Gas Content Regulates the Life Cycle of Star Formation and Black Hole Accretion in Galaxies

Hassen M. Yesuf$^{1,2}$ and Luis C. Ho$^{1,3}$

$^1$Kavli Institute for Astronomy and Astrophysics, Peking University, Beijing 100871, People’s Republic of China
$^2$Kavli Institute for the Physics and Mathematics of the Universe, The University of Tokyo, Kashiwa, 277-8583, Japan
$^3$Department of Astronomy, School of Physics, Peking University, Beijing 100871, People’s Republic of China

Received 2020 February 15; revised 2020 July 23; accepted 2020 July 24; published 2020 September 18

Abstract

Feedback from active galactic nuclei (AGNs) is expected to impact the amount of cold gas in galaxies by driving strong galactic winds, by preventing external gas inflows, or by changing the thermodynamical state of the gas. We use estimates of molecular gas mass based on dust absorption ($\text{H}\alpha$/H$\beta$) to study gas content of large samples of type 2 AGN host galaxies in comparison with inactive galaxies. Using sparse principal component and clustering analysis, we analyze a suite of stellar and structural parameters of $\sim$27,100 face-on, central galaxies at redshift $z = 0.02–0.15$ and with stellar mass $M_\star \approx 10^{10}–2 \times 10^{11} M_\odot$. We identify four galaxy groups of similar mass and morphology (mass surface density, velocity dispersion, concentration, and Sérsic index) that can be evolutionarily linked through a life cycle wherein gas content mediates their star formation rate (SFR) and level of AGN activity. Galaxies first consume their gas mostly through bursty star formation, then enter into a transition phase of intermediate gas richness in which star formation and AGNs coexist, before settling into retirement as gas-poor, quiescent systems with residual levels of AGN activity (LINERs). Strongly accreting black holes (Seyferts) live in gas-rich, star-forming hosts, but neither their gas reservoir nor their ability to form stars seems to be impacted instantaneously (timescales $\lesssim 0.5$ Gyr) by AGN feedback. Our results are inconsistent with AGN feedback models that predict that central, bulge-dominated, Seyfert-like AGNs in massive galaxies have significantly lower molecular gas fractions than inactive galaxies of similar mass, morphology, and SFR.

Unified Astronomy Thesaurus concepts: Active galactic nuclei (16); Interstellar dust extinction (837); Molecular gas (1073); Galaxy evolution (594); Seyfert galaxies (1447); Supermassive black holes (1663); AGN host galaxies (2017)

1. Introduction

There are two main types of galaxies in the local universe: star-forming, gas-rich spiral galaxies, and quiescent, gas-poor elliptical and lenticular galaxies. Nearly all massive local galaxies harbor supermassive black holes ($\sim 10^8–10^{10} M_\odot$) at their centers. Through feeding and feedback mechanisms, they evolve with their host galaxies, and perhaps modulate each other’s growth (Kormendy & Ho 2013).

The exact nature of (feedback) mechanisms that transform galaxies from star-forming to quiescent is unknown. Star formation can be quenched by removing or heating gas in galaxies or their surrounding halos (Dekel & Silk 1986; Silk & Rees 1998; Di Matteo et al. 2005; Dekel & Birnboim 2006; Hopkins et al. 2006; Martig et al. 2009; Peng et al. 2010; Fabian 2012; Peng et al. 2015; Dubois et al. 2016; Pillepich et al. 2018), either through external mechanisms that are governed by the dark matter halo or the environment, or through internally driven processes such as feedback from active galactic nuclei (AGNs), stellar feedback, and morphological quenching.

Studying gas in the interstellar medium is critical for understanding the coevolution of supermassive black holes and their host galaxies. However, directly measuring gas content using radio telescopes for large, representative galaxy samples is difficult and time consuming. Thus, to date, small and heterogenous samples of AGN hosts with cold gas measurements have yielded mixed results on the impact of AGN feedback. While some studies report suppressed gas content in AGNs (Haan et al. 2008; Brusa et al. 2015; Kakkad et al. 2017; Perna et al. 2018), others find that active galaxies have gas content similar to or even higher than inactive galaxies (Maiolino et al. 1997; Bertram et al. 2007; Ho et al. 2008; Fabello et al. 2011; Gérèb et al. 2015; Husemann et al. 2017; Saintonge et al. 2017; Yesuf et al. 2017a; Rosario et al. 2018; Shangguan et al. 2018; Ellison et al. 2019; Shangguan & Ho 2019; Jarvis et al. 2020; Zhuang & Ho 2020).

We recently proposed a cost-effective method to predict molecular gas masses using dust absorption as inferred from $\text{H}\alpha$/H$\beta$; Interstellar dust extinction (837); Molecular gas (1073); Galaxy evolution (594); Seyfert galaxies (1447); Supermassive black holes (1663); AGN host galaxies (2017).

Perna et al. 2018).
on the black hole accretion rate, AGN feedback is modeled using two main modes (although see Schaye et al. 2015). At high accretion rates, a radiative mode (also called quasar mode) is used (e.g., Di Matteo et al. 2005; Hopkins et al. 2006), while at low accretion rates, a kinetic mode (also called radio mode), producing high-velocity winds and/or radio jets, is used (e.g., Croton et al. 2006; Weinberger et al. 2017). Galaxy formation models that do not include AGN feedback form stars too efficiently and fail to reproduce even basic properties of massive galaxies.

Previous work related to IllustrisTNG simulations suggests that the transition between star-forming and quenched galaxies is tied to the onset of kinetic mode feedback (Weinberger et al. 2017; Habouzit et al. 2019; Nelson et al. 2019; Terrazas et al. 2020; Zinger et al. 2020). Kinetic feedback removes gas from the star-forming regions and also leads to a hotter, more dilute, and higher entropy circumgalactic medium (CGM) with long cooling times (Zinger et al. 2020). Hence, the AGN feedback in TNG simulations ejects and heats up the gas within and around galaxies. The cumulative wind energy from a low-accretion-rate black hole is crucial to producing gas-deficient, compact, quiescent galaxies, above a threshold of $M_\text{h} \approx 2 \times 10^8 M_\odot$ or $M_\text{h} \approx 3 \times 10^9 M_\odot$. Below this threshold most simulated central galaxies are star-forming, and above the threshold most are quiescent (Terrazas et al. 2020; Zinger et al. 2020).

In contrast, in the EAGLE simulations (Bower et al. 2017) it is proposed that star formation-driven outflows regulate the amount of gas reaching the black holes and set the characteristic stellar mass $M_\text{s} \approx 3 \times 10^{10} M_\odot$ or halo mass $M_\text{h} \approx 10^{12} M_\odot$. Outflows are efficient in low-mass galaxies, making it difficult for cool gas to build up in the central regions of galaxies. In massive galaxies, which are surrounded by hot halos, outflows cease to be buoyant and are less efficient, and thus the unrestrained central buildup of gas leads to rapid black hole growth. In turn, the black holes become effective at heating the halo gas, disrupt the infalling cool gas supply, and slowly starve the galaxies of the fuel for future star formation.

Compared to those in the EAGLE simulations, the winds in TNG simulations eject significantly larger amounts of gas from the centers of massive galaxies (Nelson et al. 2019; Mitchell et al. 2020). Likewise, EAGLE simulations do not predict a negative trend between specific SFR (SSFR) and AGN luminosity; AGN feedback does not reduce galaxy-wide instantaneous SFRs of luminous, simulated AGN host galaxies (Scholtz et al. 2018). The signature of AGN feedback is instead imprinted on the overall SSFR distributions of massive galaxies. Thus, there may be implicit/subtle relationships between the SFRs, gas fractions, and AGN luminosities because the timescale of an AGN episode is shorter than the timescale for the suppression of star formation by possibly multiple AGN episodes (Harrison et al. 2017; McAlpine et al. 2017; Scholtz et al. 2018; Schulze et al. 2019).

In Section 2, we describe the data, sample selection, and the methods used to identify our proposed evolutionary sequence. In Section 3, we present relationships among nuclear activity, gas, stellar, and structural properties, along the sequence. Section 4 presents a discussion of our results. A summary of this work and its main conclusions are given in Section 5.

### 2. Data and Methodology

#### 2.1. Data

Our galaxy sample is taken from the Seventh Data Release of the Sloan Digital Sky Survey (SDSS DR7; Abazajian et al. 2009; Alam et al. 2015). The publicly available Catalog Archive Server (CAS)\(^4\) is used to collate some of the measurements used in this work (e.g., emission-line fluxes and spectral indices). These data are supplemented with stellar mass and SFR from version 2 of the Galaxy Evolution Explorer–SDSS–Wide-field Infrared Survey Explorer Legacy Catalog\(^5\) (GSWLC-2; Salim et al. 2016, 2018), along with structural parameters (Sérsic index and ellipticity/axis ratio) derived from single-component Sérsic function fits (Simard et al. 2011), and environmental information from the SDSS group catalog (Lim et al. 2017). To assess the importance of bars on AGN fueling, we use a bar classification catalog based on machine learning (Domínguez Sánchez et al. 2018). The deep learning model was trained on visual classifications of SDSS images by expert astronomers (Nair & Abraham 2010). For the $i$th galaxy in a given sample of size $n$, the machine classification outputs a probability $p_i$ that it has a bar. Assuming a Poisson-binomial distribution for the number of barred galaxies in the given sample, we estimate its mean bar fraction, $\mu_{\bar{b}} = \frac{\sum_i p_i}{n}$, and its standard deviation, $\sigma_{\bar{b}} = \sqrt{\sum_i p_i (1 - p_i)/n}$. The binomial distribution is a special case of the Poisson-binomial distribution, when all probabilities are the same.

We define the effective mass surface density as $\Sigma_e = M_e/(2\pi R_{50,e})$, where $M_e$ is the total stellar mass and $R_{50,e}$ is the $z$-band half-light radius. The concentration index is defined as the ratio of 90th Petrosian radius to the 50th Petrosian radius in the $r$ band, $C = R_{90}/R_{50}.$

In our adopted group catalog\(^6\) (Lim et al. 2017), the groups were identified with an iterative halo-based group finder (Yang et al. 2007), which used the stellar mass of a central galaxy as well as the difference in stellar mass between the central galaxy and the $n$th most massive satellite as proxies for halo mass. An abundance matching technique was used to assign final halo masses to distinct groups. For groups that were not assigned masses by the abundance matching, due to mass incompleteness, halo masses were assigned based on the mean relation between the halo mass and its proxies obtained from the group finder. Mock galaxy samples constructed from EAGLE simulations were used to test and calibrate the group finder.

#### 2.2. The Main Sample Selection

The following steps are taken to select the main sample used in this work.

1. Select galaxies in the redshift range $z = 0.02-0.15$ and stellar mass range $\log (M_\ast/M_\odot) = 10.1-11.3$ from the four

\(^4\) [http://skyserver.sdss.org/casjobs/](http://skyserver.sdss.org/casjobs/) We use CAS in the context of data release 13 (DR13) to retrieve various measurements in different catalogs. However, the sample of galaxies we use is restricted to those in DR7, because the measurements of axis ratio and Sérsic index from Simard et al. (2011) are only available for DR7 galaxies. The following tables are queried: photoobjall, galSpecIndx, galSpecInfo, galSpecLine, galSpecExtra, and specDR7.

\(^5\) [http://gax.sjtu.edu.cn/data/Group.html](http://gax.sjtu.edu.cn/data/Group.html) The group catalog is based on SDSS DR13.
narrow mass samples shown in Figure 1. The four samples, denoted by S1–S4, have stellar mass log($M_*/M_\odot$) = 10.1–
10.4, 10.4–10.7, 10.7–11.0, and 11.0–11.3, respectively, and $z = 0.02$–0.09, 0.02–0.12, 0.02–0.15, and 0.02–0.15, respectively.

2. Include only central galaxies (not satellites) in the group catalog (Lim et al. 2017). The results are similar if we only use isolated galaxies, with no nearby neighbors.

3. Exclude edge-on galaxies (axis ratio $b/a < 0.5$). The orientation effect may complicate the estimation of gas mass from dust absorption in edge-on galaxies.

4. Exclude type 1 AGNs by restricting the velocity dispersions of the Balmer emission lines to values less than $\sigma_{H\alpha} = 400$ km s$^{-1}$, below which very few type 1 AGNs exist (Greene & Ho 2007). For these objects, the AGN may significantly affect the measurements of the host galaxy properties.

5. Select galaxies with signal-to-noise ratio (S/N) > 3 for the density fluxes of H$\alpha$, H$\beta$, [O $\text{III}$] $\lambda$5007, [O $\text{II}$] $\lambda$3726, 3729, and [O $\text{I}$] $\lambda$6300.$^7$ These criteria ensure that the molecular gas mass, AGN luminosity, and ionization parameter are reliable.

6. Select galaxies with well-measured (fiber) stellar velocity dispersion (S/N > 3). The main results are similar for all four samples, and we take the sample S2 (log($M_*/M_\odot$) = 10.4–10.7) as the fiducial sample. To avoid redundancy, we selectively present some of the results for the other samples.

2.3. Estimation of Molecular Gas Masses

We calculate the V-band dust absorption using the observed H$\alpha$/H$\beta$ ratio and the dust attenuation curve (Charlot & Fall 2000, see also Wild et al. 2011), as follows ($\lambda$ in Å):

$$Q_\lambda = 0.6(\lambda/5500)^{-1.3} + 0.4(\lambda/5500)^{-0.7}.$$  (1)

Assuming that the intrinsic Balmer decrement H$\alpha$/H$\beta$ = 2.86 for inactive galaxies and H$\alpha$/H$\beta$ = 3.1 for AGNs (e.g., Ferland & Netzer 1983; Gaskell & Ferland 1984),

$$A_V = \frac{2.5}{Q_{5000} - Q_{6500}} \times \log \frac{H\alpha/\beta}{3.1 \text{ or } 2.86},$$  (2)

where $Q_{5000} - Q_{6500} = 0.31$. If the observed ratio of an object is below the intrinsic ratio (2.86 or 3.1), we set $A_V = 0$ mag.$^8$

For reference, $A_V = [0.5, 1, 1.5, 2, 3]$ mag corresponds to H$\alpha$/H$\beta \approx [3.3, 3.8, 4.4, 5.1, 6.8]$.

We estimate the molecular gas mass using our empirical (median) estimator presented in our previous work (Yesuf & Ho 2019). In particular, we use nebular dust absorption, the average gas-phase metallicity, which is inferred from the stellar mass–metallicity relation (Tremonti et al. 2004), and the half-light radius to estimate $M_{\text{HI}}$, with uncertainty of ~0.4 dex. The results for the gas masses are similar if we use the dust absorption derived from fitting the stellar continuum alone or in combination with H$\alpha$/H$\beta$.

2.4. Estimation of Black Hole Accretion Rate

We combine dust-corrected [O $\text{III}$] $\lambda$5007 and [O $\text{I}$] $\lambda$6003 luminosities to estimate the total bolometric luminosity ($L_{\text{bol}}$) of the AGN, with an uncertainty of ~0.5 dex (Netzer 2009). We calculate the Eddington luminosity, $L_{\text{Edd}} = 1.26 \times 10^{38}(M_*/M_\odot)$ erg s$^{-1}$, using black hole masses ($M_*$) derived from the fiber stellar velocity dispersions (Kormendy & Ho 2013). The SDSS fiber spans different regions of galaxies for different redshifts; we do not correct for this effect in our black hole estimates, which has a nominal uncertainty of ~0.3 dex. Thus, the resulting Eddington ratio ($\lambda_{\text{Edd}} = L_{\text{bol}}/L_{\text{Edd}}$), a rough approximation of the mass accretion rate, has an uncertainty of ~0.6 dex.

$^7$ This might cause the object’s predicted gas mass to be lower by up to ~0.1 dex than it actually is. Such small changes are not important for our application. An object with H$\alpha$/H$\beta < 2.9$–3.1 might have log($M_{\text{HI}}/M_\odot$) $\leq 8.5$–8.8.

$^8$ If the requirement of the faint [O $\text{I}$] $\lambda$6300 line is omitted, the sample size increases significantly but the results that do not depend on [O $\text{I}$] do not change significantly.
2.5. SPCA and Clustering Analysis in Narrow Mass Ranges

Star formation and AGN activities evolve on timescales that are much shorter than a galaxy merger timescale (Fang et al. 2013; Yesuf et al. 2014; Rodriguez-Gomez et al. 2015; Trayford et al. 2016; Hahn et al. 2017). Most local star-forming galaxies are unlikely to grow more than 0.3 dex in stellar (or black hole) mass by mergers in the time it takes them to cease forming stars. Galaxies that have similar masses but different internal mass distributions and environments are also unlikely to be immediate progenitors and descendants. Thus, to select candidate galaxies that are likely to be directly evolutionarily linked, we choose central, massive galaxies in four narrow stellar mass ranges.

Even in a restricted mass range and environment, galaxies are diverse in their gas content, SFR, and star formation history (SFH), stellar population age, internal kinematics, morphology (mass/light distribution), and AGN activity. Some of these properties not only correlate with each other but also may be sensitive to multiple, slightly different indicators. To explore statistically the evolution of central, massive galaxies in the four samples of approximately constant mass, we perform sparse principal component analysis (SPCA) and clustering analysis of 12 variables that describe the morphology and the AGN and star formation activities of galaxies (Figure 2). SPCA seeks a reduced set of new, uncorrelated variables, called principal components (PCs), which are linear combinations of the original variables. PCs sequentially capture the maximum variance in the data, and the first PC accounts for the highest variance. Because each PC is a linear combination of all variables, usually with non-zero weights, it is generally difficult to interpret the derived PCs. SPCA is a modern variant of PCA in which the PCs are derived from only a few of the most important variables. Furthermore, the clustering analysis aims to partition the reduced set of variables into “clusters,” such that galaxies assigned to the same cluster are less dissimilar than those in different clusters.9 After efficiently identifying the clusters using all the relevant information, we visualize how they are distributed in two-dimensional spaces of well-known diagrams to reveal evolutionary paths. Stellar spectral indices are used to gauge the average stellar population age of each cluster, and to determine its position on the evolutionary sequence.

We implement SPCA using the sparsepca package (Erichson et al. 2018) in R. SPCA can be formulated as a regularized regression-type problem (Zou et al. 2006). Given the data matrix $X$, SPCA attempts to minimize the function $\frac{1}{2}\|X - XB\|^2 + \alpha\|B\| + \frac{1}{2}\|\beta A\|^2$, subject to $A'A = I$, where $B$ is a sparse weight matrix, $A$ is an orthonormal matrix, and $I$ is the identity matrix. So, the PC matrix is $XB$.

The last two terms of the function are known as the elastic net regularizer. The first term shrinks the weights toward zero if $\alpha$ is large enough. We use four PC components, accounting for ~85% of the variance, and $\alpha = \beta = 10^{-3}$. The main results do not change significantly if we instead use a value of $10^{-4}$ for both $\alpha$ and $\beta$. We also center and scale the variables to have unit variance.

We perform $k$-means clustering of the PC scores in $\mathbb{R}$ to group the data into $k = 5$ or $k = 6$ clusters. One of the simplest and most popular clustering algorithms, $k$-means aims to iteratively partition a given data set of $n$ observations into $k$ optimal clusters. In the end, each observation is assigned to the nearest cluster, such that the squared Euclidean distances of the observations from the cluster centers are minimized. Determining objectively the optimal number of clusters in a data set is not an easy problem. Different metrics (the gap statistics, the elbow method, etc.) indicate $k \approx 3$–6. Using $k = 3$ gives similar results in which three of the four bulge-dominated groups (C1–C3) are merged together. While this reinforces the fact that these clusters are evolutionarily related, this choice, unlike $k = 5$ or $k = 6$, does not provide sufficiently fine distinction to illustrate the evolutionary stages discussed in Section 3.

2.6. Toy Stellar Population Models

To illustrate how a recent burst of star formation evolves in the $D_n(4000)-H\delta_A$ diagram (Balogh et al. 1999; Kauffmann et al. 2003), we use the updated version of the stellar population code of Bruzual & Charlot (2003) to model the SFHs of galaxies as a superposition of an old stellar population initially formed at time $t_1 = 5$ Gyr ($z \approx 1.2$) and a young population that formed in a recent burst at $t_2 = 12$ Gyr ($z \approx 0.12$). The recent burst fraction is 5%, 10%, or 30%. The old population formed following a delayed exponential SFH of the form $\psi \propto \exp(-t/\tau_1)$ with $e$-folding time $\tau_1 = 1$ Gyr, while the young population has SFH of the form $\psi \propto \exp(-t/\tau_2)$ with $\tau_2 = 0.1$ Gyr (Yesuf et al. 2014). The models with a single stellar population assume a stellar initial mass function according to Chabrier (2003) and solar

---

9 The “clusters” identified through the clustering analysis should not be confused with gravitationally bound galaxy clusters.
metallicity before the recent burst and 2.5 times solar after the burst.

3. Results

3.1. Relationships among Nuclear Activity, Gas, Stellar, and Structural Properties

Figure 3 shows how the morphology (mass surface density, velocity dispersion, concentration or Sérsic $n$) of the five clusters relates to their SFRs. The cluster C0 mostly contains less dense ($\Sigma_g \lesssim 10^{8.5} M_\odot$ kpc$^{-2}$), lower velocity dispersion ($\sigma_v < 100$ km s$^{-1}$), and less concentrated ($C < 2.5$ and $n < 2$) galaxies than the other clusters. The galaxies in C0 are exclusively star-forming galaxies. The clusters from C1 to C4 span a wide range in SFR and gas mass, at a constant morphology. C1 has the highest average SFR while C4 has the lowest, and those of C2 and C3 are intermediate between them. The galaxies in C0 are likely pseudobulges, and the rest...
classical bulges or ellipticals (Yesuf et al. 2020a). Although our current theoretical understanding of bulges is incomplete, the common view is that classical bulges and ellipticals are formed mainly by early dissipative mergers, while pseudobulges are formed by disk-related, internal secular processes (Kormendy & Kennicutt 2004; Brooks & Christensen 2016). Black holes also correlate differently with the properties of the two bulge types (Kormendy & Ho 2013). The correlations with pseudobulge properties are weak and imply no close coevolution. Thus, galaxies in C0 may have undergone different formation histories than those in the other clusters. While they may eventually evolve to become galaxies in C1, this is unlikely to happen in a short period of time.

Simulations show that the evolution tracks in the SFR–Σ, plane (Figure 3) have a characteristic “L” shape (Zolotov et al. 2015; Tacchella et al. 2016; Choi et al. 2018). Gas consumption by star formation and gas ejection by stellar and supernova feedback lead to quenching of star formation at a critical value of Σ. Further gas accretion is prevented because the galaxies have reached the critical halo mass to form a stable virial shock, which keeps the CGM at the virial temperature (Dekel & Birnboim 2006). The halo mass for our fiducial sample is log(M_h/10^12M_☉) = 12 ± 0.1 (Lim et al. 2017). In the EAGLE simulations, Correa et al. (2019) found that the time when galaxies move to the red sequence depends on their morphology, and AGN feedback is important for quenching central ellipticals, but not for disks. Next, we show the nature of the nuclear activities in groups C0 to C4.

The diagnostic diagrams of optical emission-line intensity ratio effectively discriminate whether the source of ionization in a galaxy arises from recently formed hot, massive stars or accretion onto a black hole (Baldwin et al. 1981; Veilleux & Osterbrock 1987; Ho et al. 1997a; Kewley et al. 2001, 2006; Kauffmann et al. 2003). Figure 4(a) shows the most sensitive of the line-ratio diagrams, involving [O III] λ5007/Hβ and [O I] λ6300/Hα. AGN-dominated galaxies occupy the upper right corner of the diagram, while star formation-dominated galaxies lie in the lower left corner. Galaxies hosting AGNs lie on two branches associated with Seyferts and LINERs, whose difference in ionization stems from their difference in accretion rate (Ho 2009). Our clusters occupy distinct loci normally assigned to standard galaxy spectral classes, even though the line-ratio classification criteria were not among the parameters used in the SPCA. For the fiducial sample with stellar mass log(M_*/10^10M_☉) = 10.4–10.7, nearly all of the members in C1 (98%) and 86% of C2 are classified as star-forming galaxies, with the former systematically more luminous than the latter owing to their larger SFRs. By contrast, AGNs comprise most (75%–90%) of the galaxies in C3 and C4, predominantly Seyferts in C3 (74%) and LINERs in C4 (74%), whose difference in AGN power is reflected in their [O III] luminosity (Heckman et al. 2004). The four classes segregate quite cleanly in a plot of L ([O III]) versus [O III]/Hβ (Figure 4(b)). An additional, similar result for [N II] λ6584/Hα and [S II] λ6716, 6731/Hα diagrams is given in the Appendix.

Table 1 quantifies the AGN fractions (divided into Seyferts and LINERs) and star-forming galaxy fractions for the four samples. Throughout this work, we use the [O III]/Hβ and [O I]/Hα line intensity ratios to define the activity types (Kewley et al. 2006). For a cluster of a given sample, we estimate the standard errors of the fractions of the activity types in the cluster as √f(1−f)/n, where f is the fraction of either Seyferts or LINERs or star-forming galaxies in the cluster and n is the sample size of the cluster. For all samples, the C2 and C1 clusters are dominated by star-forming galaxies, C3 by Seyferts, and C4 by LINERs. In other words, there is a broad correspondence between our classification and those of previous work. However, there are significant differences because our clusters contain mixtures of activity types of varying proportions. The aim of the previous classifications was to cleanly separate the activity types based on emission-line ratios using photoionization models or empirical data.

Figure 4. (a) Evolution on spectral diagnostic diagrams for galaxies with stellar mass log(M_*/10^10M_☉) = 10.4–10.7. The solid curve demarcates the theoretically maximum boundary for star-forming galaxies; galaxies above this curve are dominated by AGN emission (Kewley et al. 2001). The dotted line separates strong AGNs (Seyferts) from weak AGNs (LINERs) (Kauffmann et al. 2003; Kewley et al. 2006). Without using the standard emission-line ratios, our clustering analysis identifies four clusters that broadly correspond to star-forming galaxies, Seyferts, and LINERs. The clusters evolve from C1 (blue) to C4 (maroon). In panel (b), C3 is distinguishable from C2 and C4 because it has stronger L([O III]) from black hole accretion; likewise, C1 has higher SFRs and therefore produces stronger L([O III]) than C2.

The Astrophysical Journal, 901:42 (14pp), 2020 September 20

Yesuf & Ho
analysis (Kauffmann et al. 2003; Kewley et al. 2006; de Souza et al. 2017). In contrast, our aim is to study the correlation among activity type, gas content, SFR, and structural properties. The clear correspondence between the four galaxy populations identified through the SPCA with distinct spectral classes and luminosity states strongly suggests that some physical parameter controls the relative dominance and evolutionary trajectory of the two principal energy mechanisms (star formation and black hole accretion) during the life cycle of galaxies. What could that be?

Answer: gas content. Figure 5(a) shows that AGN-dominated galaxies (C3 and C4) exhibit a moderately strong relationship (Spearman correlation coefficient $\rho \approx 0.5$) between the molecular gas fraction ($f_{\text{H}_2} \equiv M_{\text{H}_2}/M_*$) and the Eddington ratio, such that more gas results in higher levels of black hole accretion. Seyferts (C3) possess roughly twice as much molecular gas as LINERs (C4). Thus, while the quiescence of LINERs has been attributed to their radiatively inefficient central engines (Ho 2008, 2009b), their host galaxies also suffer from a starvation diet. The sizeable scatter in Figure 5(a) cautions that the causal link between global gas supply and black hole feeding is not one-to-one, but likely stochastic (Hopkins & Hernquist 2006; Novak et al. 2011; Yuan et al. 2018). After all, galaxies in C2 and C3 enjoy similar gas reservoirs ($f_{\text{H}_2} \approx 4\%$) and yet those in C3 have higher $\lambda_{\text{Edd}}$, likewise, C1 is more gas-rich than C3 despite having lower $\lambda_{\text{Edd}}$ on average.

As expected (Genzel et al. 2015; Saintonge et al. 2017), the SSFR (SSFR $\equiv SFR/M_*$) of star-forming galaxies scales with $f_{\text{H}_2}$. But so too do the hosts of bright AGNs (C3), which completely overlap with galaxies with moderate levels of star formation (C2) that lie at the lower end of the blue cloud or main sequence (Figure 5(b)). Indeed, the trend continues to hold even for gas-poor passive galaxies (C4). The observation that AGNs preferentially sit below the main sequence has been routinely (mis)interpreted as evidence for AGN feedback (e.g., Schawinski et al. 2007; Leslie et al. 2016). That C2 and C3 have similar $f_{\text{H}_2}$ and SSFR, despite their difference in nuclear activity, argues against instantaneous negative AGN feedback, perhaps contravening the idea that AGN-driven outflows remove large amounts of gas from galaxies in a short period of time. Galaxies in C1 to C3 likely evolve to C4 when star formation depletes their gas on a timescale $\tau_{\text{dep}} \equiv M_{\text{H}_2}/SFR \lesssim 1$ Gyr. Galaxies in C3 and C2 have median (16%, 84%) depletion times $\log(\tau_{\text{dep}}/\text{yr}) = 8.9 (8.5, 9.2)$ and 8.9 (8.7, 9.3), respectively. The amount of gas needed to power the Seyfert activity observed in C3 is small (<0.03 $M_\odot \text{yr}^{-1}$), and even continuous accretion onto the black hole would not deplete $\sim 10^9 M_\odot$ in several gigayears. AGN-driven outflows with mass outflow rates of several solar masses per year can do the job, but the mass outflow rate is highly uncertain (Harrison et al. 2018) and difficult to estimate. Fluetsch et al. (2019) found depletion times of molecular gas between a few times $10^6$ and $10^8$ yr for small samples of local AGNs.

Black hole accretion and star formation are closely intertwined but not coeval. The significant scatter in the $D_{\text{c}}(4000)$–$H_\alpha$ diagram (Figure 5(c)) requires a bursty SFH (Kauffmann et al. 2003). Our toy stellar population models can reproduce the data with recent bursts with mass fractions of $\sim 5$–30% superposed on an underlying old population. The AGN activity may be linked but delayed with respect to the most recent burst (Kauffmann et al. 2003; Yesuf et al. 2014) by $\sim 50$–250 Myr (Davies et al. 2007; Wild et al. 2010; Yesuf et al. 2014). This is corroborated by the clear separation

| Sample | Cluster | Seyferts | LINERs | Star-forming Galaxies | $\log (L_{\text{O}3\alpha}/L_\odot) > 7$ |
|--------|--------|----------|--------|-----------------------|----------------------------------|
| C0a    | 0.003 ± 0.001 | 0.009 ± 0.002 | 0.988 ± 0.003 | 0.0 |
| C0b    | 0.009 ± 0.002 | 0.979 ± 0.003 | 0.020 ± 0.003 |
| C0     | 0.010 ± 0.002 | 0.014 ± 0.002 | 0.976 ± 0.003 | 0.573 ± 0.01 |
| C1     | 0.036 ± 0.004 | 0.031 ± 0.003 | 0.932 ± 0.005 | 0.239 ± 0.009 |
| C2     | 0.045 ± 0.005 | 0.036 ± 0.004 | 0.919 ± 0.006 | 0.930 ± 0.006 |
| C3     | 0.056 ± 0.005 | 0.195 ± 0.009 | 0.749 ± 0.010 | 0.280 ± 0.010 |
| C4     | 0.018 ± 0.013 | 0.737 ± 0.016 | 0.242 ± 0.016 | 0.963 ± 0.007 |
| C5     | 0.028 ± 0.012 | 0.011 ± 0.003 | 0.187 ± 0.012 | 0.994 ± 0.002 |
| C6     | 0.012 ± 0.009 | 0.833 ± 0.010 | 0.039 ± 0.005 | 0.081 ± 0.007 |
| C7     | 0.017 ± 0.011 | 0.173 ± 0.013 | 0.720 ± 0.015 | 0.362 ± 0.016 |
| C8     | 0.079 ± 0.010 | 0.095 ± 0.011 | 0.826 ± 0.014 | 0.971 ± 0.006 |
| C9     | 0.089 ± 0.014 | 0.021 ± 0.006 | 0.090 ± 0.013 | 0.996 ± 0.003 |
| C10    | 0.136 ± 0.012 | 0.533 ± 0.018 | 0.331 ± 0.017 | 0.539 ± 0.018 |
| C11    | 0.069 ± 0.009 | 0.917 ± 0.009 | 0.014 ± 0.004 | 0.053 ± 0.008 |

Table 1: The Fraction of AGNs and Star-forming Galaxies

Note. There are some variations between clusters in different mass samples. We use numbers 0 to 4 to group the clusters into five approximately distinct categories, across different mass ranges. The cluster C4a/C2 of the most massive sample shows overlapping properties with both clusters C2 and C4 in the fiducial sample, S2, and cluster C4b in the same mass range.
of the distributions of $D_n(4000)$ and $\text{H}\delta_A$ for C1 and C3. The time delay can be attributed to stellar feedback or dynamical and viscous lags experienced by the gas as it journeys from kiloparsec scales en route to the central black hole (Wild et al. 2010; Hopkins 2012; Blank & Duschl 2016).

Tables 2 provides summary statistics for the distributions of $\text{H}\alpha/\text{H}\beta$, gas mass, gas fraction, SFR, the 4000 Å break, morphological parameters, and the numbers of galaxies in the five clusters for each of the four samples. As expected, for a given cluster, SFRs and molecular gas masses increase and gas fractions decrease as the mass range of the sample increases. C2 and C3 have similar gas, stellar, and structural properties in all four samples. Their gas mass is $M_{\text{HI}} \approx (1 - 2) \times 10^9 M_\odot$. This indicates that AGN feedback does not significantly impact the cold gas content of galaxies instantaneously. Substantial amounts of cold gas still remain in strong AGNs. Note that the total gas mass, including HI, is likely $\sim (4 - 5) \times M_{\text{HI}}$ (Catinella et al. 2018).

Table 3 presents the Spearman correlation coefficients ($\rho$) for the quantities plotted in Figure 5 and for the four samples. We provide $\rho$ values of the relations including all four clusters,
| Sample | Cluster | Hα/Hβ | log $M_{\text{BH}}$ | log $f_{\text{H} \beta}$ | log SFR | $D_e(4000)$ | Hβ | C | $\sigma$ | Sérsic n | log $\Sigma_*$ | N Gal. |
|--------|---------|--------|---------------------|------------------------|---------|-------------|------|---|--------|---------|----------------|-------|
| C0a    | 3.9 (3.6, 4.3) | 9.1 (8.9, 9.2) | -1.2 (-1.3, -1.1) | 0.2 (0.1, 0.4) | 1.3 (1.2, 1.4) | 4.0 (2.8, 5.1) | 2.1 (1.9, 2.3) | 66 (49, 82) | 0.9 (0.7, 1.4) | 8.3 (8.1, 8.6) | | 1478 |
| C0b    | 4.1 (3.8, 4.6) | 9.1 (9.0, 9.2) | -1.1 (-1.3, -1.0) | 0.5 (0.3, 0.6) | 1.2 (1.2, 1.3) | 5.2 (4.3, 6.0) | 2.2 (2.0, 2.4) | 76 (59, 94) | 1.0 (0.7, 1.5) | 8.5 (8.3, 8.7) | | 1729 |
| S1     | 4.5 (3.9, 5.2) | 9.1 (8.9, 9.3) | -1.1 (-1.3, -1.0) | 0.6 (0.4, 0.8) | 1.2 (1.2, 1.3) | 5.4 (4.3, 6.3) | 2.6 (2.3, 2.8) | 92 (72, 113) | 2.1 (1.2, 4.3) | 8.8 (8.6, 9.0) | | 938 |
| C2     | 4.1 (3.6, 4.8) | 9.0 (8.8, 9.2) | -1.3 (-1.4, -1.1) | 0.1 (-0.1, 0.3) | 1.3 (1.3, 1.4) | 3.6 (2.5, 4.7) | 2.6 (2.4, 2.9) | 86 (68, 106) | 2.4 (1.5, 4.4) | 8.8 (8.6, 9.0) | | 1139 |
| C3     | 4.0 (3.6, 4.6) | 8.9 (8.7, 9.1) | -1.3 (-1.5, -1.2) | 0.0 (-0.3, 0.5) | 1.4 (1.3, 1.6) | 2.7 (0.9, 4.0) | 2.6 (2.2, 2.9) | 90 (69, 112) | 2.8 (1.4, 5.3) | 8.8 (8.6, 9.0) | | 301 |
| C4     | 3.6 (3.1, 4.1) | 8.7 (8.5, 8.9) | -1.5 (-1.7, -1.3) | -0.7 (-1.5, -0.3) | 1.6 (1.5, 1.9) | 0.6 (-1.1, 2.1) | 2.7 (2.5, 3.1) | 95 (75, 121) | 3.6 (2.2, 5.9) | 9.0 (8.7, 9.2) | | 385 |
| C0     | 4.3 (3.9, 4.7) | 9.3 (9.1, 9.4) | -1.2 (-1.4, -1.1) | 0.5 (0.3, 0.7) | 1.3 (1.2, 1.4) | 4.2 (2.9, 5.3) | 2.1 (2.0, 2.3) | 82 (65, 101) | 1.0 (0.7, 1.5) | 8.5 (8.3, 8.7) | | 2824 |
| C1     | 4.7 (4.2, 5.6) | 9.4 (9.2, 9.5) | -1.2 (-1.3, -1.0) | 0.8 (0.5, 1.0) | 1.2 (1.2, 1.3) | 5.1 (4.1, 6.0) | 2.4 (2.2, 2.8) | 109 (86, 133) | 1.6 (1.0, 3.1) | 8.8 (8.6, 9.0) | | 2374 |
| S2     | 4.3 (3.7, 5.0) | 9.2 (9.0, 9.4) | -1.4 (-1.5, -1.2) | 0.2 (0.0, 0.5) | 1.4 (1.3, 1.5) | 3.0 (1.8, 4.2) | 2.7 (2.4, 3.0) | 105 (86, 128) | 2.7 (1.7, 4.9) | 8.9 (8.7, 9.1) | | 1808 |
| C3     | 4.3 (3.8, 5.1) | 9.1 (8.9, 9.3) | -1.4 (-1.6, -1.2) | 0.3 (-0.1, 0.6) | 1.4 (1.3, 1.6) | 2.5 (0.9, 4.0) | 2.7 (2.3, 3.0) | 114 (94, 138) | 3.4 (1.6, 6.3) | 8.9 (8.7, 9.1) | | 699 |
| C4     | 3.6 (2.9, 4.4) | 8.9 (8.7, 9.1) | -1.7 (-1.9, -1.5) | -0.7 (-1.6, -0.2) | 1.7 (1.6, 1.9) | -0.5 (-2.2, 1.0) | 2.9 (2.6, 3.2) | 124 (101,158) | 4.4 (2.8, 6.5) | 9.0 (8.8, 9.2) | | 796 |
| C0     | 4.5 (4.0, 5.0) | 9.4 (9.3, 9.6) | -1.4 (-1.5, -1.2) | 0.8 (0.5, 1.0) | 1.3 (1.2, 1.4) | 3.8 (2.3, 5.0) | 2.2 (2.0, 2.4) | 104 (82, 127) | 1.2 (0.8, 2.0) | 8.6 (8.5, 8.8) | | 2509 |
| C1     | 5.3 (4.6, 6.3) | 9.6 (9.4, 9.8) | -1.3 (-1.5, -1.1) | 1.0 (0.7, 1.2) | 1.3 (1.2, 1.4) | 4.8 (3.6, 5.9) | 2.6 (2.3, 2.9) | 139 (114, 166) | 2.1 (1.3, 4.1) | 8.9 (8.7, 9.1) | | 1991 |
| S3     | 4.5 (3.9, 5.2) | 9.3 (9.2, 9.5) | -1.5 (-1.7, -1.3) | 0.4 (0.1, 0.7) | 1.5 (1.4, 1.6) | 2.5 (1.0, 3.8) | 2.8 (2.5, 3.1) | 131 (108, 156) | 3.6 (2.2, 6.1) | 8.9 (8.7, 9.1) | | 1969 |
| C3     | 4.3 (3.8, 5.0) | 9.3 (9.1, 9.5) | -1.6 (-1.8, -1.4) | 0.4 (0.1, 0.7) | 1.5 (1.3, 1.6) | 2.3 (0.3, 3.9) | 2.7 (2.3, 3.0) | 134 (107, 160) | 3.4 (1.8, 6.2) | 8.9 (8.7, 9.1) | | 1061 |
| C4     | 3.6 (2.9, 4.3) | 9.0 (8.8, 9.2) | -1.9 (-2.1, -1.6) | -0.6 (-1.4, 0.0) | 1.8 (1.7, 1.9) | -1.1 (-2.3, 0.3) | 3.0 (2.6, 3.3) | 152 (121, 188) | 4.6 (2.9, 6.5) | 9.0 (8.8, 9.2) | | 1345 |

Note. (1) Narrow mass samples. (2) Cluster names from the SPCa+ clustering analysis. (3) Balmer decrement. (4) Molecular gas mass in $M_{\odot}$. (5) Gas fraction, $f_{\text{H} \beta} = M_{\text{BH}}/M_{\odot}$. (6) SFR in $M_{\odot}$ yr$^{-1}$. (7) The 4000 Å break. (8) The equivalent width of Hβ absorption in Å. (9) Concentration index, $C = R_90/R_{50}$. (10) Stellar velocity dispersion in km s$^{-1}$. (11) Global Sérsic index. (12) Stellar mass density in $M_{\odot}$ kpc$^{-2}$. (13) Number of galaxies in a cluster. We use the notation $X (Y, Z)$ to denote $X = \text{median (50%)}$, $Y = 16\%$, and $Z = 84\%$ of a distribution.
only the AGN-dominated clusters C3 and C4, and only using AGNs classified according to Figure 4. In all four samples, AGNs show moderate ($\rho \approx 0.5$--0.7) correlations between $f_{\text{H}_2}$ and SSFR, and between $f_{\text{H}_2}$ and $\lambda_{\text{Edd}}$.

## 4. Discussion

### 4.1. Strong AGNs are Gas-rich and Dusty

In Yesuf & Ho (2019), we showed that dust absorption is a dirt-cheap method to estimate molecular gas masses for large samples of galaxies. This method can be applied to a variety of problems. Here, we use it to study $\sim 27,100$ central, massive galaxies in narrow stellar mass ranges in order to identify galaxy populations that are likely to be evolutionarily related. We identify four galaxy groups that can be evolutionarily linked through a life cycle wherein gas content modulates their SFR and level of AGN activity. We find that strongly accreting black holes (Seyferts) live in gas-rich, young, and star-forming hosts, unlike their weakly accreting counterparts (LINERs). Thus, we do not find evidence for rapid or instantaneous impact of AGN feedback on the molecular gas and dust in nearby galaxies. Our result supports the observational (Netzer 2009; Diamond-Stanic & Rieke 2012; Esquej et al. 2014; Delvecchio et al. 2015; Izumi et al. 2016; Shimizu et al. 2017; Shangguan et al. 2020; Zhuang & Ho 2020) and theoretical (Hopkins & Quataert 2010; Thacker et al. 2014; Volonteri et al. 2015; McAlpine et al. 2017) claim that strong AGNs cohabit in (gas-rich) star-forming galaxies.

In a companion paper (Yesuf & Ho 2020), we also apply our method to study a large sample of post-starburst galaxies (PSBs). They are a rare but important class of galaxies that are rapidly quenching. We find that large reservoirs of molecular gas are present in significant numbers of post-starburst AGNs, and that they are not removed or destroyed by AGN feedback in some PSBs. Next, we discuss the nature of outflows in the AGNs in the current sample, and whether AGNs are preferentially powered by bar-driven inflows.

### 4.2. AGN-driven Outflows

Simulations have shown that AGN feedback may drive strong outflows with little impact on the dense gas in the galaxy disk (Gabor & Bournaud 2014; Roos et al. 2015). AGNs mainly affect the diffuse gas in the interstellar medium and CGM. AGN-driven outflows in these simulations do not cause rapid quenching of star formation, but they may remove significant amounts of gas over long timescales ($\gtrsim 10^9$ yr). It is widely accepted that AGN feedback plays an important role in keeping the CGM hot and in maintaining long-term quenching (Croton et al. 2006; Fabian 2012). Although based on small samples and uncertain assumptions, Fluetsch et al. (2019) found scaling relations with the bolometric luminosity of AGN or SFR that predict average gas outflow rates in AGN hosts. Using these relations with median $\log(SFR/M_\odot \text{yr}^{-1}) = 0.3$ or median $\log(L_{\text{bol}}/\text{erg s}^{-1}) = 43.8$, we estimate an outflow rate for the C3 cluster in the S2 sample of $\sim 11 M_\odot \text{yr}^{-1}$ or $\sim 20 M_\odot \text{yr}^{-1}$, respectively. These estimates seem too high for the observed change in gas mass from C3 to C4 of $\sim 5 \times 10^8 M_\odot$. If the average time between C3 and C4 is $\Delta t \gtrsim 10^9$ yr, then the AGN outflow rate is $\lesssim 5 M_\odot \text{yr}^{-1}$, even ignoring the gas consumption by a SFR of $\sim 2 M_\odot \text{yr}^{-1}$.

Modeling the observed SFR by a SFH $\psi \propto \exp(-t/\tau)$ with $\tau = 100$–200 Myr, $\Delta t = 1.8\tau$. Likewise, in rapidly quenching galaxies, the time interval between the Seyfert and LINER post-starburst phases is $\gtrsim 3 \times 10^8$ yr (Yesuf et al. 2014). Furthermore, the AGN lifetime is likely $\sim 5 \times 10^7$–$5 \times 10^8$ yr for AGNs similar to those in C3 (Martini & Weinberg 2001; Gonçalves et al. 2008; Hopkins & Hernquist 2009; Borisova et al. 2016). Thus, while the current work cannot rule out significant gas loss due to AGN-driven outflows, we do not find a compelling reason to invoke it, in addition to gas consumption by star formation. Likewise, Jarvis et al. (2020) recently showed that some $\sim 0.1$ type 2 quasars with strong, kiloparsec-scale ionized gas outflows and jets are star-forming and gas-rich, and that there is no evidence for immediate appreciable impact of AGN feedback on the global molecular gas reservoirs of these powerful AGNs. Next, we briefly discuss the role of AGNs in regulating new gas accretion according to cosmological simulations (Davies et al. 2020; Oppenheimer et al. 2020; Terrazas et al. 2020; Zinger et al. 2020). Although there are quantitative differences between them, both EAGLE and TNG simulations indicate that the expulsion of rapidly cooling gas may lead to a long CGM cooling time, effectively depleting the halo of gas that would otherwise replenish the interstellar medium and fuel future star formation (Davies et al. 2020). In both sets of simulations, central halos with high CGM gas fractions ($f_{\text{CGM}}$) preferentially host galaxies that are more star-forming and have greater rotational support than typical halos. In addition, halos with lower $f_{\text{CGM}}$ at fixed mass have more massive black holes and higher ratios of cumulative black hole feedback energy to halo binding energy (Oppenheimer et al. 2020; Terrazas et al. 2020). At a fixed halo mass, $f_{\text{CGM}}$ is lower in the EAGLE simulations because the black holes, which are hosted by early-assembled halos, reach high accretion rates sooner. In contrast, in TNG simulations $f_{\text{CGM}}$ is lower at fixed mass when the central black holes reach the mass threshold at which AGN feedback switches from the thermal to kinetic mode. In EAGLE simulations, the scatter about the median $f_{\text{CGM}}$ is not correlated with the AGN luminosity/accretion rate, but TNG simulations predict strong correlation between these quantities, especially for the characteristic halo mass $M_h \approx 10^{12} M_\odot$. How these differences translate to the molecular gas in the interstellar medium and which simulations agree better with the trends presented in this work are yet to be seen in future studies. To interpret our results
in the meantime, both simulations indicate that the CGM cooling times for the present-day galaxies near the characteristic mass are $\gtrsim 1$ Gyr (Davies et al. 2020; Zinger et al. 2020). Hence, it may be assumed that the net inflow rates to the interstellar medium are not changing rapidly as local massive galaxies quench. The evolution of gas fraction already in the interstellar medium governs the observed rates of star formation and black hole accretion.

4.3. The Effect of Galactic Bars

Stellar bars can efficiently transfer angular momentum and instigate radial inflow of gas to trigger nuclear activity (Shlosman et al. 1989). However, the four evolutionary classes defined in our study have similar bar fractions (Table 4). Within sample S2, the bar fractions are $f_{\text{bar}} \approx 0.30 \pm 0.02$ for clusters C1 to C3, $0.25 \pm 0.01$ for C4, and $0.38 \pm 0.01$ for C0. Bar classifications using low-resolution images are highly uncertain. The values of $f_{\text{bar}}$ may increase by up to $\sim 10\%$ for $z < 0.07$. Similar to our results, previous attempts to connect AGN fueling to bars have yielded negative results (Ho et al. 1997b; Cisternas et al. 2013; Cheung et al. 2015; Galloway et al. 2015; Goulding et al. 2017; Neumann et al. 2019) or at best ambiguous ones (Laine et al. 2002; Alonso et al. 2018).

| Sample | Clusters | $f_{\text{bar}}$ | $f_{\text{bar}}(z < 0.07)$ |
|--------|----------|----------------|---------------------------|
| S1     | C0a      | 0.39 ± 0.01    | 0.41 ± 0.02               |
|        | C0b      | 0.36 ± 0.01    | 0.39 ± 0.02               |
|        | C1       | 0.26 ± 0.01    | 0.26 ± 0.02               |
|        | C2       | 0.27 ± 0.01    | 0.27 ± 0.02               |
|        | C3       | 0.29 ± 0.02    | 0.31 ± 0.03               |
|        | C4       | 0.23 ± 0.02    | 0.23 ± 0.02               |
| S2     | C0       | 0.38 ± 0.01    | 0.47 ± 0.02               |
|        | C1       | 0.31 ± 0.01    | 0.34 ± 0.02               |
|        | C2       | 0.31 ± 0.01    | 0.34 ± 0.02               |
|        | C3       | 0.30 ± 0.02    | 0.35 ± 0.03               |
|        | C4       | 0.25 ± 0.01    | 0.25 ± 0.02               |
| S3     | C0       | 0.40 ± 0.01    | 0.55 ± 0.03               |
|        | C1       | 0.31 ± 0.01    | 0.42 ± 0.04               |
|        | C2       | 0.34 ± 0.01    | 0.41 ± 0.02               |
|        | C3       | 0.36 ± 0.01    | 0.41 ± 0.04               |
|        | C4       | 0.27 ± 0.01    | 0.29 ± 0.02               |
| S4     | C0       | 0.49 ± 0.02    | 0.61 ± 0.04               |
|        | C1       | 0.35 ± 0.02    | 0.48 ± 0.06               |
|        | C3       | 0.44 ± 0.02    | 0.56 ± 0.06               |
|        | C4a/C2   | 0.29 ± 0.01    | 0.30 ± 0.04               |
|        | C4b      | 0.25 ± 0.01    | 0.28 ± 0.02               |

Figure 6. Schematic diagram summarizing the results of this study. We identify four “clusters” (C1–C4) of central, massive, and bulge-dominated galaxies that are evolutionarily linked. The life cycle begins with C1 and ends with C4, with the SFR and molecular gas fraction decreasing along the sequence. The intermediate-stage clusters C2 and C3 have similar gas and stellar properties, but C3 is dominated by strong AGNs while C2 has many fewer AGNs. The double-headed arrow connecting C2 and C3 indicates the possibility that galaxies may go back and forth between these two groups as the AGNs flicker. The most powerful AGNs (Seyferts) emerge during phase C3, terminating as LINERs in C4. The molecular gas is depleted within $\sim 1$ Gyr mainly by star formation, with negligible impact from AGN feedback.
4.4. Caveats

Our method can only predict gas masses within a factor of 2.5 (Yesuf & Ho 2019). For this reason, the intrinsic correlation of AGN activity or star formation with gas fraction may be stronger than reported here. Furthermore, we can only estimate the global gas content. We do not have information on the interstellar medium on nuclear scales. Lastly, our current study does not include quasars and unobscured, type 1 AGNs. It is unclear whether our main conclusions extend to these AGN populations. We should bear in mind, however, that a recent far-infrared study indicates that the gas contents of the host galaxies of type 1 and type 2 quasars are similar, and neither type is gas-deficient relative to normal, inactive galaxies (Shangguan & Ho 2019). This is supported by the work of Zhuang & Ho (2020), which is also based on the method of Yesuf & Ho (2019).

5. Summary and Conclusions

Utilizing our recent method to estimate the molecular gas mass using dust absorption (Yesuf & Ho 2019), we study the connection between gas content, star formation, and black hole accretion. Gas-rich galaxies first consume their fuel through vigorous star formation (C1). This is followed by an intermediate phase during which most gas supply sustains a reduced rate of star formation (C2) in concert with strong Seyfert activity mediated by stochastic accretion (C3). AGN feedback, however, does not seem to have instantaneous, observable effects on the molecular gas content or SFR. Finally, gas exhaustion extinguishes star formation and reduces black hole accretion to the levels seen in LINERs (C4).

AGN-dominated galaxies (C3 and C4) exhibit a moderately strong relationship (Spearman $\rho \approx 0.6$) between the molecular gas fraction and the Eddington ratio, such that more gas results in higher levels of black hole accretion. Seyferts live in young, star-forming, and gas-rich ($M_{\text{H}_2} \approx 10^{10} M_\odot$) hosts, with median (16%, 84%) depletion times $t_{\text{dep}}/\text{yr} = 8.9 (8.5, 9.2)$.

Our study adopts a data-driven approach and is not meant to test specific AGN feedback models. Future comparisons with our results may be useful to discriminate between different models. Our results would be inconsistent with AGN feedback models that predict that central, bulge-dominated, Seyfert-like AGNs in massive galaxies have significantly lower molecular gas fractions than inactive galaxies of similar mass, morphology, and SFR. In addition, future observational work can expand upon the current work by (1) directly measuring the molecular gas masses in a carefully selected, representative sample of galaxies (including low-luminosity AGNs and low-SFR AGNs) and repeating the correlations reported here, (2) estimating resolved gas masses either directly or indirectly and studying the effects of AGNs on the gas in nuclear regions, (3) using hard X-ray luminosity, which is less affected by strong star formation, instead of $L_{\text{X,0.1-4.5 keV}}$ as a proxy for black hole accretion rate, and (4) utilizing bar classifications based on images of higher resolution than those provided by SDSS to robustly study the connection between AGN fueling and bars.

We are very thankful to the anonymous referee for the helpful feedback, comments, and suggestions that significantly improved the content and presentation of the paper.

This work was supported by the National Science Foundation of China (11721303, 11991052, 11950410492) and the National Key R&D Program of China (2016YFA0400702).

Funding for SDSS 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 website 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.

Appendix

Additional Diagnostics

For completeness, Figure A1 plots the [N II] $\lambda 6584/\text{H}$ and [S II] $\lambda\lambda 6716, 6731/\text{H}$ diagrams (Kewley et al. 2001, 2006; Kauffmann et al. 2003; Schawinski et al. 2007) color-coded by the four clusters. The results are similar to Figure 4(a). AGN-dominated galaxies occupy the upper right corners of the diagrams, while star-forming galaxies occupy the lower left corners. In the [N II]/H$\alpha$ diagram, composite/transition galaxies with some mixture of AGN and star formation activity lie in between (Ho et al. 1993). The composite region is populated by all four classes. In other words, composite galaxies are heterogeneous and span a wide range in star formation, age, and gas content.

Figure A2(a) shows that, overall, the stellar velocity dispersion is correlated ($\rho = 0.7$) with gas velocity dispersion measured using all forbidden lines simultaneously. We simply denote it as $\sigma_{\text{O III}}$. This correlation is also observed in previous work (Greene & Ho 2005; Ho 2009a). Galaxies in C3, however, exhibit more turbulent gas motions relative to their stellar velocity dispersions. This may be due to AGN-driven outflows (Ho 2009a; Kong & Ho 2018), which are commonly observed in AGN hosts (see Fiore et al. 2017; Yesuf et al. 2017b, 2020b; Harrison et al. 2018, and references therein). However, without spatially resolved kinematics, the possibility of inflows of low-angular-momentum gas powering nuclear activity in C3 galaxies cannot be ruled out (Ho et al. 2003). Data with high spatial resolution show that both inflow and outflow can coexist in the same AGN host (Audiibert et al. 2019). The fact that the distributions of stellar velocity dispersions for the four clusters are similar implies that their black hole mass distributions are similar (Kormendy & Ho 2013). Their inferred median (16%, 84%) black hole mass is $\log(M_{\text{BH}}/M_\odot) \approx 7.4 (6.9, 7.7)$. 

12
Figure A2. Shows dust-corrected $[\text{O II}]$ luminosity and $[\text{O III}]/[\text{O II}]$ ratio. This figure aims to particularly show that the $[\text{O II}]$ and $[\text{O III}]$ luminosities of C1 galaxies are consistent with ionization by newly formed, massive stars, and that the $[\text{O II}]$ luminosities of these galaxies agree with the expectation from their SFRs (Zhuang & Ho 2019) given in Table 2.

ORCID iDs

Hassen M. Yesuf @ https://orcid.org/0000-0002-4176-9145
Luis C. Ho @ https://orcid.org/0000-0001-6947-5846

References

Abazajian, K. N., Adelman-McCarthy, J. K., Agüeros, M. A., et al. 2009, ApJS, 182, 543
Alam, S., Albareti, F. D., Allende Prieto, C., et al. 2015, ApJS, 219, 12
Alonso, S., Coldwell, G., Duplancic, F., Mesa, V., & Lambas, D. G. 2018, A&A, 618, A149
Audibert, A., Combes, F., García-Burillo, S., et al. 2019, A&A, 632, A33
Baldwin, J. A., Phillips, M. M., & Terlevich, R. 1981, PASP, 93, 5
Balogh, M. L., Morris, S. L., Yee, H. K. C., Carlberg, R. G., & Ellingson, E. 1999, ApJ, 527, 54
Bertram, T., Eckart, A., Fischer, S., et al. 2007, A&A, 470, 571
Blank, M., & Duschl, W. J. 2016, MNRAS, 462, 2246

Figure A1. Similar to Figure 4. The solid curve marks the theoretical maximum boundary for star-forming galaxies (Kewley et al. 2001), while the dashed curve represents the empirical boundary of pure star-forming galaxies (Kauffmann et al. 2003); composite galaxies between the two curves have both star formation and AGN activity. The dotted line in the AGN region separates Seyferts from LINERs (Kewley et al. 2006; Schawinski et al. 2007).
Fabian, A. C. 2012, ARA&A, 50, 455
Fabello, S., Kauffmann, G., Catinella, B., et al. 2011, MNRAS, 416, 1739
Esquej, P., Alonso-Herrero, A., González-Martín, O., et al. 2014, ApJ, 780, 86
Ho, L. C., Filippenko, A. V., & Sargent, W. L. W. 1993, ApJ, 417, 63
Ferland, G. J., & Netzer, H. 1983, ApJ, 264, 105
Borisova, E., Lilly, S. J., Cantalupo, S., et al. 2016, ApJ, 830, 120
Bower, R. G., Schaye, J., Frenk, C. S., et al. 2017, MNRAS, 465, 32
Brooks, A., & Christensen, C. 2016, Galactic Bulges, Astrophysics and Space Science Library, Vol. 418 (Cham: Springer), 317
Brusa, M., Fengilio, C., Cresci, G., et al. 2015, A&A, 578, A11
Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000
Catinella, B., Saintonge, A.,Janowiecki, S., et al. 2018, MNRAS, 476, 875
Chabrier, G. 2003, PASP, 115, 763
Charlot, S., & Fall, S. M. 2000, ApJ, 539, 718
Cheung, E., Trump, J. R., Athanassoula, E. et al., 2015, MNRAS, 447, 506
Choi, E., Somerville, R. S., Ostriker, J. P., Naah, T., & Hirschmann, M. 2018, ApJ, 866, 91
Cisternas, M., Gadotti, D. A., Knapp, J. H., et al. 2013, MNRAS, 449, 373
Di Matteo, T., Springel, V., & Hernquist, L. 2005, Natur, 433, 604
Dekel, A. & Birnboim, Y. 2006, MNRAS, 368, 2
Dekel, A., & Silk, J. 1986, ApJ, 303, 39
Delvecchio, L., Lutz, D., Berta, S., et al. 2015, MNRAS, 449, 373
Dekel, A., & Birnboim, Y. 2006, MNRAS, 368, 2
Dekel, A., & Silk, J. 1986, ApJ, 303, 39
Netzer, H. 2009, MNRAS, 399, 1907
Neumann, J., Gadotti, D. A., Wisotzki, L., et al. 2019, A&A, 627, A26
Novak, G. S., Ostriker, J. F., & Ciotti, L. 2011, ApJ, 737, 26
Oppenheimer, B. D., van den Bosch, F. C., et al. 2020, MNRAS, 491, 2939
Perna, M., Sargent, M. T., Brusa, M., et al. 2018, A&A, 619, A90
Pillepich, A., Springel, V., Nelson, D., et al. 2018, MNRAS, 473, 4077
Rodríguez-Gomez, V., Genel, S., Vogelsberger, M., et al. 2015, MNRAS, 455, 149
Shlosman, I., Frank, J., & Begelman, M. C. 1989, Natur, 338, 45
Silk, J., & Rees, M. J. 1998, A&A, 331, L1
Shen, W., Ho, L. C., Mao, S., & Li, Y. 2013, ApJ, 776, 78
Saintonge, A., Catinella, B., Tacconi, L. J., et al. 2017, ApJS, 233, 22
Salim, S., Boquien, M., & Lee, J. C. 2018, ApJ, 859, 11
Salim, S., Lee, J. C., Janowiecki, S., et al. 2016, ApJS, 227, 2
Yesuf, H. M., & Ho, L. C. 2019, ApJ, 884, 177
Yesuf, H. M., Faber, S. M., Trump, J. R., et al. 2014, ApJ, 792, 84
Yesuf, H. M., Koo, D. C., Faber, S. M., et al. 2020a, ApJ, 889, 14
Yesuf, H. M., Faber, S. M., Trump, J. R., et al. 2014, ApJ, 792, 84
Yesuf, H. M., & Ho, L. C. 2019, ApJ, 882, 89
Yesuf, H. M., & Ho, L. C. 2017a, ApJ, 839, 14
Yesuf, H. M., French, K. D., Faber, S. M., Koo, D. C. 2017b, ApJ, 841, 6
Yesuf, H. M., & Ho, L. C. 2015, ApJ, 800, 20
Yesuf, H. M., Koo, D. C., Faber, S. M., et al. 2020b, ApJ, 896, 108
Yesuf, H. M., Faber, S. M., Koo, D. C., & Lee, L. L. 2020b, MNRAS, submitted (arXiv:1706.09500)
Yesuf, H. M., Faber, S. M., Trump, J. R., et al. 2014, ApJ, 792, 84
Yesuf, H. M., French, K. D., Faber, S. M., & Koo, D. C. 2017a, ApJ, 839, 14
Yesuf, H. M., Faber, S. M., Trump, J. R., et al. 2014, ApJ, 792, 84
Yesuf, H. M., French, K. D., Faber, S. M., & Koo, D. C. 2017a, ApJ, 839, 14
Yesuf, H. M., & Ho, L. C. 2019, ApJ, 884, 177
Yesuf, H. M., & Ho, L. C. 2020, ApJ, 900, 107
Yesuf, H. M., Koo, D. C., Faber, S. M., et al. 2017b, ApJ, 841, 83
Yuan, F., Yoon, D., Li, Y.-P., et al. 2018, ApJ, 857, 121
Zhang, M.-Y., & Ho, L. C. 2019, ApJ, 882, 89
Zhang, M.-Y., & Ho, L. C. 2020, ApJ, 896, 108
Zinger, E., Pillepich, A., Nelson, D., et al. 2020, arXiv:2004.06132
Zou, H., Hastie, T., & Tibshirani, R. 2006, Journal of Computational and Graphical Statistics, 15, 265

Yesuf & Ho