg s ⁻¹)	log(M_{BH}) (M _{$\odot$})
J000338.94+160220.6	3	0.11668	91 ⁺¹³ ₋₁₃	-129 ⁺¹¹ ₋₁₁	107 ⁺⁴ ₋₄	41.01 ^{+0.02} _{-0.02}	-331 ⁺²⁴ ₋₂₆	397 ⁺²¹ ₋₂₁	40.90 ^{+0.03} _{-0.03}	-712 ⁺³⁵ ₋₃₆	3513 ⁺¹¹⁶ ₋₁₁₄	43.38 ^{+0.01} _{-0.01}	7.63 ^{+0.19} _{-0.13}
J001335.38-095120.9	1	0.06196	55 ⁺¹⁸ ₋₁₈	-38 ⁺¹⁴ ₋₁₃	127 ⁺¹¹ ₋₁₁	40.25 ^{+0.05} _{-0.05}	-284 ⁺⁴⁵ ₋₃₆	392 ⁺⁴⁰ ₋₃₀	40.28 ^{+0.05} _{-0.06}	-749 ⁺⁶⁴ ₋₇₅	3587 ⁺⁵³ ₋₅₀	43.34 ^{+0.01} _{-0.01}	7.62 ^{+0.19} _{-0.13}
J010939.01+005950.4	1	0.09376	105 ⁺¹⁵ ₋₁₄	-198 ⁺¹¹ ₋₁₁	119 ⁺³ ₋₃	41.26 ^{+0.01} _{-0.01}	-398 ⁺¹⁶ ₋₁₆	393 ⁺¹³ ₋₁₂	41.17 ^{+0.01} _{-0.01}	-705 ⁺²⁰ ₋₂₀	3279 ⁺¹⁰² ₋₉₆	43.23 ^{+0.02} _{-0.02}	7.47 ^{+0.12} _{-0.10}
J012159.81-010224.4	1	0.05484	116 ⁺¹² ₋₁₂	-130 ⁺¹² ₋₁₂	117 ⁺³ ₋₃	41.38 ^{+0.02} _{-0.02}	-303 ⁺¹² ₋₁₃	272 ⁺³ ₋₃	41.50 ^{+0.01} _{-0.01}	-521 ⁺⁷ ₋₇	4210 ⁺⁵¹ ₋₄₇	43.53 ^{+0.00} _{-0.00}	7.86 ^{+0.20} _{-0.13}
J021257.59+140610.0	1	0.06228	120 ⁺¹⁰ ₋₉	-88 ⁺⁵ ₋₅	138 ⁺³ ₋₃	40.83 ^{+0.01} _{-0.02}	-290 ⁺²⁵ ₋₂₇	413 ⁺³⁵ ₋₃₀	40.51 ^{+0.03} _{-0.03}	-732 ⁺⁴⁶ ₋₄₆	4221 ⁺⁹⁰ ₋₈₉	43.14 ^{+0.01} _{-0.01}	7.65 ^{+0.18} _{-0.13}
J024912.86-081525.7	2	0.02983	20 ⁺¹¹ ₋₁₁	-72 ⁺⁵ ₋₅	62 ⁺⁴ ₋₄	39.60 ^{+0.02} _{-0.02}	-212 ⁺²² ₋₂₂	303 ⁺²³ ₋₂₀	39.42 ^{+0.03} _{-0.03}	-529 ⁺³⁴ ₋₃₅	891 ⁺³⁸ ₋₃₈	41.76 ^{+0.02} _{-0.03}	5.56 ^{+0.22} _{-0.13}
J030124.26+011022.8	1	0.07216	94 ⁺¹¹ ₋₁₁	-172 ⁺¹⁰ ₋₁₀	120 ⁺⁶ ₋₆	40.50 ^{+0.02} _{-0.02}	-560 ⁺³⁷ ₋₄₁	487 ⁺³² ₋₃₀	40.51 ^{+0.03} _{-0.03}	-1013 ⁺⁵³ ₋₅₆	3263 ⁺⁶⁵ ₋₆₆	43.22 ^{+0.02} _{-0.02}	7.47 ^{+0.19} _{-0.13}
J030144.19+011530.8	1	0.07558	98 ⁺¹⁰ ₋₁₀	-208 ⁺⁸ ₋₈	131 ⁺⁵ ₋₅	40.89 ^{+0.02} _{-0.02}	-517 ⁺¹⁶ ₋₁₈	345 ⁺⁸ ₋₈	41.00 ^{+0.02} _{-0.02}	-752 ⁺¹⁹ ₋₂₁	3622 ⁺⁴⁷ ₋₄₅	43.42 ^{+0.01} _{-0.01}	7.68 ^{+0.21} _{-0.14}
J030417.78+002827.2	2	0.04488	55 ⁺⁸ ₋₈	-81 ⁺⁵ ₋₅	63 ⁺⁴ ₋₄	40.27 ^{+0.03} _{-0.03}	-172 ⁺¹⁰ ₋₁₁	166 ⁺⁶ ₋₆	40.26 ^{+0.03} _{-0.03}	-304 ⁺¹² ₋₁₃	1505 ⁺²⁷ ₋₂₇	42.80 ^{+0.01} _{-0.01}	6.58 ^{+0.22} _{-0.14}
J073106.86+392644.5	2	0.04894	32 ⁺¹⁰ ₋₁₀	-134 ⁺⁶ ₋₆	107 ⁺³ ₋₃	40.17 ^{+0.01} _{-0.01}	-363 ⁺¹⁰ ₋₁₀	318 ⁺⁵ ₋₅	40.28 ^{+0.01} _{-0.01}	-637 ⁺¹¹ ₋₁₁	1492 ⁺³⁵ ₋₃₄	42.38 ^{+0.01} _{-0.01}	6.36 ^{+0.14} _{-0.14}
J073505.65+423545.7	3	0.08644	45 ⁺¹² ₋₁₃	-55 ⁺⁷ ₋₇	76 ⁺⁶ ₋₆	40.59 ^{+0.05} _{-0.05}	-156 ⁺¹¹ ₋₁₂	174 ⁺⁷ ₋₆	40.70 ^{+0.04} _{-0.04}	-324 ⁺¹³ ₋₁₄	1712 ⁺⁶¹ ₋₆₀	42.85 ^{+0.02} _{-0.02}	6.72 ^{+0.19} _{-0.13}
J073703.28+424414.6	1	0.08861	102 ⁺¹⁰ ₋₁₀	-46 ⁺⁸ ₋₉	135 ⁺³ ₋₃	41.27 ^{+0.01} _{-0.02}	-249 ⁺²³ ₋₂₇	318 ⁺²⁴ ₋₂₂	40.83 ^{+0.04} _{-0.04}	-610 ⁺³⁶ ₋₃₈	4004 ⁺⁶³ ₋₆₅	43.33 ^{+0.01} _{-0.01}	7.71 ^{+0.21} _{-0.14}
J073714.28+292634.1	2	0.08029	83 ⁺⁹ ₋₁₀	-160 ⁺¹⁰ ₋₁₀	108 ⁺⁹ ₋₁₂	40.41 ^{+0.06} _{-0.07}	-353 ⁺⁶⁹ ₋₁₀₃	231 ⁺⁴¹ ₋₄₂	40.07 ^{+0.17} _{-0.15}	-488 ⁺⁸⁸ ₋₁₁₇	2395 ⁺¹⁸⁹ ₋₁₈₉	42.62 ^{+0.05} _{-0.04}	6.89 ^{+0.21} _{-0.15}
J080243.40+310403.3	1	0.04130	99 ⁺⁸ ₋₈	-26 ⁺⁶ ₋₆	104 ⁺³ ₋₃	40.69 ^{+0.01} _{-0.02}	-199 ⁺³⁶ ₋₄₄	301 ⁺⁴⁸ ₋₄₁	40.07 ^{+0.06} _{-0.06}	-559 ⁺⁶⁴ ₋₆₉	5511 ⁺⁸¹ ₋₈₀	43.22 ^{+0.01} _{-0.01}	7.93 ^{+0.22} _{-0.14}
J081718.55+520147.7	2	0.03911	42 ⁺¹² ₋₁₃	-156 ⁺⁷ ₋₇	59 ⁺³ ₋₄	40.20 ^{+0.02} _{-0.02}	-101 ⁺²⁵ ₋₂₀	205 ⁺²⁸ ₋₂₄	39.63 ^{+0.09} _{-0.07}	318 ⁺³⁹ ₋₃₆	1941 ⁺⁵⁸ ₋₅₅	42.39 ^{+0.01} _{-0.01}	6.59 ^{+0.19} _{-0.14}
J082912.68+500652.3	2	0.04373	75 ⁺⁸ ₋₇	-56 ⁺⁶ ₋₆	77 ⁺¹ ₋₁	40.79 ^{+0.00} _{-0.00}	-225 ⁺¹⁰ ₋₁₀	358 ⁺¹⁰ ₋₉	40.38 ^{+0.01} _{-0.01}	-627 ⁺¹⁴ ₋₁₄	1017 ⁺²³ ₋₂₂	42.54 ^{+0.01} _{-0.01}	6.10 ^{+0.22} _{-0.14}
J085504.16+525248.3	2	0.08994	76 ⁺⁹ ₋₉	-171 ⁺¹⁵ ₋₁₅	206 ⁺¹⁸ ₋₂₂	40.60 ^{+0.07} _{-0.09}	-309 ⁺³² ₋₄₂	522 ⁺⁸⁰ ₋₆₈	40.60 ^{+0.08} _{-0.07}	-808 ⁺⁹⁴ ₋₉₃	2092 ⁺¹⁰⁴ ₋₉₃	42.96 ^{+0.02} _{-0.02}	6.95 ^{+0.22} _{-0.14}
J090902.35+133019.4	1	0.05005	88 ⁺¹⁰ ₋₁₀	-69 ⁺¹⁰ ₋₁₀	92 ⁺¹⁵ ₋₁₃	39.72 ^{+0.10} _{-0.07}	4 ⁺²⁴ ₋₁₉	249 ⁺²⁷ ₋₂₇	39.89 ^{+0.05} _{-0.07}	392 ⁺³⁴ ₋₃₁	3524 ⁺¹⁴⁷ ₋₁₃₄	42.49 ^{+0.02} _{-0.02}	7.15 ^{+0.20} _{-0.13}
J092343.00+225432.7	1	0.03357	137 ⁺⁷ ₋₆	-157 ⁺⁸ ₋₈	146 ⁺³ ₋₃	41.16 ^{+0.01} _{-0.01}	-233 ⁺⁸ ₋₈	417 ⁺⁷ ₋₆	41.37 ^{+0.01} _{-0.01}	-612 ⁺⁹ ₋₉	3598 ⁺²⁶ ₋₂₆	43.69 ^{+0.00} _{-0.00}	7.81 ^{+0.21} _{-0.14}
J093259.60+040506.0	1	0.05990	103 ⁺⁴ ₋₅	-133 ⁺⁵ ₋₅	89 ⁺⁴ ₋₄	40.43 ^{+0.03} _{-0.03}	-282 ⁺²⁰ ₋₂₅	235 ⁺²² ₋₂₀	40.20 ^{+0.05} _{-0.05}	-450 ⁺³³ ₋₃₆	4829 ⁺²¹⁴ ₋₁₉₄	42.74 ^{+0.02} _{-0.02}	7.56 ^{+0.24} _{-0.14}
J094057.19+032401.2	2	0.06122	63 ⁺¹² ₋₁₂	-134 ⁺⁹ ₋₉	90 ⁺⁹ ₋₉	40.39 ^{+0.04} _{-0.04}	-291 ⁺³² ₋₃₈	335 ⁺³² ₋₂₈	40.34 ^{+0.05} _{-0.05}	-587 ⁺⁴⁸ ₋₅₂	1577 ⁺¹⁰¹ ₋₉₉	42.69 ^{+0.02} _{-0.02}	6.56 ^{+0.24} _{-0.14}
J094529.36+093610.4	2	0.01394	80 ⁺⁶ ₋₆	-167 ⁺⁴ ₋₄	114 ⁺² ₋₂	40.15 ^{+0.01} _{-0.01}	-247 ⁺⁶ ₋₆	300 ⁺⁷ ₋₇	39.95 ^{+0.02} _{-0.02}	-465 ⁺¹⁰ ₋₁₀	2084 ⁺⁵¹ ₋₅₀	41.95 ^{+0.01} _{-0.02}	6.41 ^{+0.22} _{-0.15}
J094838.43+403043.5	1	0.04771	92 ⁺¹¹ ₋₁₁	-176 ⁺⁷ ₋₇	103 ⁺² ₋₃	40.78 ^{+0.01} _{-0.01}	-304 ⁺⁴⁰ ₋₄₉	457 ⁺⁶³ ₋₅₆	40.12 ^{+0.05} _{-0.05}	-714 ⁺⁸² ₋₈₆	3374 ⁺⁶⁶ ₋₆₄	43.07 ^{+0.01} _{-0.01}	7.43 ^{+0.18} _{-0.14}
J104925.39+245123.7	1	0.05543	105 ⁺¹¹ ₋₁₁	-81 ⁺⁸ ₋₈	98 ⁺² ₋₂	41.15 ^{+0.01} _{-0.01}	-166 ⁺¹³ ₋₁₃	281 ⁺¹⁵ ₋₁₃	40.72 ^{+0.03} _{-0.03}	-445 ⁺¹⁹ ₋₂₀	5072 ⁺⁴⁷ ₋₄₅	43.48 ^{+0.00} _{-0.00}	8.00 ^{+0.20} _{-0.14}
J110016.03+461615.2	2	0.03257	63 ⁺⁵ ₋₅	-156 ⁺⁴ ₋₄	79 ⁺³ ₋₃	40.30 ^{+0.02} _{-0.02}	-256 ⁺⁷ ₋₈	213 ⁺⁶ ₋₅	40.23 ^{+0.02} _{-0.02}	-374 ⁺⁹ ₋₁₀	1433 ⁺⁵³ ₋₅₂	42.17 ^{+0.02} _{-0.02}	6.20 ^{+0.21} _{-0.14}
J110101.78+110248.8	1	0.03596	104 ⁺⁹ ₋₉	-69 ⁺⁶ ₋₆	115 ⁺³ ₋₃	40.93 ^{+0.01} _{-0.01}	-24 ⁺⁷ ₋₈	320 ⁺⁹ ₋₉	40.80 ^{+0.02} _{-0.02}	455 ⁺¹² ₋₁₂	6092 ⁺⁸⁰ ₋₈₂	43.13 ^{+0.01} _{-0.01}	7.97 ^{+0.20} _{-0.14}
J110456.03+433409.1	1	0.04952	70 ⁺⁷ ₋₇	-40 ⁺⁵ ₋₅	71 ⁺⁴ ₋₄	40.57 ^{+0.02} _{-0.02}	44 ⁺¹² ₋₁₀	218 ⁺¹³ ₋₁₂	40.34 ^{+0.04} _{-0.04}	363 ⁺¹⁹ ₋₁₉	4031 ⁺²⁵³ ₋₂₃₇	42.46 ^{+0.04} _{-0.03}	7.25 ^{+0.23} _{-0.14}
J112526.51+022039.0	2	0.04897	76 ⁺¹¹ ₋₁₁	-57 ⁺⁷ ₋₇	75 ⁺⁵ ₋₅	40.31 ^{+0.03} _{-0.04}	-99 ⁺¹⁸ ₋₂₂	229 ⁺²⁹ ₋₂₄	39.97 ^{+0.07} _{-0.07}	-336 ⁺³⁵ ₋₃₈	1618 ⁺¹²¹ ₋₁₁₆	42.31 ^{+0.03} _{-0.03}	6.38 ^{+0.23} _{-0.15}
J114545.18+554759.6	1	0.05419	96 ⁺¹² ₋₁₂	-159 ⁺⁸ ₋₈	79 ⁺¹¹ ₋₁₁	40.18 ^{+0.08} _{-0.07}	-122 ⁺¹³ ₋₁₂	248 ⁺²⁸ ₋₂₂	40.30 ^{+0.05} _{-0.05}	356 ⁺³⁰ ₋₃₀	3765 ⁺¹⁶⁸ ₋₁₆₀	42.67 ^{+0.03} _{-0.03}	7.31 ^{+0.19} _{-0.14}
J115333.22+095408.4	2	0.06965	99 ⁺⁹ ₋₉	-134 ⁺⁷ ₋₇	123 ⁺² ₋₃	41.19 ^{+0.01} _{-0.01}	-260 ⁺¹⁶ ₋₁₈	345 ⁺¹⁹ ₋₁₈	40.74 ^{+0.03} _{-0.03}	-569 ⁺²⁷ ₋₂₉	1937 ⁺⁶⁹ ₋₆₆	42.90 ^{+0.02} _{-0.02}	6.86 ^{+0.19} _{-0.13}
J120556.01+495956.4	1	0.06376	120 ⁺⁸ ₋₇	-170 ⁺⁶ ₋₆	148 ⁺² ₋₂	41.67 ^{+0.01} _{-0.01}	-208 ⁺⁹ ₋₉	390 ⁺¹⁸ ₋₁₇	41.05 ^{+0.03} _{-0.03}	-539 ⁺²² ₋₂₂	7451 ⁺¹⁷⁷ ₋₁₆₆	43.28 ^{+0.01} _{-0.01}	8.24 ^{+0.19} _{-0.13}
J120626.29+424426.1	1	0.05234	119 ⁺⁸ ₋₈	-100 ⁺⁷ ₋₇	110 ⁺⁴ ₋₄	40.47 ^{+0.01} _{-0.01}	-290 ⁺⁴² ₋₄₅	544 ⁺⁶⁵ ₋₅₇	40.14 ^{+0.04} _{-0.04}	-889 ⁺⁸⁴ ₋₈₆	3819 ⁺⁶² ₋₅₈	43.12 ^{+0.01} _{-0.01}	7.57 ^{+0.17} _{-0.14}
J121044.27+382010.3	1	0.02319	97 ⁺⁷ ₋₆	-75 ⁺⁶ ₋₆	103 ⁺⁴ ₋₄	40.70 ^{+0.02} _{-0.02}	-63 ⁺⁸ ₋₈	26					

Table 1 – *continued*

Object (SDSS)	Ref.	z	σ_* (km s ⁻¹)	v_{core} (km s ⁻¹)	σ_{core} (km s ⁻¹)	L_{core} (erg s ⁻¹)	v_{outflow} (km s ⁻¹)	σ_{outflow} (km s ⁻¹)	L_{outflow} (erg s ⁻¹)	v_{max} (km s ⁻¹)	FWHM _{Hβ} (km s ⁻¹)	$\lambda L_{5100\text{\AA}}$ (erg s ⁻¹)	log(M_{BH}) ($M_{\odot}$)
J160746.00+345048.9	2	0.05478	86 ⁺⁸ ₋₈	-200 ⁺⁷ ₋₈	108 ⁺⁴ ₋₄	40.61 ^{+0.02} _{-0.02}	-411 ⁺¹¹ ₋₁₁	424 ⁺⁹ ₋₈	40.84 ^{+0.01} _{-0.01}	-755 ⁺¹⁴ ₋₁₄	1651 ⁺⁴⁷ ₋₄₄	42.75 ^{+0.01} _{-0.01}	6.63 ^{+0.18} _{-0.13}
J161156.30+521116.8	1	0.04149	108 ⁺⁷ ₋₇	-44 ⁺⁶ ₋₆	131 ⁺⁴ ₋₄	40.50 ^{+0.02} _{-0.02}	-368 ⁺²⁹ ₋₃₁	426 ⁺²¹ ₋₂₁	40.33 ^{+0.03} _{-0.03}	-870 ⁺³⁹ ₋₄₁	3727 ⁺¹⁴¹ ₋₁₃₅	42.72 ^{+0.02} _{-0.02}	7.33 ^{+0.19} _{-0.13}
J163159.59+243740.2	2	0.04384	72 ⁺⁶ ₋₆	-58 ⁺⁶ ₋₆	94 ⁺² ₋₂	40.65 ^{+0.01} _{-0.01}	-207 ⁺²¹ ₋₂₄	278 ⁺¹⁶ ₋₁₆	39.97 ^{+0.04} _{-0.04}	-506 ⁺²⁸ ₋₃₁	1065 ⁺⁴⁸ ₋₄₄	42.37 ^{+0.02} _{-0.02}	6.07 ^{+0.18} _{-0.14}
J163501.46+305412.1	2	0.05460	95 ⁺¹⁵ ₋₁₄	-108 ⁺⁹ ₋₉	122 ⁺⁶ ₋₆	40.79 ^{+0.03} _{-0.02}	-223 ⁺¹¹ ₋₁₂	393 ⁺¹³ ₋₁₁	40.97 ^{+0.01} _{-0.02}	-620 ⁺¹⁷ ₋₁₇	2333 ⁺¹²⁰ ₋₁₁₀	42.73 ^{+0.02} _{-0.02}	6.94 ^{+0.18} _{-0.14}
J170859.15+215308.1	1	0.07277	123 ⁺¹² ₋₁₂	-85 ⁺⁹ ₋₉	164 ⁺⁴ ₋₄	40.94 ^{+0.01} _{-0.01}	2 ⁺¹⁵ ₋₁₅	551 ⁺²⁴ ₋₂₂	40.86 ^{+0.02} _{-0.02}	795 ⁺³¹ ₋₃₁	6325 ⁺¹²² ₋₁₁₅	43.37 ^{+0.01} _{-0.01}	8.13 ^{+0.18} _{-0.13}
J172759.14+542147.0	2	0.09989	40 ⁺¹⁹ ₋₁₆	-89 ⁺¹² ₋₁₃	73 ⁺³ ₋₃	40.80 ^{+0.02} _{-0.02}	-185 ⁺²⁹ ₋₃₄	241 ⁺³⁶ ₋₃₁	40.26 ^{+0.07} _{-0.06}	-406 ⁺⁴⁷ ₋₅₁	1295 ⁺⁷⁸ ₋₇₈	42.66 ^{+0.02} _{-0.02}	6.38 ^{+0.20} _{-0.13}
J205822.14-065004.3	2	0.07413	35 ⁺¹³ ₋₁₃	-59 ⁺⁸ ₋₈	96 ⁺² ₋₂	41.05 ^{+0.01} _{-0.01}	-225 ⁺¹² ₋₁₃	287 ⁺⁸ ₋₈	40.78 ^{+0.02} _{-0.02}	-533 ⁺¹⁴ ₋₁₄	1302 ⁺³² ₋₃₂	42.87 ^{+0.01} _{-0.01}	6.48 ^{+0.19} _{-0.13}
J210226.54+000702.3	2	0.05222	73 ⁺⁹ ₋₈	-92 ⁺⁸ ₋₉	77 ⁺⁷ ₋₈	39.93 ^{+0.05} _{-0.05}	-275 ⁺³⁹ ₋₄₉	199 ⁺³¹ ₋₂₈	39.73 ^{+0.09} _{-0.09}	-438 ⁺⁵³ ₋₆₀	1984 ⁺¹²⁹ ₋₁₁₆	42.19 ^{+0.03} _{-0.03}	6.50 ^{+0.20} _{-0.14}
J222246.61-081943.9	1	0.08312	111 ⁺⁸ ₋₈	-222 ⁺⁷ ₋₇	172 ⁺⁴ ₋₄	41.16 ^{+0.01} _{-0.01}	-599 ⁺¹³ ₋₁₃	570 ⁺⁸ ₋₈	41.35 ^{+0.01} _{-0.01}	-1109 ⁺¹⁵ ₋₁₆	3933 ⁺¹¹⁰ ₋₁₀₅	43.27 ^{+0.01} _{-0.01}	7.67 ^{+0.18} _{-0.13}
J223338.42+131243.5	1	0.09438	123 ⁺¹⁶ ₋₁₅	-203 ⁺¹⁰ ₋₉	150 ⁺⁴ ₋₄	41.40 ^{+0.02} _{-0.02}	-258 ⁺¹² ₋₁₃	483 ⁺²⁰ ₋₂₀	41.29 ^{+0.02} _{-0.02}	-675 ⁺²⁸ ₋₂₈	4326 ⁺⁵⁵ ₋₅₁	43.66 ^{+0.00} _{-0.00}	7.96 ^{+0.19} _{-0.13}
J235128.75+155259.1	1	0.09675	136 ⁺¹² ₋₁₁	-23 ⁺⁸ ₋₉	98 ⁺⁴ ₋₄	41.21 ^{+0.01} _{-0.01}	-135 ⁺⁸ ₋₈	249 ⁺² ₋₂	41.55 ^{+0.00} _{-0.00}	-431 ⁺⁴ ₋₄	7236 ⁺¹⁸⁰ ₋₁₇₅	43.43 ^{+0.01} _{-0.01}	8.28 ^{+0.19} _{-0.13}

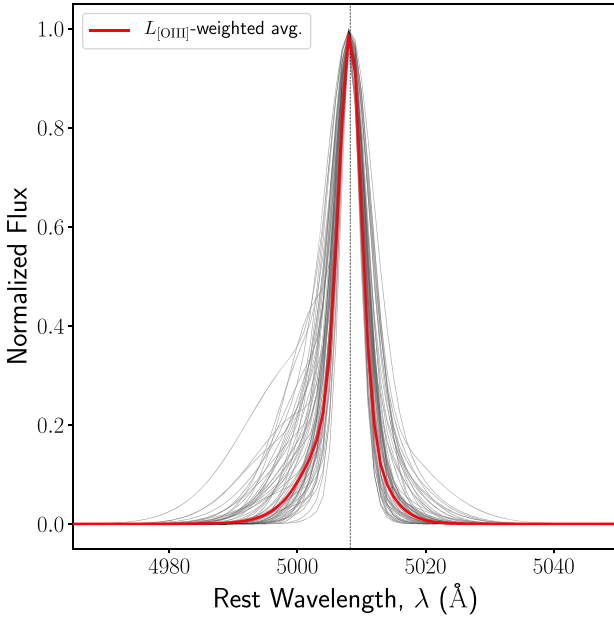


Figure 13. Superimposed [O III] profiles of all 63 outflow objects in our sample shown in black. Individual [O III] profiles are aligned at the rest-frame velocity of the core component and normalized by the maximum amplitude of the full [O III] profile. The luminosity-weighted average is shown by the red profile.

empirical relation between the luminosity of the broad H β emission line and $\lambda L_{5100\text{\AA}}$ from Greene & Ho (2005), and M_{BH} is calculated using the relation from S19 based on the mass recalibration from reverberation mapping measurements from Woo et al. (2015). Black hole masses for the 63 outflow objects span nearly three orders of magnitude from $(5.6 \leq \log_{10}(M_{\text{BH}}/M_{\odot}) \leq 8.4)$. To quantify the maximal velocity of the outflows, we adopt the relation from Harrison et al. (2014) given by

$$v_{\text{max}} = \Delta v_0 + \frac{W_{80}}{2}, \quad (11)$$

where v_0 is the velocity offset of the outflow component measured with respect to the velocity offset of the core component, and $W_{80} = 1.09$ FWHM, which represents the width containing 80 per cent of the Gaussian flux of the outflow component. Values of v_{max} listed in Table 1 are negative if the outflow is blueshifted with respect to the core component, and positive if redshifted with respect to the

core component. Velocities for the core and outflow components, v_{core} and v_{outflow} , are reported as velocities with respect to the systemic (stellar) velocity. All reported dispersion are corrected for the SDSS redshift-dependent instrumental dispersion during the fitting process by our code. The vast majority of our objects ($N = 55$) have blueshifted outflow components with respect to their core component.

We use these measurements to further investigate correlations of outflows with their host galaxy and AGN to understand their relationship, if any. For reference, we plot a heatmap of the Spearman’s rank correlation coefficient r_s for all relevant and possibly interesting quantities measured with our algorithm in Fig. 14. We discuss the most notable correlations in detail in the following subsections. In the following figures, we adopt a consistent colourscale shown in Fig. 15, which represents the absolute value of v_{max} .

3.4.1 Correlations with velocity

We first investigate correlations with the velocities of the [O III] core and outflow components. Following Woo et al. (2016), we plot velocity–velocity dispersion (VVD) diagrams of the core and outflow components in Fig. 15. The Spearman correlation coefficient for the [O III] core VVD diagram is $r_s = -0.19$, indicating no significant correlation, while the [O III] outflow VVD diagram shows stronger correlation of $r_s = -0.36$. The [O III] outflow VVD diagram also exhibits the same ‘fan’ shape characterized by Woo et al. (2016), which according to 3D biconical outflow models (Bae & Woo 2016), is caused by an increasing extinction due to the presence of an obscuring dust plane. In theory, if there exists a dust plane that bisects a biconical outflow for which one of the cones points towards the observer along the LOS (the blueshifted cone), the dust plane will obscure the cone on the far side (the redshifted cone) causing an observed blueshifted flux excess. The fan-shaped VVD diagram for outflows does not appear to extend to the core component. We also confirm the result from Rakshit & Woo (2018) that blueshifted outflows are significantly more common than redshifted outflows for type 1 AGN.

There does appear to be strong correlation ($r_s = 0.56$) between v_{core} and v_{outflow} measured with respect to the systemic (stellar) velocity, as shown in Fig. 16, which shows v_{core} scales linearly with v_{outflow} . Our data suggest that for blueshifted outflows (right of the dashed line in Fig. 16) there is an average offset of 120 km s⁻¹ between the core and outflow components, and there appears to be an increasing offset from v_{core} with v_{max} . A larger sample with $v_{\text{max}} > 900$ km

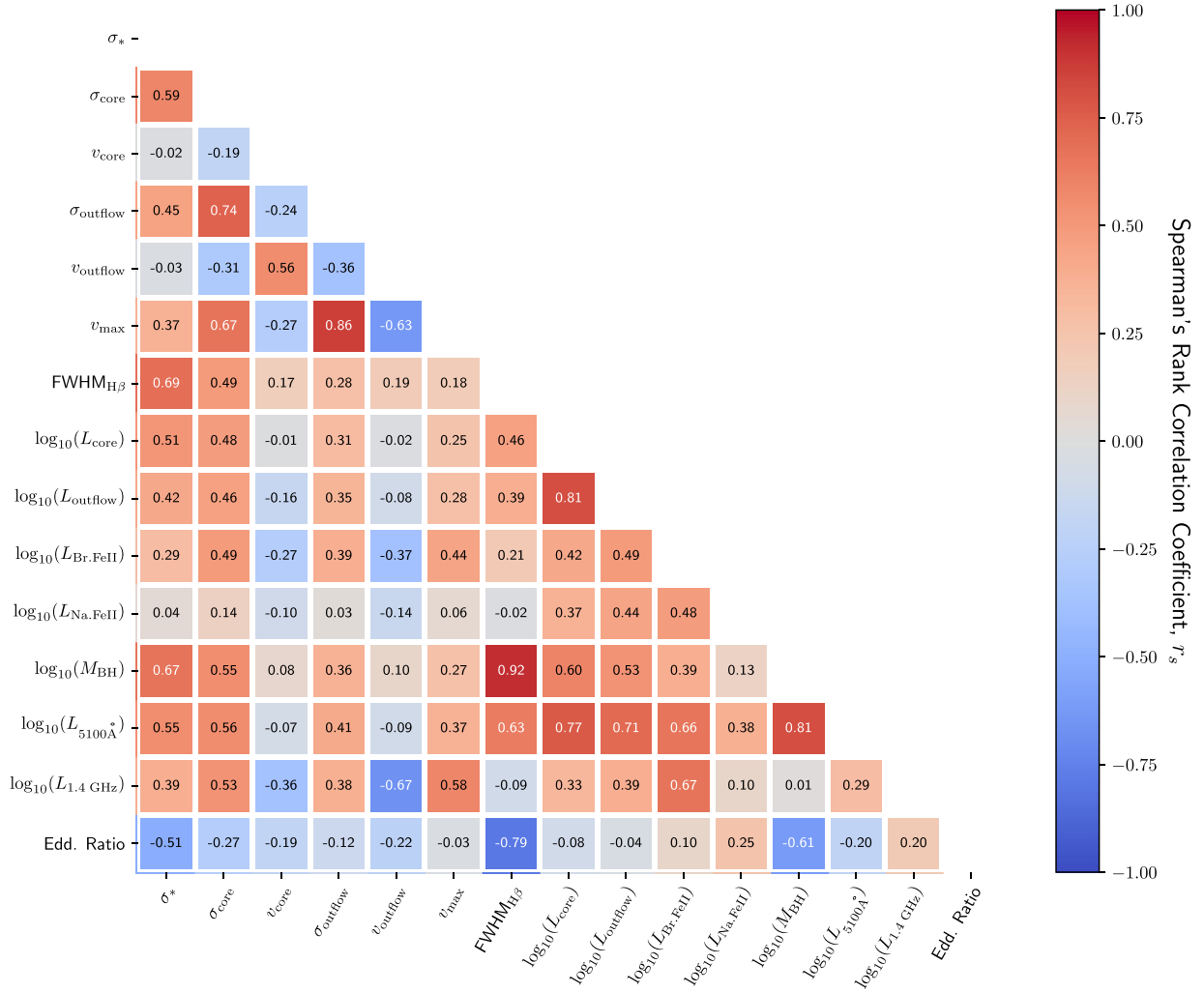


Figure 14. Correlation matrix of relevant quantities. The colourscale represents the absolute Spearman’s rank correlation coefficient r_s . The average uncertainty for all calculated values of r_s is 0.04.

s^{-1} and objects with redshifted outflows is needed to conclusively determine whether this trend holds, or if there is a value of v_{max} for which this trend no longer holds true.

3.4.2 Correlations with dispersion

In Fig. 17, we plot the single-Gaussian ‘no-outflow’ model [O III] dispersion σ_{single} alongside the double-Gaussian ‘outflow’ model core dispersion σ_{core} , both as a function of σ_* . We confirm the results from Bennert et al. (2018) that σ_{core} correlates more strongly with σ_* once the secondary outflow component is accounted for. Values of σ_{core} are scattered about the perfect correlation with σ_* , with a root-mean-square error of $27 \pm 2 \text{ km s}^{-1}$. The mean of this distribution of σ_{core} values is $18 \pm 2 \text{ km s}^{-1}$, caused primarily by objects with $v_{\text{max}} > 700 \text{ km s}^{-1}$.

In Fig. 18, we plot the difference ($\sigma_{\text{core}} - \sigma_*$) as a function of σ_{outflow} and find that there is some dependence on how well σ_{core} can recover the gravitational component σ_* as a function of σ_{outflow} and v_{max} , however a larger sample will also be needed to confirm this trend.

There also exists a strong correlation between the [O III] core component dispersion σ_{core} and the outflow component dispersion

σ_{outflow} as first reported by Zhang & Feng (2017), which we plot in Fig. 19 for our sample. This correlation has so far been largely overlooked, mostly due to the parametrization other studies have used to quantify outflows. For instance, Woo et al. (2016) performed double-Gaussian decomposition of the [O III], but did not study individual dispersions and instead adopted a flux-weighted integrated dispersion for the full line (core + outflow) profile. The Spearman rank correlation coefficient for this relation is $r_s = 0.74$, implying a very strong correlation, and stronger than the $\sigma_{\text{core}} - \sigma_*$ correlation ($r_s = 0.59$). We perform linear regression using EMCEE following the same methods used in S19, and determine a best-fitting slope of $m = 0.26 \pm 0.03$, intercept of $b = 25.00^{+9.05}_{-8.89} \text{ km s}^{-1}$, and intrinsic scatter of $f = 19.11^{+2.16}_{-2.08}$, which we plot in Fig. 19.

It is important to emphasize that the $\sigma_{\text{core}} - \sigma_{\text{outflow}}$ correlation is not a result of our definition of an ‘outflow’ or our selection criteria for outflows given in Section 2.2.2. Although we define an outflow to have $\sigma_{\text{outflow}} > \sigma_{\text{core}}$, which excludes objects above the dashed line ($\sigma_{\text{core}} = \sigma_{\text{outflow}}$) in Fig. 19 by design, this does not explain the tightness in the correlation below the dashed line. The outflow criterion for dispersion given in Section 2.2.2 also does not select objects based on the ratio of σ_{outflow} and σ_{core} , but by the ratio of the difference of σ_{outflow} and σ_{core} and their relative uncertainties.

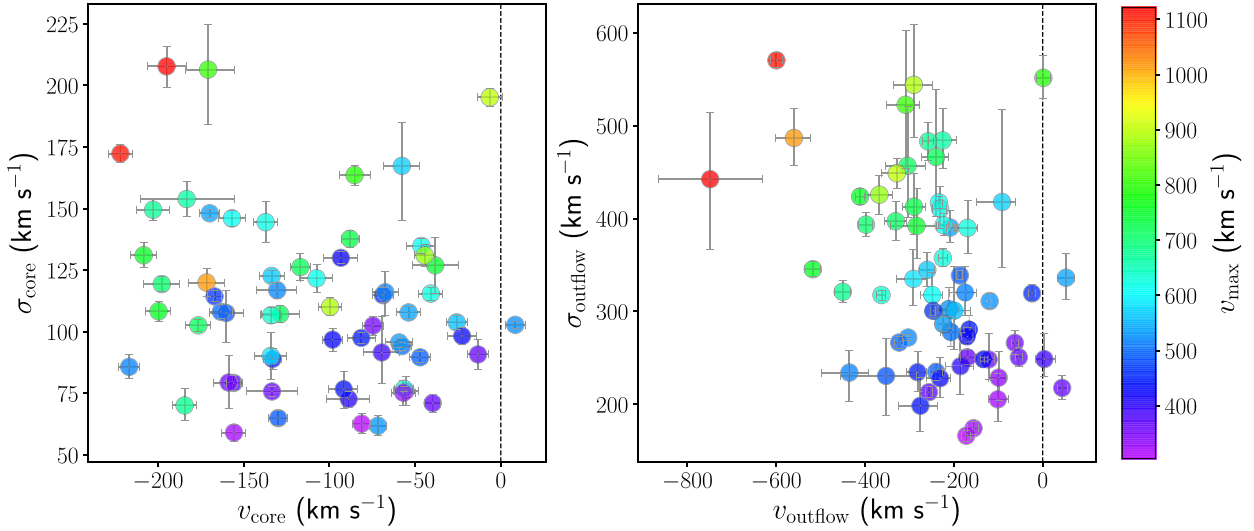


Figure 15. VVD diagrams for the [O III] core and outflow components. The black dashed line indicates zero velocity offset with respect to the systemic (stellar) velocity. *Left:* There is no correlation between σ_{core} and v_{core} , however the majority of velocities are blueshifted with respect to the systemic velocity. *Right:* There is a significant correlation between σ_{outflow} and v_{outflow} , characterized by the distinct ‘fan’ shape described by Woo et al. (2016), caused by increasing extinction with increasing outflow velocity.

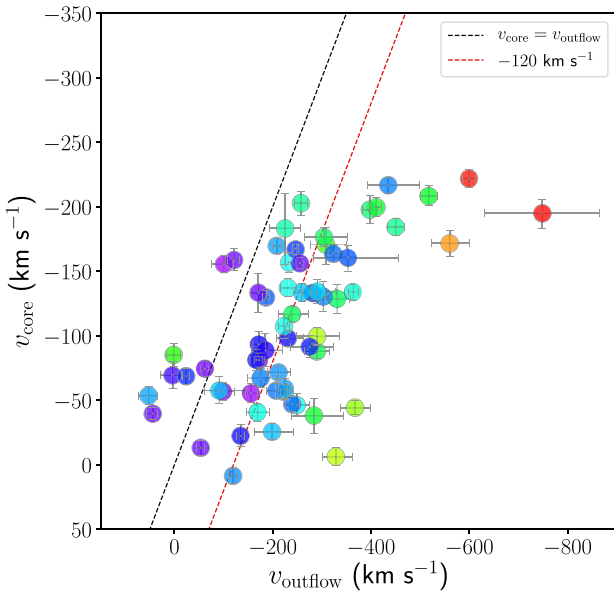


Figure 16. The strong correlation between v_{core} and v_{outflow} measured with respect to the systemic (stellar) velocity. The black dashed line indicates the perfect correlation $v_{\text{core}} = v_{\text{outflow}}$, and the red dashed line indicates an average -120 km s^{-1} offset from the perfect correlation for blueshifted outflows. Objects with larger v_{max} appear to deviate from this correlation, however a larger sample of objects with $v_{\text{max}} > 900 \text{ km s}^{-1}$ is needed to conclusively determine if the correlation holds true, and likewise for redshifted outflows.

Recall that 90 per cent of the objects with outflows were first identified visually from their strong asymmetric profile, implying that we are not overfitting [O III] profiles which do not require double-Gaussian decomposition. We therefore are confident that the $\sigma_{\text{core}} - \sigma_{\text{outflow}}$ correlation is real and not an artefact of the fitting process.

3.4.3 Correlations with luminosity

There is strong correlation between the AGN luminosity at $5100 \text{ \AA}$ ($L_{5100 \text{ \AA}}$) and L_{core} ($r_s = 0.77$), and a slightly weaker correlation for L_{outflow} ($r_s = 0.71$), although the weaker correlation with L_{outflow} is likely due to larger uncertainties. Because we are estimating $L_{5100 \text{ \AA}}$ using the luminosity of the broad H β emission line (Greene & Ho 2005), correlations with $L_{5100 \text{ \AA}}$ presented here are comparable to correlations with M_{BH} , which is estimated using both the luminosity and width of the broad H β emission line. However, the correlation between $L_{5100 \text{ \AA}}$ and outflow kinematics are comparatively weaker.

To investigate correlations with the radio luminosity at 1.4 GHz , we obtain $L_{1.4 \text{ GHz}}$ measurements from the VLA First Survey Catalog (White et al. 1997), which covers 10575 deg^2 of sky for a total of 946 432 radio sources, from which 18 of our 63 outflow objects have measurements. Referring to Fig. 14, the correlations between $L_{1.4 \text{ GHz}}$ and σ_{core} and σ_{outflow} are comparable to their correlations with $L_{5100 \text{ \AA}}$. However, when compared to systemic velocities, there is much stronger correlation between $L_{1.4 \text{ GHz}}$ and v_{outflow} ($r_s = -0.67$) than for v_{core} ($r_s = -0.36$). In Fig. 20, we plot $L_{1.4 \text{ GHz}}$ as a function of v_{core} and v_{outflow} .

The only other notable correlations found between luminosities are those with the luminosity of the broad Fe II. Both $L_{5100 \text{ \AA}}$ and $L_{1.4 \text{ GHz}}$ correlate strongly with $L_{\text{Br-Fe II}}$, and with nearly identical degrees of correlation of $r_s \sim 0.67$.

3.4.4 The $M_{\text{BH}} - \sigma_*$ relation

The ultimate goal of our analysis is to determine the effect – if any – of outflow kinematics on the $M_{\text{BH}} - \sigma_*$ relation. In Fig. 21, we plot the $M_{\text{BH}} - \sigma$ relation using both σ_* and σ_{core} , and plot the local relation derived from S19 (black dashed line) and the 0.43 dex scatter (orange dotted line) for comparison.

There is considerable scatter in the $M_{\text{BH}} - \sigma_*$ relation (left of Fig. 21) for our objects, however the majority of our sample falls within or close to the expected scatter of the relation, with the exception of some outliers above the relation by as much as ~ 1

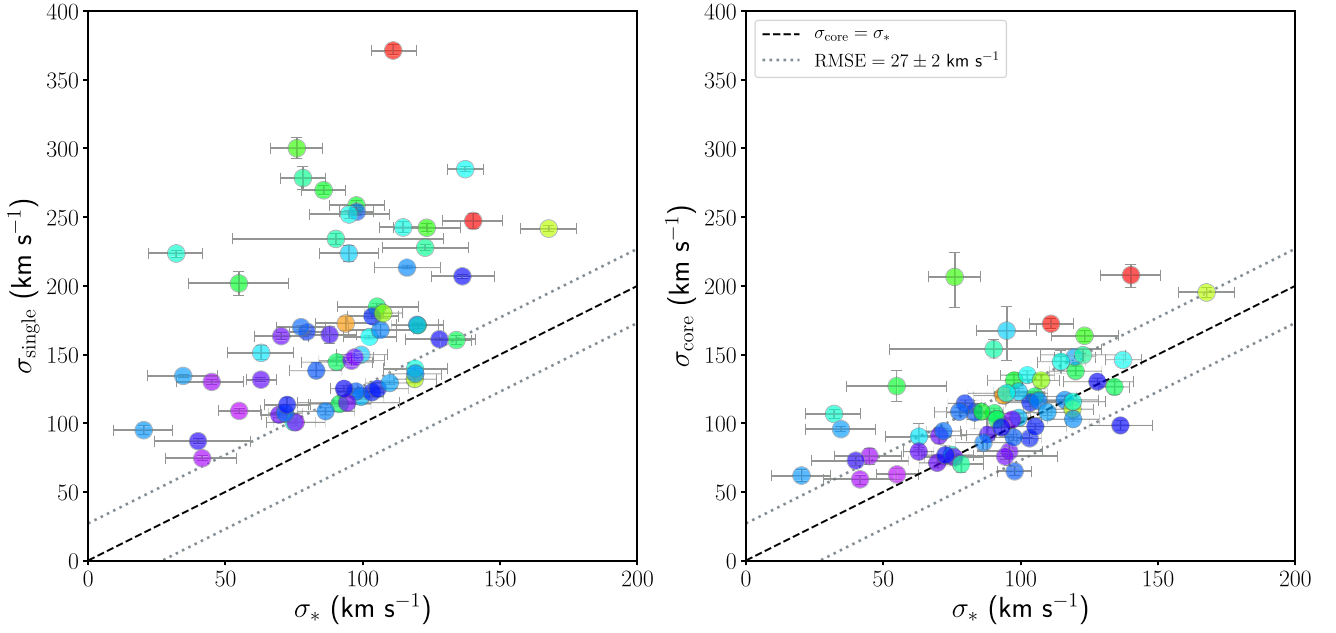


Figure 17. *Left:* The single Gaussian no-outflow model [O III] dispersion σ_{single} as a function of stellar velocity dispersion σ_* . There is a clear offset in objects which exhibit strong outflows with large v_{max} . *Right:* The double-Gaussian outflow model [O III] dispersion as a function of σ_* . The grey dotted lines in both plots give the scatter of the $\sigma_{\text{core}} - \sigma_*$ relation for comparison. The colourscale is the same as in Fig. 15. The black dashed line represents perfect correlation, i.e. $\sigma_{\text{core}} = \sigma_*$.

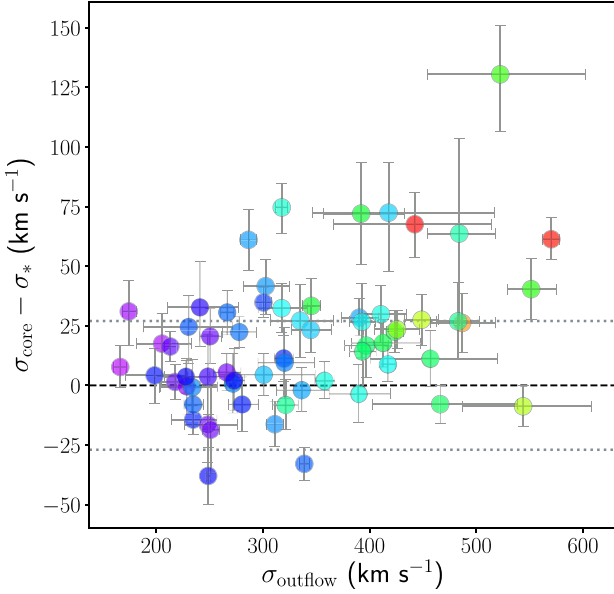


Figure 18. The difference between the decomposed σ_{core} and σ_* as a function of σ_{outflow} . There is a clear dependence on how well σ_{core} traces the gravitational component σ_* , which appears to scale with σ_{outflow} .

dex. The total scatter about the relation for our objects is 0.6 dex. There are considerable uncertainties we cannot account for given the nature of SDSS data that may affect our measurements of σ_* . Despite our efforts to correct for the effects of inclination, it is possible that the bulge + disc decomposition performed by Simard et al. (2011) resulted in a poor match to the image point spread function (PSF), since the decompositions do not take into account the PSF of the AGN, which would in turn affect measured disc quantities such as

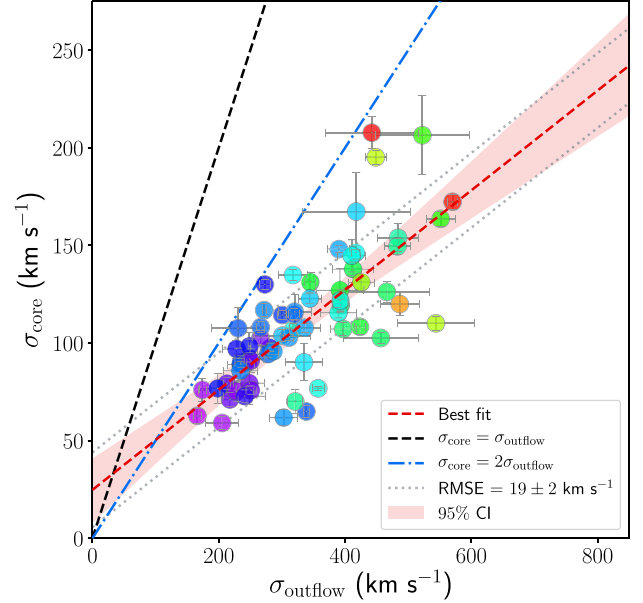


Figure 19. The correlation between σ_{core} and σ_{outflow} . The best-fitting regression line is given by the red dashed line, and the shaded red region corresponds to 95 per cent confidence interval. The grey dotted lines correspond to the scatter in the relation. The identity correlation ($\sigma_{\text{core}} = \sigma_{\text{outflow}}$) is shown by the black dashed line, and the $\sigma_{\text{core}} = 2\sigma_{\text{outflow}}$ relation is shown by the blue dash-dotted line for comparison.

ellipticity (b/a) and inclination. As mentioned in Section 3.3.2, the 3 arcsec diameter SDSS fibre can cover a significant fraction of the host galaxy to include contamination from non-bulge components, which can bias measurements of σ_* , but also decrease the fraction of light from the AGN. If the fraction of light from the AGN decreased

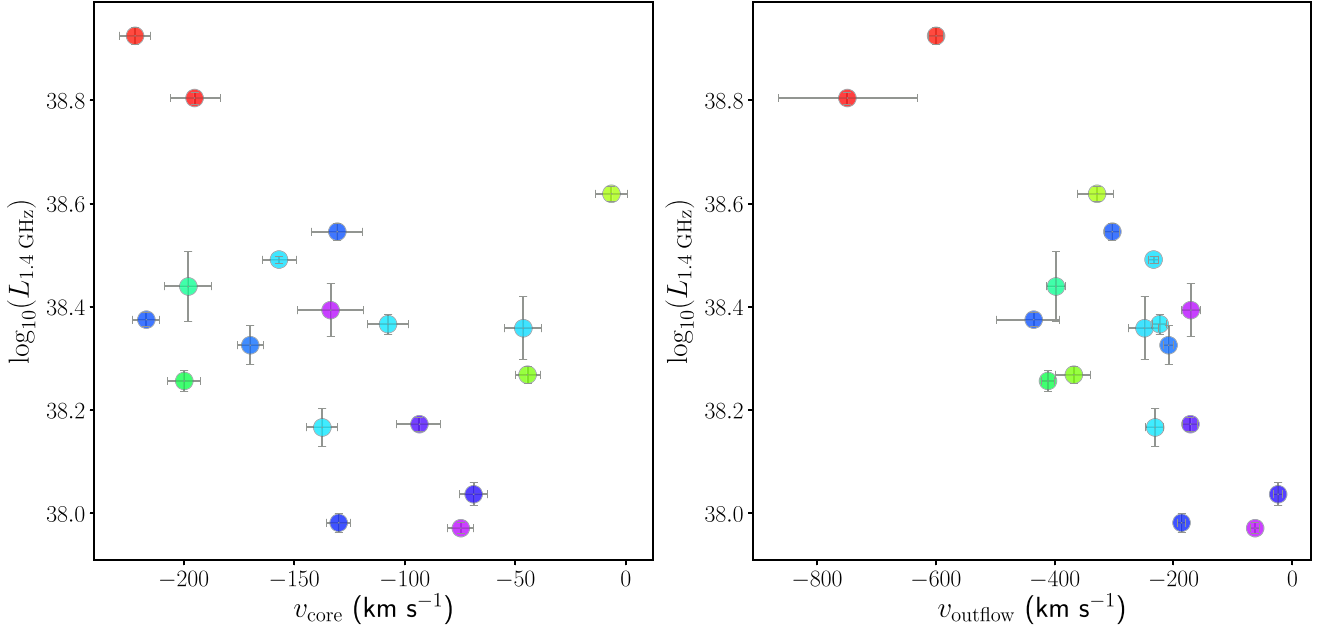


Figure 20. Correlations of $L_{1.4\text{GHz}}$ versus v_{core} (left) and v_{outflow} (right) for 18 objects with available measurements. The correlation with σ_{outflow} ($r_s = -0.67$) is nearly twice that of the correlation with v_{core} .

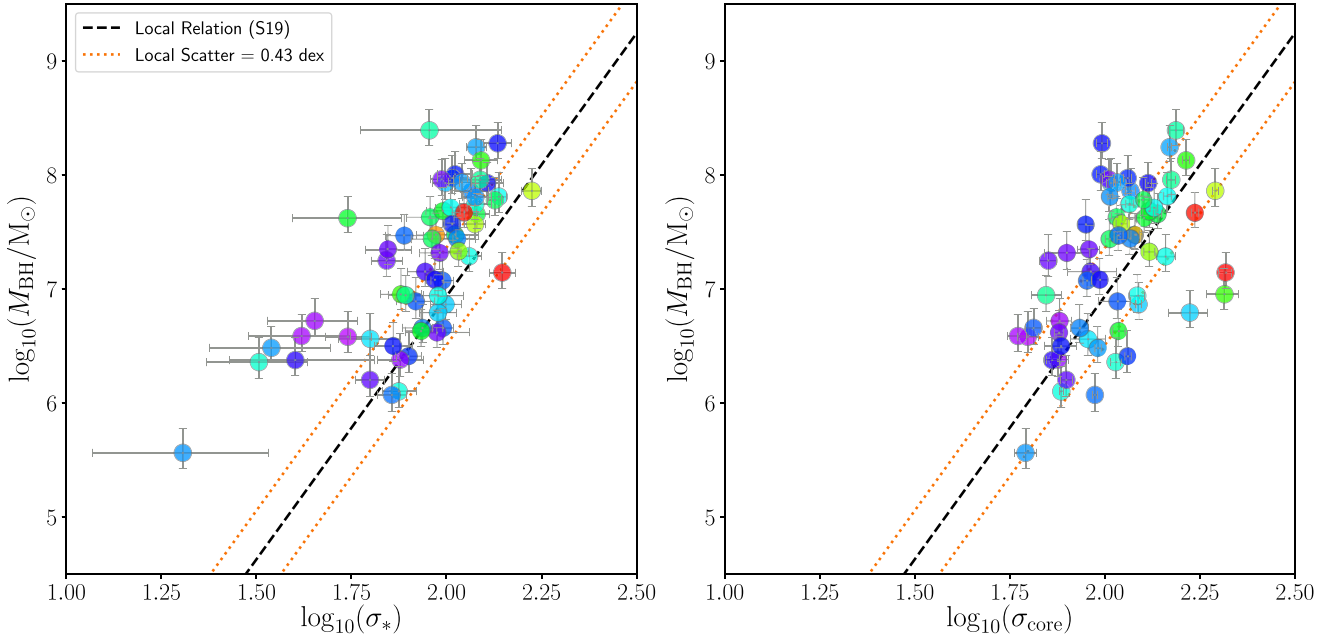


Figure 21. The $M_{\text{BH}}-\sigma$ relation as a function of σ_* (left) and σ_{core} (right). We plot the local relation calculated from S19 (black dashed line) as well as the local scatter (orange dotted lines) for comparison. The majority of objects on the $M_{\text{BH}}-\sigma_*$ relation agree with the local relation with some scatter ($f = 0.54$ dex), but the mean of the distribution is 0.57 dex above the local relation, caused by significant outliers likely due to poorly understood systematics and data quality. The $M_{\text{BH}}-\sigma_{\text{core}}$ relation also agrees with the local relation, with a comparable amount of scatter ($f = 0.58$ dex), mostly caused by stratification in σ_{core} , and a mean of only 0.17 dex above the local relation.

due to significant host galaxy absorption, we would underestimate the amplitude and therefore overestimate the FWHM of the broad H β emission line, leading to an overestimation of M_{BH} . Dependencies on AGN continuum dilution also play a role in how well σ_* can be recovered from absorption features, as we showed in Fig. 11. It remains that measurements σ_* are one of the most uncertain

measurements in BH scaling relations, due to both data limitations and poorly understood systematics.

Since we are interested in using σ_{core} as a surrogate for σ_* on the $M_{\text{BH}}-\sigma_*$ relation, we plot $M_{\text{BH}}-\sigma_{\text{core}}$ on the right in Fig. 21. We find that the scatter of the $M_{\text{BH}}-\sigma_*$ relation is 0.54 dex, slightly less than that of the $M_{\text{BH}}-\sigma_{\text{core}}$ relation with a scatter of 0.57 dex. However,

the mean of $M_{\text{BH}}-\sigma_*$ relation is 0.57 dex above the local relation, driven by clear outliers between 1 and 2 dex above the local relation. The mean of the $M_{\text{BH}}-\sigma_{\text{core}}$ is 0.17 and more-evenly distributed about the local relation.

It is clear from Fig. 21 that the scatter in $M_{\text{BH}}-\sigma_{\text{core}}$ is due primarily to stratification in σ_{core} , with $\sigma_{\text{core}} < 600 \text{ km s}^{-1}$ primarily above the relation, and $\sigma_{\text{core}} > 900 \text{ km s}^{-1}$ below the relation. It is possible that this separation in dispersion across the local relation could be attributed to the core broadening as a function of v_{max} we see in Fig. 18. It is also worthy to note that this stratification in σ_* is not as obvious in the $M_{\text{BH}}-\sigma_*$ relation, although there is similar trend for $\sigma_* < 600 \text{ km s}^{-1}$.

3.5 Discussion

In the following sections, we discuss the different correlations and their possible interpretation. We emphasize that although we can only speculate on the physical interpretation of these correlations, these observations represent observational constraints that should be considered when developing models that describe AGN-driven outflows.

3.5.1 Correlations with velocity

The differences in the VVD diagrams shown in Fig. 15 indicate that the core and outflow components of the NLR are kinematically distinct. According 3D biconical outflow modelling from Bae & Woo (2016), the fan-shaped distribution of the $\sigma_{\text{outflow}}-v_{\text{outflow}}$ relation is due to a number of factors, the most important of which are bicone inclination, ejection velocity, and dust extinction along the line of sight. For a symmetric biconical outflow in the absence of any dust extinction, we would expect to measure zero velocity offset along line of sight due to the cancelling of velocities in opposite directions. With the addition of extinction effects, the obscuration of one side of the bicone would lead to a shift in observed velocity offset. The large number of blueshifted outflows in our sample can be explained as varying obscuration of the receding (redshifted) cone. One interpretation of the strong correlation in the outflow VVD diagram is evidence of collimation, that is, we expect to see an increase in σ_{outflow} with an increase in v_{outflow} along the line of sight if the flow subtends relatively small solid angle (such as a cone) and has a preferred inclination. For example, we expect to see larger velocities as well as a larger velocity dispersion for a flow that is directed along the LOS, as opposed to a flow directed at some angle with respect to the LOS, which would produce a smaller observed velocity and thus smaller dispersion. This interpretation is consistent with the model grids for 3D biconical outflow models from Bae & Woo (2016).

We do not observe the same strong correlation for the $\sigma_{\text{core}}-v_{\text{core}}$ VVD relation. There is an overall blueshift of v_{core} which correlates with v_{outflow} , as shown in Fig. 16, however there is a larger spread in v_{core} for a given value of σ_{core} . Given our above argument for the outflow component, the lack of correlation for the core component would imply that the source of the core gas emission is less collimated and more spherically symmetric. This would agree with the interpretation that the core component represents the original NLR gas that is still strongly coupled to the gravitational potential. The core VVD diagram clearly does not exhibit the same kinematic properties of outflows, and should be treated as separate kinematic component when trying to model outflows in AGN.

The linear increase between v_{core} and v_{outflow} and constant 120 km s^{-1} offset in velocity shown in Fig. 16 could indicate that the two

components are locked in velocity, at least up to a certain value of v_{max} . We can only speculate the physical interpretation of this trend, but it could mean that the core NLR gas can be coupled to the outflowing gas below a certain velocity threshold, causing it to become entrained and expand with the outflow. At the highest velocities, the core gas may decouple from the outflowing gas, causing this trend to plateau as shown in Fig. 16. We would require a larger sample of objects with outflows with $v_{\text{max}} > 900 \text{ km s}^{-1}$ to determine if this occurs.

3.5.2 Correlations with dispersion

In an idealized gravitationally bound system, such as in an undisturbed NLR, we should expect the gas and stellar components to have similar velocity distributions. If a secondary component in the same region as the source of NLR emission is present, and exhibits some collimation (increase in velocity and velocity dispersion), we would expect a distribution similar to that shown on the left of Fig. 17. We do recover the core NLR gas within some scatter about the perfect correlation with σ_* after correction, as shown on the right in Fig. 17. Objects with $v_{\text{max}} < 600 \text{ km s}^{-1}$ are more evenly distributed about the perfect correlation with σ_* after correcting for σ_{outflow} . The majority of objects with $v_{\text{max}} > 600 \text{ km s}^{-1}$ tend to fall above with the relation even after correcting for σ_{outflow} , which may indicate that the presence of outflows may introduce additional non-gravitational broadening which may only be detected for the strongest cases. Some scatter is expected, as we cannot fully account for all non-gravitational interactions nor fully account for systematics involving the measurements of σ_* given the nature of SDSS data, such as inclination, aperture effects, or merger history. This correlation is enough to suggest that the core component of the [O III] profile traces the original NLR gas that is dominated by the gravitational potential of the stellar component.

It is not unreasonable to suggest that if a secondary outflowing component arises from within the NLR, it must start out with the same velocity distribution as the core components. If that velocity distribution then undergoes some interaction, we expect the original distribution to broaden. We can interpret the $\sigma_{\text{core}}-\sigma_{\text{outflow}}$ correlation shown in Fig. 19 to be the broadening of the original NLR core gas due to the outflowing gas. What is still puzzling is the linear rate at which the outflow dispersion grows with the core dispersion and its small scatter. We can interpret this as the outflow component having a strong dependence on the original NLR gas from which it is believed to have originated, and the strong linear dependence describes the manner by which the flow propagates through the ambient medium. Another possible interpretation is that the strength of the outflow component (parametrized by v_{max}) causes a broadening of the core component, such that σ_{core} approaches its respective value of σ_* . There is some correlation shown in Fig. 18 that suggests that the core component broadens as a function of σ_{outflow} (and therefore v_{max}), however a larger sample of objects with $v_{\text{max}} > 900$ is needed to determine if this trend is real or simply increased scatter.

The $\sigma_{\text{core}}-\sigma_*$ and $\sigma_{\text{core}}-\sigma_{\text{outflow}}$ imply that there is some connection between σ_* and σ_{outflow} . Ideally, if there is a constant linear relationship between σ_{core} and σ_{outflow} , and if σ_{core} traces σ_* , then σ_{outflow} should also scale with σ_* but positively offset by some constant. We can fit the interdependence of the three dispersions as a plane of the form

$$a \log_{10}(\sigma_*) + b \log_{10}(\sigma_{\text{core}}) + c \log_{10}(\sigma_{\text{outflow}}) + d = 0. \quad (12)$$

We perform orthogonal regression using EMCEE following the methods of S19 and obtain best-fitting coefficients of $a = -2.39^{+0.54}_{-0.51}$,

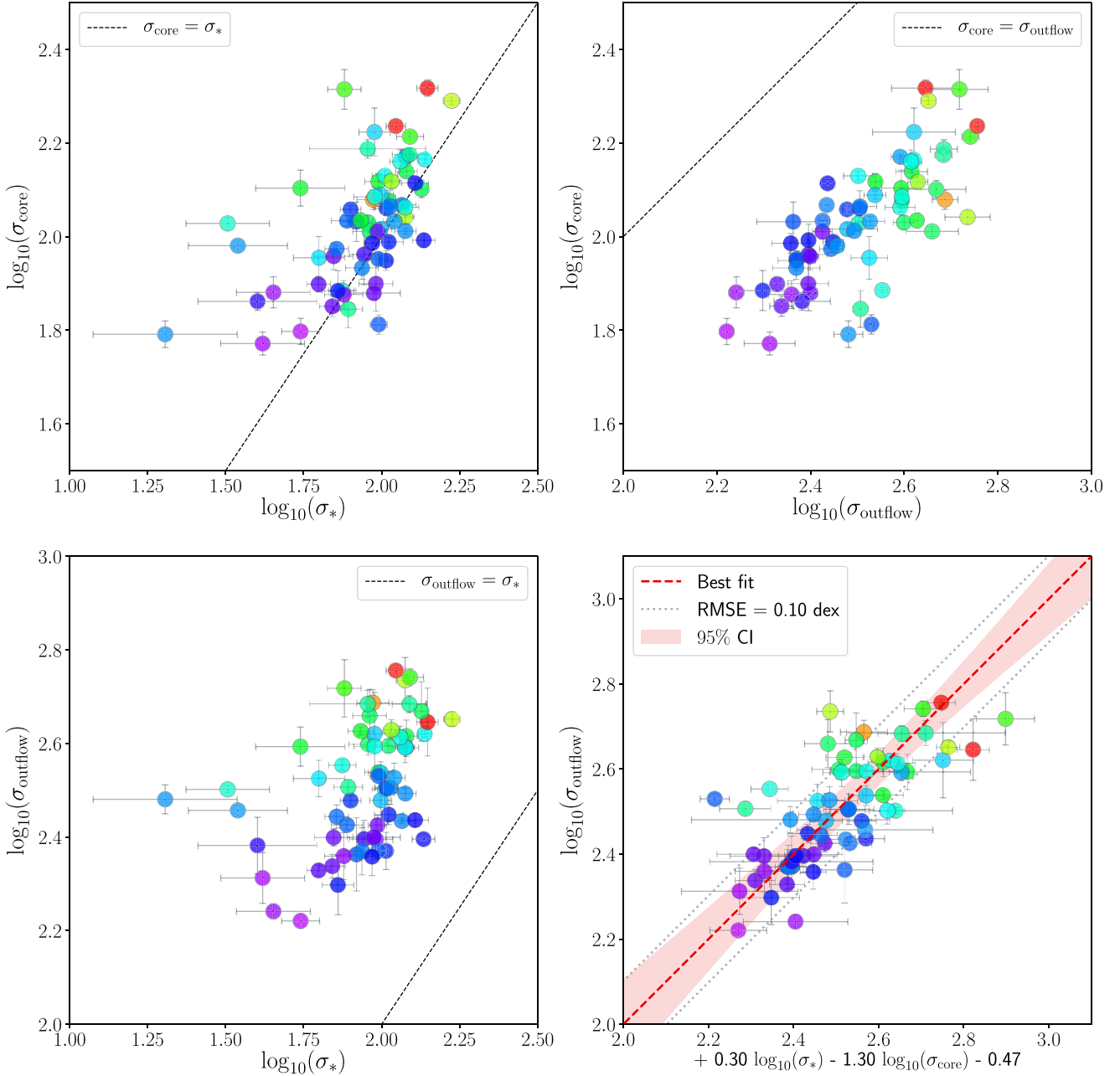


Figure 22. The three projections of the $\sigma_{\text{outflow}}-\sigma_{\text{core}}-\sigma_*$ relation, and the best-fitting relation projected parallel to the best-fitting plane. The identity correlations are given by the black dashed line in each plot. The scatter about the best-fitting plane relation is $f = 0.10$, which is considerably smaller than the scatter in the $\sigma_{\text{core}}-\sigma_*$ ($f = 0.17$ dex) and $\sigma_{\text{outflow}}-\sigma_*$ ($f = 0.19$ dex) relations, and comparable to the $\sigma_{\text{core}}-\sigma_{\text{outflow}}$ relation ($f = 0.09$ dex).

$b = 10.26^{+1.00}_{-1.19}$ and $c = -7.91^{+1.12}_{-0.96}$, $d = 3.63^{+1.46}_{-1.60}$, and a scatter about the best-fitting plane of $f = 0.10$ dex. We plot the projections of the three dispersions, and the projection along the parallel axis of the plane in Fig. 22. Despite the decreased scatter, there is still a large uncertainty in a , i.e. the slope of the $\sigma_{\text{outflow}}-\sigma_*$ relation, which is caused by large scatter. Further study with a larger sample is needed to better constrain this slope before the functional form of 12 can be used to calculate σ_* using both σ_{core} and σ_{outflow} .

The physical interpretation of the plane relationship between the three dispersions does not necessarily imply that σ_{outflow} can somehow influence σ_* or vice versa, neither does it answer the

proverbial ‘chicken or egg’ problem, that is, we do not know if outflows are the causal explanation for the broadening of σ_{core} or if σ_{outflow} correlates with σ_{core} because it originated from an already-broad gas velocity distribution. A larger sample, along with integral field spectroscopy, to determine if these relationships hold true.

3.5.3 Correlations with luminosity

It has been known for some time that the incidence of [O III] outflows correlates with radio emission in both type 1 and type 2 AGNs (Wilson & Willis 1980; Whittle 1985; Whittle et al. 1988; Nelson

& Whittle 1996). More recent studies suggest that the strongest correlation with luminosity is between the [O III] width and the radio luminosity at 1.4 GHz ($L_{1.4\text{GHz}}$), especially in objects with high-velocity outflows and at much higher redshifts. (Mullaney et al. 2013; Zakamska & Greene 2014; Zakamska et al. 2016; Hwang et al. 2018; Perrotta et al. 2019).

As mentioned in Section 3.4.3, the strong correlation between the core and outflow components, L_{core} and L_{outflow} , and the optical AGN luminosity $L_{5100\text{\AA}}$ is not surprising if the core and outflow components originate in close proximity to the ionizing source. There is still some correlation with core and outflow dispersion, but even lesser so for the core and outflow velocities. We see similar lack of correlation when we compare σ_{core} and v_{core} to $L_{1.4\text{GHz}}$. By far, the strongest correlation we find between any measured luminosities and kinematics is with $L_{1.4\text{GHz}}$ and v_{outflow} .

Previous studies by Woo et al. (2016) and Rakshit & Woo (2018) used a total (core + outflow) integrated [O III] velocity dispersion parametrization and normalized it by the stellar velocity dispersion to quantify non-gravitational kinematics to compare to radio luminosity, finding no strong correlations with radio activity. In this study, the outflow component is designated as the only non-gravitational component, for which we do find strong correlation with radio luminosity, although with a much smaller sample size. Our findings agree with Mullaney et al. (2013), who similarly found strong correlation between objects with high $L_{1.4\text{GHz}}$ and objects with the broadest [O III] profiles.

3.5.4 The $M_{\text{BH}}-\sigma_*$ relation

Fig. 21 shows that when corrected for the outflow component, σ_{core} can be used as a surrogate for σ_* on the $M_{\text{BH}}-\sigma_*$ relation with comparable scatter, and agree with the results found by Bennert et al. (2018). However, if we are to use σ_{core} for studies on the non-local $M_{\text{BH}}-\sigma_*$, we do not have the luxury of comparing it σ_* to ensure we have evidence of non-gravitational kinematics as we have done in our sample. Performing a double-Gaussian decomposition of the [O III] profile when there is no evidence of an additional non-gravitational component, while always producing a better fit, can cause one to measure a smaller σ_{core} than what σ_* suggests, which can give the impression that one is measuring BHs that are overmassive relative to the local $M_{\text{BH}}-\sigma_*$ relation.

We advise that if σ_{core} is used as a surrogate for σ_* , that one always fit a double-Gaussian component and check that the object falls within the acceptable scatter of the $\sigma_{\text{core}}-\sigma_{\text{outflow}}$ relation. Furthermore, for $\sigma_{\text{outflow}} < 200 \text{ km s}^{-1}$, the scatter of the $\sigma_{\text{core}}-\sigma_{\text{outflow}}$ relation begins to intersect with that of the $\sigma_{\text{outflow}}-\sigma_*$ relation, and it becomes increasingly unclear if there are additional non-gravitational kinematics present in the [O III] profile with respect to σ_* . Therefore, we recommend that for $\sigma_{\text{outflow}} < 200 \text{ km s}^{-1}$ one does *not* use a double-Gaussian decomposition for the risk of severely overfitting the [O III] profile and significantly underestimating σ_* . Likewise, if a single-Gaussian fit to the [O III] profile exceeds $\sim 200 \text{ km s}^{-1}$, it is recommended to perform a double-Gaussian decomposition and assess the quality of the fit. In this regard, the outflow confidence calculated by our algorithm by performing an F -statistic model comparison makes it clear when a double-Gaussian fit is warranted by the data.

4 CONCLUSION

To summarize, a new, thoroughly tested, and powerful fitting software for optical SDSS spectra that is open source and specialized

for fitting AGN spectra. Since our algorithm can fit numerous components simultaneously, it can be generalized to fit not just AGN spectra, but non-AGN host galaxies as well. The use of MCMC allows the user to fit objects with unprecedented detail, obtain robust uncertainties, and determine the quality of fits using a broad range of metrics and outputs. Our code also utilizes multiprocessing to efficiently fit large samples of objects without excessive memory overhead.

Currently, our code is being used for a variety of research projects and number of collaborations. For instance, our code is being run in a cluster environment to fit over 19 000 SDSS galaxies to determine the significance of outflows as a function of separation distance and as a function of environment. Our code will also be used to fit a larger sample of type 1 AGN to follow-up on results presented here.

Performance tests with our algorithm we have presented here in the recovery of σ_* can also be applied to other fitting routines which attempt to measure the LOSVD. We summarize the results of these tests below:

(i) In non-AGN host galaxies and type 2 AGN, where significant Fe II and AGN continuum dilution is absent, the LOSVD can be recovered in the Mg IB/Fe II region (4400–5800 Å) with less than 10 per cent error and uncertainty for S/N > 20. For objects which exhibit active star formation, the steep continuum from young stellar populations complicates measurements of σ_* , and we recommend to include more O- and B-type template stars and disable the AGN power-law component from the fit.

(ii) Measurements of σ_* are not significantly affected by the inclusion of Fe II, and are more affected by S/N level.

(iii) Dilution of stellar absorption features by strong AGN continuum contributes to the largest error and uncertainty in measuring σ_* . We find that while that strong Fe II and steep AGN power-law slope can be indicative of strong continuum dilution, they are not the root cause. Continuum dilution caused to a large fraction of continuum flux being dominated by the AGN is the root cause of large uncertainties in the estimation of σ_* , and extra caution should be given in the estimate of the LOSVD to objects which exhibit strong Fe II or steep power-law slope, such as NLS1 or BAL objects.

As an application of our algorithm, we fit a sample of 63 SDSS type 1 AGN with strong evidence of outflows in the [O III] $\lambda 5007$ emission line and performed a correlation analysis of kinematics to determine the relationships between outflows, the AGN, and the host galaxy, expanding upon previous similar studies. We summarize our most important results below:

(i) By performing a double-Gaussian decomposition of the [O III] $\lambda 5007$ emission line profile into separate core and outflow components, we find that the core dispersion of the [O III] profile (σ_{core}) is a suitable surrogate for stellar velocity dispersion (σ_*) in a statistical context but should not be used on an object-to-object basis. There is some evidence that the measured difference $\sigma_{\text{core}} - \sigma_*$ scales with increasing outflow component dispersion (σ_{outflow}), which may imply that there is some broadening of the NLR gas due to the presence of outflows, causing the scatter we see in the $\sigma_{\text{core}}-\sigma_*$ relation.

(ii) VVD diagrams of the outflow component resemble the ‘fan-shaped’ VVD profiles exhibited by 3D biconical outflow models from Bae & Woo (2016), indicating possible orientation-dependent or collimated flow. The core component does not exhibit the same VVD shape as the outflow component, indicating that it is a kinematically distinct component of the [O III] gas, more strongly coupled to the gravitational potential.

(iii) There is a systematic broadening of the σ_{core} component which scales with σ_{outflow} , resulting a tight correlation between σ_{core} and σ_{outflow} . This tight correlation implies a very specific relationship between outflow kinematics and the kinematics of the NLR, which could be used to constrain theoretical models of AGN outflows.

(iv) We present a new planar relationship between σ_* , σ_{core} , and σ_{outflow} with a scatter about the best-fitting plane of 0.10 dex. However, a larger sample is still needed to constrain the relationship between σ_{outflow} and σ_* before it can be used to obtain values for σ_* .

(v) We recover the strong correlation between $L_{1.4\text{ GHz}}$ and properties of outflows found in previous studies. We do not observe strong correlations between outflow kinematics and the optical AGN luminosity $L_{5100\text{ Å}}$.

(vi) We find that σ_{core} is a suitable surrogate for σ_* on the $M_{\text{BH}}-\sigma_*$ relation with comparable scatter in a statistical context in agreement with Bennert et al. (2018). Additionally, we present recommendations and caveats for using σ_{core} for studies of the $M_{\text{BH}}-\sigma_*$ relation in the non-local universe for which σ_* cannot be measured.

The correlations we have presented here showcase a number of observational constraints that theoretical models of AGN outflows should satisfy. Further investigation into these correlations and their causes will be necessary with larger samples, and we have shown here that our algorithm is capable of such detailed analyses. As newer and larger surveys begin to come online, tools such as our code, which underscore the need for a generalized open-source framework for fitting a variety of objects with advanced statistical techniques, will be needed for increasingly detailed analysis of astronomical spectra in the coming decade.

ACKNOWLEDGEMENTS

We thank the anonymous referee for their careful reading of our manuscript, as well as their helpful suggestions in improving the manuscript and code.

ROS acknowledges financial support from the NASA MIRO program through the Fellowships and Internships for Extremely Large Data Sets (FIELDS) in the form of a Graduate Student Fellowship. ROS personally thanks Dr. Bahram Mobasher for his generous support as a FIELDS graduate student fellow.

Partial support for this project was provided by the National Science Foundation, under grant no. AST 1817233.

Funding for SDSS-III has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, and the U.S. Department of Energy Office of Science. The SDSS-III web site is <http://www.sdss3.org/>.

SDSS-III is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS-III Collaboration including the University of Arizona, the Brazilian Participation Group, Brookhaven National Laboratory, Carnegie Mellon University, University of Florida, the French Participation Group, the German Participation Group, Harvard University, the Instituto de Astrofísica de Canarias, the Michigan State/Notre Dame/JINA Participation Group, Johns Hopkins University, Lawrence Berkeley National Laboratory, Max Planck Institute for Astrophysics, Max Planck Institute for Extraterrestrial Physics, New Mexico State University, New York University, Ohio State University, Pennsylvania State University, University of Portsmouth, Princeton University, the Spanish Participation Group, University of Tokyo, University of Utah, Vanderbilt University, University of Virginia, University of Washington, and Yale University.

DATA AVAILABILITY STATEMENT

The code referenced throughout this manuscript is available at <https://github.com/remingtonsexton> for both PYTHON 2.7 and PYTHON 3.6. The spectral data used in the preparation of this manuscript is publicly available at <http://www.sdss3.org/>.

REFERENCES

- Afanasev V. L., Lipovetskii V. A., Markarian B. E., Stepanian D. A., 1980, *Astrofizika*, 16, 193
- Astropy Collaboration, 2013, *A&A*, 558, A33
- Bae H.-J., Woo J.-H., 2016, *ApJ*, 828, 97
- Bae H.-J., Woo J.-H., Karouzos M., Gallo E., Flohic H., Shen Y., Yoon S.-J., 2017, *ApJ*, 837, 91
- Barth A. J. et al., 2013, *ApJ*, 769, 128
- Beifiori A., Maraston C., Thomas D., Johansson J., 2011, *A&A*, 531, A109
- Bellovary J. M., Holley-Bockelmann K., Gültekin K., Christensen C. R., Governato F., Brooks A. M., Loebman S., Munshi F., 2014, *MNRAS*, 445, 2667
- Bennert V. N. et al., 2018, *MNRAS*, 481, 138
- Berton M., Björklund I., Lähteenmäki A., Congiu E., Järvelä E., Terreran G., La Mura G., 2020, *Contrib. Astron. Obs. Skalnaté Pleso*, 50, 270
- Boroson T. A., 2003, *ApJ*, 585, 647
- Boroson T. A., Meyers K. A., 1992, *ApJ*, 397, 442
- Brotherton M. S., 1996, *ApJS*, 102, 1
- Calderone G., Nicastro L., Ghisellini G., Dotti M., Sbarrato T., Shankar F., Colpi M., 2017, *MNRAS*, 472, 4051
- Cappellari M., 2017, *MNRAS*, 466, 798
- Cappellari M., Emsellem E., 2004, *PASP*, 116, 138
- Cicone C., Maiolino R., Marconi A., 2016, *A&A*, 588, A41
- Costa T., Pakmor R., Springel V., 2020, *MNRAS*, 497, 5229
- Croton D. J. et al., 2006, *MNRAS*, 365, 11
- Davies R. L. et al., 2019, *ApJ*, 873, 122
- Davies R. L. et al., 2020, *ApJ*, 894, 28
- DeGraf C., Di Matteo T., Treu T., Feng Y., Woo J. H., Park D., 2015, *MNRAS*, 454, 913
- De Robertis M. M., Osterbrock D. E., 1984, *ApJ*, 286, 171
- DiPompeo M. A., Hickox R. C., Carroll C. M., Runnoe J. C., Mullaney J. R., Fischer T. C., 2018, *ApJ*, 856, 76
- Dubois Y., Gavazzi R., Peirani S., Silk J., 2013, *MNRAS*, 433, 3297
- Fabian A. C., 2012, *ARA&A*, 50, 455
- Falcón-Barroso J. et al., 2017, *A&A*, 597, A48
- Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, *PASP*, 125, 306
- Freitas I. C. et al., 2018, *MNRAS*, 476, 2760
- Grandi S. A., 1977, *Lick Obs. Bull.*, 754, 1
- Greene J. E., Ho L. C., 2005, *ApJ*, 627, 721
- Greene J. E., Ho L. C., 2006, *ApJ*, 641, 117
- Guo H., Shen Y., Wang S., 2018, *Astrophysics Source Code Library*, record ascl:1809.008
- Hall M., Courteau S., Dutton A. A., McDonald M., Zhu Y., 2012, *MNRAS*, 425, 2741
- Harrison C. M., Alexander D. M., Mullaney J. R., Swinbank A. M., 2014, *MNRAS*, 441, 3306
- Hartmann M., Debattista V. P., Cole D. R., Valluri M., Widrow L. M., Shen J., 2014, *MNRAS*, 441, 1243
- Heckman T. M., Miley G. K., van Breugel W. J. M., Butcher H. R., 1980, *BAAS*, 12, 809
- Heckman T. M., Miley G. K., Green R. F., 1984, *ApJ*, 281, 525
- Hwang H.-C., Zakamska N. L., Alexandroff R. M., Hamann F., Greene J. E., Perrotta S., Richards G. T., 2018, *MNRAS*, 477, 830
- Ivezić Ž. et al., 2019, *ApJ*, 873, 111
- Jackson N., Browne I. W. A., 1991, *MNRAS*, 250, 414
- King A., Pounds K., 2015, *ARA&A*, 53, 115

- Kluyver T. et al., 2016, in Loizides F., Schmidt B., eds, Positioning and Power in Academic Publishing: Players, Agents and Agendas. IOS Press, the Netherlands, p. 87
- Komossa S., Xu D., 2007, *ApJ*, 667, L33
- Komossa S., Xu D. W., Wagner A. Y., 2018, *MNRAS*, 477, 5115
- Kormendy J., Ho L. C., 2013, *ARA&A*, 51, 511
- Kovačević J., Popović L. Č., Dimitrijević M. S., 2010, *ApJS*, 189, 15
- Kraft D., 1988, Technical Report DFVLR-FB 88-28, Institut für Dynamik der Flugsysteme. Oberpfaffenhofen
- Leighly K. M., 1999, *ApJS*, 125, 317
- Malkan M. A., 1983, *ApJ*, 268, 582
- Manzano-King C. M., Canalizo G., Sales L. V., 2019, *ApJ*, 884, 54
- Maraston C., Strömbäck G., Thomas D., Wake D. A., Nichol R. C., 2009, *MNRAS*, 394, L107
- Melioli C., de Gouveia Dal Pino E. M., 2015, *ApJ*, 812, 90
- Moran E. C., Halpern J. P., Helfand D. J., 1996, *ApJS*, 106, 341
- Mullaney J. R., Alexander D. M., Fine S., Goulding A. D., Harrison C. M., Hickox R. C., 2013, *MNRAS*, 433, 622
- Müller-Sánchez F., Prieto M. A., Hicks E. K. S., Vives-Arias H., Davies R. I., Malkan M., Tacconi L. J., Genzel R., 2011, *ApJ*, 739, 69
- Nelson C. H., Whittle M., 1996, *ApJ*, 465, 96
- Perrotta S., Hamann F., Zakamska N. L., Alexand roff R. M., Rupke D., Wylezalek D., 2019, *MNRAS*, 488, 4126
- Rakshit S., Woo J.-H., 2018, *ApJ*, 865, 5
- Rakshit S., Stalin C. S., Chand H., Zhang X.-G., 2017, *ApJS*, 229, 39
- Sarzi M. et al., 2006, *MNRAS*, 366, 1151
- Sexton R. O., Canalizo G., Hiner K. D., Komossa S., Woo J.-H., Treister E., Hiner Dimassimo S. L., 2019, *ApJ*, 878, 101
- Simard L., Trevor Mendel J., Patton D. R., Ellison S. L., McConnachie A. W., 2011, *ApJS*, 196, 11
- Valdes F., 1984, *BAAS*, 16, 497
- Valdes F., Gupta R., Rose J. A., Singh H. P., Bell D. J., 2004, *ApJS*, 152, 251
- Vazdekis A., Sánchez-Blázquez P., Falcón-Barroso J., Cenarro A. J., Beasley M. A., Cardiel N., Gorgas J., Peletier R. F., 2010, *MNRAS*, 404, 1639
- Veilleux S., 1991, *ApJS*, 75, 383
- Véron-Cetty M. P., Véron P., Gonçalves A. C., 2001, *A&A*, 372, 730
- Véron-Cetty M. P., Joly M., Véron P., 2004, *A&A*, 417, 515
- Virtanen P. et al., 2020, *Nat. Methods*, 17, 261
- Wang T., Lu Y., 2001, *A&A*, 377, 52
- Wang J., Xu D. W., Wei J. Y., 2018, *ApJ*, 852, 26
- White R. L., Becker R. H., Helfand D. J., Gregg M. D., 1997, *ApJ*, 475, 479
- Whittle M., 1985, *MNRAS*, 213, 1
- Whittle M., Pedlar A., Meurs E. J. A., Unger S. W., Axon D. J., Ward M. J., 1988, *ApJ*, 326, 125
- Wilson A. S., Willis A. G., 1980, *ApJ*, 240, 429
- Woo J.-H., Treu T., Malkan M. A., Blandford R. D., 2006, *ApJ*, 645, 900
- Woo J.-H., Yoon Y., Park S., Park D., Kim S. C., 2015, *ApJ*, 801, 38
- Woo J.-H., Bae H.-J., Son D., Karouzos M., 2016, *ApJ*, 817, 108
- Wylezalek D., Flores A. M., Zakamska N. L., Greene J. E., Riffel R. A., 2020, *MNRAS*, 492, 4680
- Xu D., Komossa S., Zhou H., Lu H., Li C., Grupe D., Wang J., Yuan W., 2012, *AJ*, 143, 83
- York D. G. et al., 2000, *AJ*, 120, 1579
- Zakamska N. L., Greene J. E., 2014, *MNRAS*, 442, 784
- Zakamska N. L. et al., 2016, *MNRAS*, 459, 3144
- Zhang X.-G., Feng L.-L., 2017, *MNRAS*, 468, 620
- Zhang S., Wang T.-G., Wang H., Zhou H., Dong X.-B., Wang J.-G., 2010, *ApJ*, 714, 367

This paper has been typeset from a $\text{\LaTeX}$ file prepared by the author.