Powerful Outflows and Feedback from Active Galactic Nuclei

Andrew King and Ken Pounds

Department of Physics and Astronomy, University of Leicester, Leicester LE1 7RH, United Kingdom; email: ark@leicester.ac.uk, kap@leicester.ac.uk

Annu. Rev. Astron. Astrophys. 2015. 53:115–54

First published online as a Review in Advance on April 10, 2015

The *Annual Review of Astronomy and Astrophysics* is online at astro.annualreviews.org

This article's doi: 10.1146/annurev-astro-082214-122316

Keywords

supermassive black holes, accretion, M–σ relation, X-ray winds, molecular outflows, quenching of star formation

Abstract

Active galactic nuclei (AGNs) represent the growth phases of the supermassive black holes in the center of almost every galaxy. Powerful, highly ionized winds, with velocities $\sim$0.1–0.2c, are a common feature in X-ray spectra of luminous AGNs, offering a plausible physical origin for the well-known connections between the hole and properties of its host. Observability constraints suggest that the winds must be episodic and detectable only for a few percent of their lifetimes. The most powerful wind feedback, establishing the M–σ relation, is probably not directly observable at all. The M–σ relation signals a global change in the nature of AGN feedback. At black hole masses below M–σ, feedback is confined to the immediate vicinity of the hole. At the M–σ mass, it becomes much more energetic and widespread and can drive away much of the bulge gas as a fast molecular outflow.

1. INTRODUCTION

1.1. Supermassive Black Hole Scaling Relations

Astronomers now generally agree that the center of almost every galaxy but the smallest contains a supermassive black hole (SMBH). In recent years it has become clear that the mass M of the hole correlates strongly with physical properties of the host galaxy. In particular the hole mass M appears always to be a fairly constant fraction of the stellar bulge mass $M_{\rm b}$, i.e.,

$$M \sim 10^{-3} M_{\rm b} \tag{1}$$

(Häring & Rix 2004). Even more remarkably, observations give a tight relation of the form

$$M \simeq 3 \times 10^8 \ {\rm M}_\odot \sigma_{200}^{\alpha} \tag{2}$$

between the SMBH mass and the velocity dispersion $\sigma = 200\sigma_{200}$ km s^{-1} of the host galaxy's central bulge, with $\alpha \simeq 4.4 \pm 0.3$ (Ferrarese & Merritt 2000, Gebhardt et al. 2000; see Kormendy & Ho 2013 for a recent review). For many practical cases the relation is more conveniently written as

$$M \simeq 2 \times 10^7 \ {\rm M}_\odot \sigma_{100}^{\alpha}, \tag{3}$$

with now $\sigma = 100\sigma_{100}$ km s^{-1}.

Because observationally determining the SMBH mass generally involves resolving its sphere of influence, of radius

$$R_{\rm inf} \simeq \frac{GM}{\sigma^2} \simeq 8 \frac{M_8}{\sigma_{200}^2} {\rm pc} \simeq 3 \frac{M_7}{\sigma_{100}^2} {\rm pc}, \tag{4}$$

with $M = 10^8 M_8 \ {\rm M}_\odot = 10^7 M_7 \ {\rm M}_\odot$, Equation 2 may represent a maximum SMBH mass for a given velocity dispersion σ (Batcheldor 2010).

1.2. Binding Energies

At first glance, Equations 1 and 2 appear surprising. Equation 4 shows that the black hole's gravity has a completely negligible effect on its host galaxy, which in most ways must be quite unaware of its existence. But we know (Soltan 1982) that the hole grew largely through luminous accretion of gas. This released energy,

$$E_{\rm BH} \simeq \eta M c^2 \sim 2 \times 10^{61} M_8 \ {\rm erg} \tag{5}$$

(where $\eta \simeq 0.1$ is the accretion efficiency), is far larger than the binding energy

$$E_{\rm bulge} \sim M_{\rm b} \sigma^2 \sim 8 \times 10^{58} M_8 \sigma_{200}^2 \ {\rm erg} \tag{6}$$

of a host bulge of stellar mass $M_{\rm b} \sim 10^3 M$.

The vast difference in these two numbers suggests that the host must notice the presence of the hole through its energy output, even though it is utterly insignificant in all other ways. We can already see how the black hole mass could correlate with galaxy properties—the hole grows by accreting gas, but in doing this it communicates some of its huge binding energy $E_{\rm BH}$ back to the gas reservoir and so potentially limits its own growth. This suggests that the most relevant quantity to compare with $E_{\rm BH}$ is not $E_{\rm bulge}$ but instead the gravitational binding energy of the bulge gas alone, i.e.,

$$E_{\rm gas} = f_{\rm g} E_{\rm bulge}, \tag{7}$$

where $f_g < 1$ is the gas fraction. In the following, we take this as $f_g \sim 0.16$, the cosmological mean value, giving a typical relation

$$E_{\rm BH} \sim 2 \times 10^3 E_{\rm gas} \tag{8}$$

for a black hole close to the M–σ relation (the right-hand side has an implicit factor $\sim \sigma_{200}^4/M_8$).

This picture requires the black hole to communicate some of its accretion energy to its host. But this process cannot be very efficient, as otherwise the hole could disrupt the host entirely or at the very least remove a large fraction of its gas. In this sense, the galaxy bulge leads a precarious existence. For much of its life it can ignore the threat that the SMBH poses, but we shall see that in the end this is always decisive if accretion continues.

1.3. Communicating the Energy: Feedback

There are two main ways that the SMBH binding energy can potentially affect its surroundings. By far the stronger one (in principle) is through direct radiation. After all, this is how all the accretion energy is initially released. But we know from observation that most light escapes relatively freely from active galactic nuclei (AGNs). This suggests that radiation is in general not the main way the SMBH affects its host, and we discuss in detail why this is so in Section 7.4. The second form of coupling SMBH binding energy to a host bulge is mechanical. The huge SMBH accretion luminosity drives powerful gas flows into the host, making collisions and communication inevitable. One form of flow often mentioned is jets—highly collimated flows driven from the immediate vicinity of the SMBH (see Fabian 2012 for a review). To turn these into a means of affecting most of the bulge requires a way of making the interaction relatively isotropic, perhaps with changes of the jet direction over time. Here we mainly consider another form of mechanical communication that automatically has this property already. This is the observed presence in many AGNs of near-isotropic winds carrying large momentum fluxes.

1.4. Powerful Ionized Winds

Early X-ray observations of AGNs yielded soft X-ray spectra frequently showing the imprint of absorption from ionized gas, the warm absorber (WA) (Halpern 1984, Reynolds & Fabian 1995). More recent observations have found at least 50% of radio-quiet AGNs showing WAs in their soft X-ray (~0.3–2 keV) spectra. The limited spectral resolution of the *Einstein Observatory* and *Advanced Satellite for Cosmology and Astrophysics* (ASCA) observations prevented important parameters of the WAs, in particular the outflow velocity and mass rate, to be determined with useful precision. The higher resolution and high throughput afforded by contemporary X-ray observatories *Chandra*, *XMM-Newton*, and *Suzaku* has transformed that situation over the past decade, with the WA being shown, typically, to be dominated by K-shell ions of the lighter metals (C, N, O, etc.) and Fe-L, with outflow velocities of several hundred kilometers per second (Blustin et al. 2005, McKernan et al. 2007).

A more dramatic discovery made possible with the new observing capabilities was the detection of blueshifted X-ray absorption lines in the Fe K band, indicating the presence of highly ionized outflows with velocities $v \sim 0.1$–$0.25c$ (Chartas et al. 2002; Pounds et al. 2003; Reeves et al. 2003). In addition to adding an important dimension to AGN accretion studies, the mechanical power of such winds, which for a radial flow depends on v^3, was quickly recognized to have a wider potential importance in galaxy feedback.

Additional detections of high-velocity AGN winds were delayed by the weakness of absorption in such highly ionized gas combined with strongly blueshifted lines in low-redshift objects

coinciding with falling telescope sensitivity above ~7 keV. However, further extended observations, particularly with *XMM-Newton*, found evidence in five additional AGNs for outflow velocities of ~0.1–0.2c (Cappi 2006). Some doubts remained as to how common high-velocity outflows were, as the majority of detections were of a single absorption line (with consequent uncertainty of identification) that had moderate statistical significance, raising concerns of publication bias (Vaughan & Uttley 2008). In addition, only for PG1211+143 had a wide-angle outflow been directly measured, confirming a high outflow rate of mass rate and mechanical energy in that case (Pounds & Reeves 2007, 2009).

These residual doubts were finally removed following a blind search of extended AGN observations in the *XMM-Newton* archive (Tombesi et al. 2010), finding compelling evidence in 13 (of 42) radio-quiet objects for blueshifted Fe K absorption lines with implied outflow velocities of ~0.03–0.3c. A later search of the *Suzaku* data archive yielded a further group of strong detections with a median outflow velocity again of ~0.1c (Gofford et al. 2013). In addition to confirming that high-velocity, highly ionized AGN winds are common, the high yield from these archival searches shows the flows must typically have a large covering factor and therefore be likely to involve substantial mass and energy fluxes.

The observed distributions of velocity, ionization parameter, and column density are all compatible with Eddington winds—launched from close to the black hole, where the optical depth $\tau_{es} \sim 1$—that carry the local escape velocity (King & Pounds 2003). However, because the mean luminosity in most low-redshift AGNs is on average sub-Eddington, such winds are likely to be intermittent, which is a view supported by repeated observations and by the range of observed column densities.

For the best-quantified high-velocity outflow (the luminous Seyfert galaxy PG1211+143), in which a wide-angle flow was directly measured (Pounds & Reeves 2007, 2009), the wind has more energy than needed to unbind the likely gas mass of the observed stellar bulge. This suggests that the energy coupling of wind-to-bulge gas must be inefficient, as seen in the discussion following Equation 8. Evidence that the fast wind in NGC 4051 is shocked at a distance of ~0.1 pc from the black hole offers an explanation of why such powerful winds do not disrupt the bulge gas: Strong Compton cooling by the AGN radiation field removes most of the wind energy before it can be communicated.

2. THE OBSERVATIONAL EVIDENCE FOR ULTRAFAST OUTFLOWS

As noted above, the requirement of X-ray observations with high sensitivity and good spectral resolution over a wide energy band delayed the discovery of powerful, highly ionized winds from AGNs other than those with broad absorption lines (BALs) until the launches of *Chandra* and *XMM-Newton*. A decade after the first reports (Pounds et al. 2003, Reeves et al. 2003), high-velocity ($v \sim 0.1c$), highly ionized winds are now established as common in low-redshift AGNs.

2.1. The Fast Outflow in PG1211+143

Exploring the nature of the "soft excess" in a sample of luminous Palomar Green AGNs was a primary target in the Guaranteed Time program awarded to Martin Turner, Project Scientist for the European Photon Imaging Camera (EPIC) on *XMM-Newton* (Turner et al. 2001). At that time the X-ray spectrum in AGNs above ~1 keV was expected to be a rather featureless power law apart from a fluorescent emission line at ~6.4 keV from near-neutral Fe. One source, PG1211+143, showed a surprisingly "noisy" X-ray spectrum that one of us (K.P.) volunteered to explore.

PG1211+143, at a redshift of 0.0809 (Marziani et al. 1996), is one of the brightest AGNs at soft X-ray energies. It was classified (Kaspi et al. 2000) as a narrow-line Seyfert 1 galaxy (FWHM Hβ 1,800 km s^{-1}) with black hole mass $\sim 4 \times 10^7$ $M_\odot$ and bolometric luminosity 4×10^{45} erg s^{-1}, indicating a mean accretion rate close to Eddington.

Analysis of the unusual spectral structure in the 2001 *XMM-Newton* observation of PG1211+143 showed it to be dominated by blueshifted absorption lines of highly ionized metals, providing the first evidence for a high-velocity ionized outflow in a non-BAL AGN, with the initial identification of a deep blueshifted Fe Lyman-α absorption line indicating an outflow velocity of $\sim$0.09c (Pounds et al. 2003). That observation, closely followed by the detection of a still higher outflow velocity from the luminous quasi-stellar object (QSO) PDS 456 (Reeves et al. 2003), attracted wide attention, potentially involving the mechanical ejection of a significant fraction of the bolometric luminosity, which is perhaps characteristic of AGNs accreting near the Eddington rate (King & Pounds 2003).

Appropriately for such an unexpected discovery, the validity of the high velocity in PG1211+143 was not unchallenged. The near-coincidence of the observed absorption line blueshift and the redshift of the host galaxy was a concern, despite the uncomfortably high column density of heavy metals implied by a local origin. Then, in a detailed modeling of the soft X-ray Reflection Grating Spectrometer (RGS) data, Kaspi & Behar (2006) found only a much lower velocity. However, any doubts relating to the absorption being local were removed by a revised velocity of 0.13–0.15c based on the inclusion of additional absorption lines from intermediate-mass ions (Pounds & Page 2006) and on the repeated observations of PG1211+143 demonstrating that the strong Fe K absorption line was variable over several years (Reeves et al. 2008).

Here we use the initial *XMM-Newton* observation of PG1211+143 in 2001 to illustrate the two methods used then—and since—to parameterize the ionized outflow. **Figure 1** shows the ratio of EPIC pn data to a simple power-law continuum with a deep absorption line seen near 7 keV and additional spectral structure at $\sim$1–4 keV.

Fitting a negative Gaussian to the deep $\sim$7-keV absorption line (**Figure 2*a***) finds an observed line energy of 7.06 $\pm$ 0.02 keV or 7.63 $\pm$ 0.02 keV at the AGN redshift of 0.0809. The line is clearly resolved with 1σ width $\sim$100 $\pm$ 30 eV. Assuming identification with the FeXXV resonance (6.70-keV rest energy), the blueshifted line corresponds to an AGN outflow velocity of $v \sim 0.122 \pm 0.005c$. The most likely a priori alternative identification, with the FeXXVI Lyman-α line (6.97-keV rest energy), conservatively adopted in the initial analysis (Pounds et al. 2003), yields a lower outflow velocity $v \sim 0.095 \pm 0.005c$.

An alternative procedure, which also provides additional parameters of the gas flow, requires full spectral modeling, as conducted by Pounds & Page (2006) and more widely in recent outflow studies (Section 2.3). For the 2001 *XMM-Newton* pn spectrum of PG1211+143, modeling the absorption from 1–10 keV with a photoionized gas derived from the XSTAR code of Kallman et al. (1996) gives an excellent fit for a column density $N_{\rm H} \sim 3.2 \pm 0.7 \times 10^{23}$ cm^{-2} and ionization parameter $\log \xi = 2.7 \pm 0.1$ erg cm s^{-1}, with an outflow velocity (in the AGN rest frame) $v \sim 0.149 \pm 0.003c$. The model profile (**Figure 2*b***) shows significant inner shell absorption components to the low-energy wing of the 1s-2p resonance line, which explains why simply identifying the absorption near 7 keV with the 6.7-keV rest energy of FeXXV gives too low a velocity.

Although individually weaker than the Fe absorption, the combination of resonance lines of He- and H-like Mg, Si, S, and Ar in the broadband spectral fit is evidently constraining the spectral fit. This conclusion is confirmed by Gaussian fits to corresponding absorption features in **Figure 1** which yield a weighted observed blueshift of 0.055 $\pm$ 0.05 and outflow velocity (at the AGN redshift) of $v \sim 0.14 \pm 0.01c$. This is consistent with the velocity found from spectral modeling

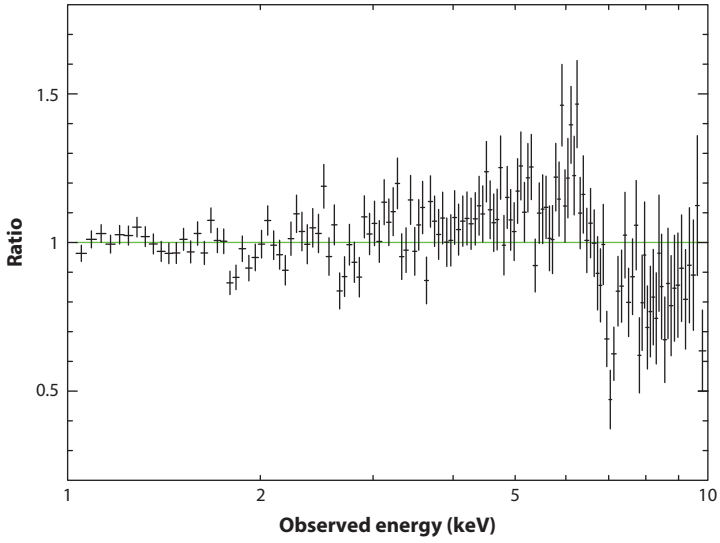

Figure 1

Ratio of EPIC (European Photon Imaging Camera) pn data to a simple power-law continuum for the 2001 *XMM-Newton* observation of PG1211+143 showing a deep absorption line near 7 keV and additional structure between ∼1 and 4 keV. Deriving an outflow velocity requires the correct identification of the individual absorption lines, which ideally requires spectral modeling with a photoionized absorber.

but significantly higher than results from simply identifying the ∼7-keV absorption line with the resonance 1s-2p transitions of either FeXXV or FeXXVI.

An interesting by-product of the XSTAR modeling in the above case is that the observed broadening of the ∼7-keV absorption line does not require high turbulence (we used grid 25 with $v_{\rm turb}$ of 200 km s^{-1}) or an accelerating/decelerating flow. Instead, intrinsically narrow absorption components remain consistent with a radial outflow, coasting post-launch.

2.2. Mass Rate and Mechanical Energy in the PG1211+143 Outflow

Although the detection of high-speed winds in a substantial fraction of bright AGNs suggests most such flows have a large covering factor, PG1211+143 is one of very few in which a wide-angle flow has been demonstrated directly.Using stacked data from four *XMM-Newton* observations made between 2001 and 2007, Pounds & Reeves (2007, 2009) examined the relative strength of ionized emission and absorption spectra modeled using XSTAR to estimate the covering factor and collimation of the outflowing ionized gas. The summed pn data of PG1211+143 also show a well-defined P Cygni profile in the Fe K band (**Figure 3**), the classic signature of an outflow, with emission and absorption components of comparable equivalent width (EW). Both methods indicated a covering factor $b(=\Omega/2\pi)$ of 0.75 $\pm$ 0.25. Analysis of a *Suzaku* observation of PG1211+143 gives a very similar result (Reeves et al. 2008) with an intrinsic emission component of ∼6.5 keV and width of $\sigma \sim 250$ eV, corresponding to a flow cone of half-angle ∼50 degrees, assuming velocity broadening in a radial flow.

The outflow mass rate and mechanical energy can then be estimated; for a uniform radial outflow of velocity v, the mass rate is

$$\dot{M}_{\rm out} \simeq 4\pi b n r^2 m_{\rm p} v, \tag{9}$$

where n is the gas density at a radial distance r, and $m_{\rm p}$ is the proton mass.

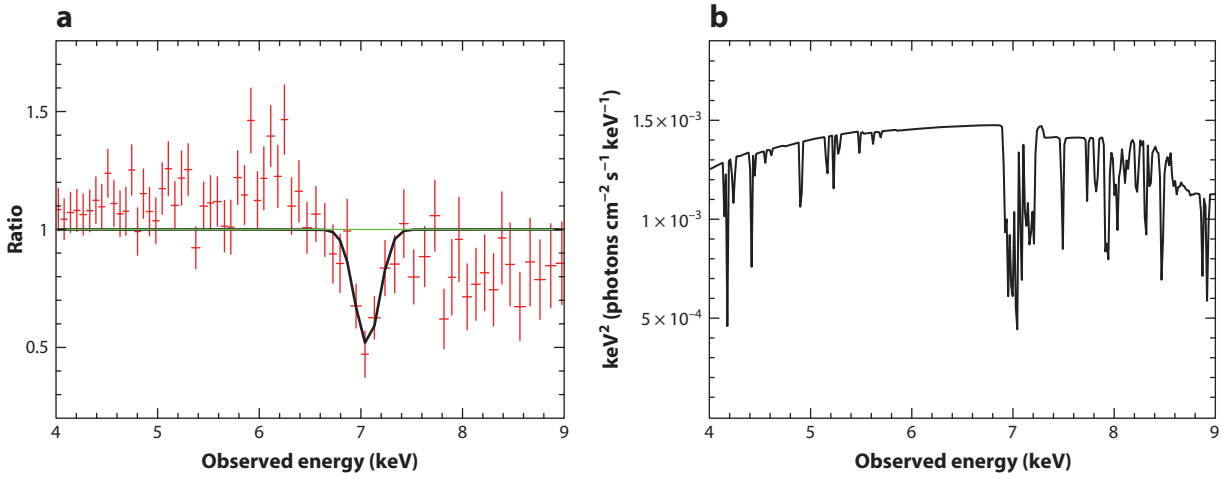

Figure 2

(*a*) A Gaussian fit to the ∼7-keV absorption feature finds a line energy of 7.06 ± 0.02 keV with (1σ) width 100 ± 30 eV. Identification with the FeXXV 1s-2p resonance line (6.70-keV rest energy) indicates an outflow velocity of $v \sim 0.12 \pm 0.01c$. (*b*) Alternative modeling with a photoionized gas over the wider 1–10 keV spectral band yields a good fit with a relatively high column density $N_{\rm H} \sim 3.2 \pm 0.7 \times 10^{23}$ cm^{-2}, moderate ionization parameter $\log \xi = 2.7 \pm 0.1$ erg cm s^{-1}, and outflow velocity of $v \sim 0.15 \pm 0.01c$. Interestingly, the FeXXV absorption line profile includes lower energy components due to the addition of one or more L-shell electrons, showing why the simple Gaussian fit gives too low a velocity.

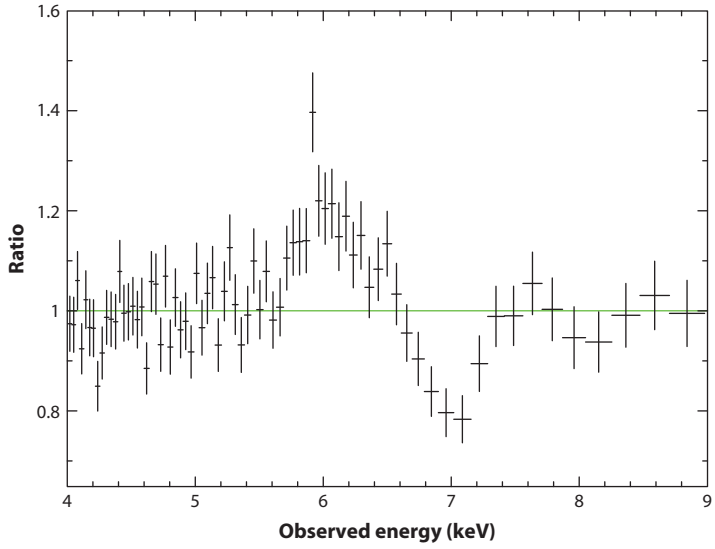

Figure 3

The P Cygni profile of FeXXV from the stacked *XMM-Newton* pn observations of PG1211+143 is characteristic of a wide-angle outflow. The comparable equivalent width of emission and blueshifted absorption components indicate the highly ionized outflow has a large covering factor. Reprinted from Pounds & Reeves (2009) with permission.

The observed values for PG1211+143 find a mass-loss rate of $\dot{M}_{\rm out} \sim 7 \times 10^{25}$ gm s^{-1} ($\sim$2.5 M$_\odot$ year^{-1}) and mechanical energy $\sim 4.5 \times 10^{44}$ erg s^{-1} (Pounds & Reeves 2009).

The mass-loss rate is comparable with the Eddington accretion rate $\dot{M}_{\rm Edd} = 1.3$ M$_\odot$ year^{-1} for an SMBH of mass $\sim 4 \times 10^7$ M$_\odot$ accreting at an efficiency of 10%, whereas the outflow mechanical energy is only $\sim$6% of the Eddington luminosity, which is close to that predicted by continuum driving (Equation 5 in Section 3 below). As noted elsewhere, an energy-flow rate like this would be more than sufficient to unbind the gas of the host galaxy bulge if all its energy were efficiently communicated.

2.3. High-Speed Winds are Common

The evidence for high-velocity winds as an important property of AGNs remained dependent on the prototype case of PG1211+143 for several years, with fast outflows in two BAL AGNs (Chartas et al. 2002) and in the most luminous low-redshift QSO PDS 456 (Reeves et al. 2003, O'Brien et al. 2005) regarded as rare objects. That began to change with the detection of a highly significant outflow of velocity $v \sim 0.1c$ in the Seyfert 1 galaxy IC 4329A (Markowitz et al. 2006) and several outflow detections in the range of $\sim$0.14c–0.2c in multiple observations of Mrk 509 (Dadina et al. 2005). A review in 2006 (Cappi 2006) listed seven non-BAL objects with outflows of $v \sim 0.1c$ and several with redshifted absorption lines.

A major step forward came with the results of an *XMM-Newton* archival search of bright AGNs by Tombesi et al. (2010), who found strong statistical evidence of blueshifted Fe K absorption lines in 15 of 42 radio-quiet objects, identification with FeXXV or XXVI resonance absorption lines implying ultrafast outflow (UFO) velocities up to $\sim$0.3c, and clustering near $v \sim 0.1c$. A later analysis based on broadband modeling with XSTAR photoionized grids (Tombesi et al. 2011) led to several revised velocities and confirmed that the outflows were typically highly ionized, with $\log \xi \sim 3 - 6$ erg cm s^{-1} and column densities in the range of $N_{\rm H} \sim 10^{22} - 10^{24}$ cm^{-2}. A similar search of the *Suzaku* data archive (Gofford et al. 2013) yielded a further group of UFO detections, finding significant absorption in the Fe K band in 20 (of 51) AGNs with velocities up to $\sim$0.3c and a flatter distribution than the *XMM-Newton* sample but a median value again of $v \sim 0.1c$.

Figure 4 brings together the results from the spectral modeling analyses of the *XMM-Newton* and *Suzaku* surveys. We follow Tombesi et al. (2011) in defining UFOs as having outflow velocities greater than 10^4 km s^{-1} to discriminate against WAs or postshock flows (Sections 2.4 and 4). The velocity plot shows a peak at $\sim$0.1c, with a tail extending to $\sim$0.3c. According to the continuum-driving black hole winds model (King & Pounds 2003), the higher velocities would imply a higher accretion efficiency suggesting the potential for such observations to provide a measure of black hole spin. Equation 22 also suggests the low-velocity tail in both the Tombesi et al. (2011) and Gofford et al. (2013) distributions could relate to primary outflows formed at a higher accretion ratio (but see Section 3.1).

Figure 4 also shows the distribution of ionization parameter and absorption column density from the surveys by Tombesi et al. (2011) and Gofford et al. (2013). The high-ionization parameter, peaking near $\log \xi \sim 4$, explains why the detection of UFOs has been almost exclusively from X-ray observations in the Fe K band, leaving open the possibility that fully ionized outflows (also consistent with continuum driving) will become detectable when the AGN luminosity (and hence ionization) falls. In assessing observational data it is important to note that for a radial outflow the observed column density is a line-of-sight integration over the flow time, dominated by the higher density at small radii, whereas the ionization parameter is governed by the current AGN luminosity. The column density, which generally lies below $N_{\rm H} \sim 10^{24}$ cm^{-2}, can vary rapidly

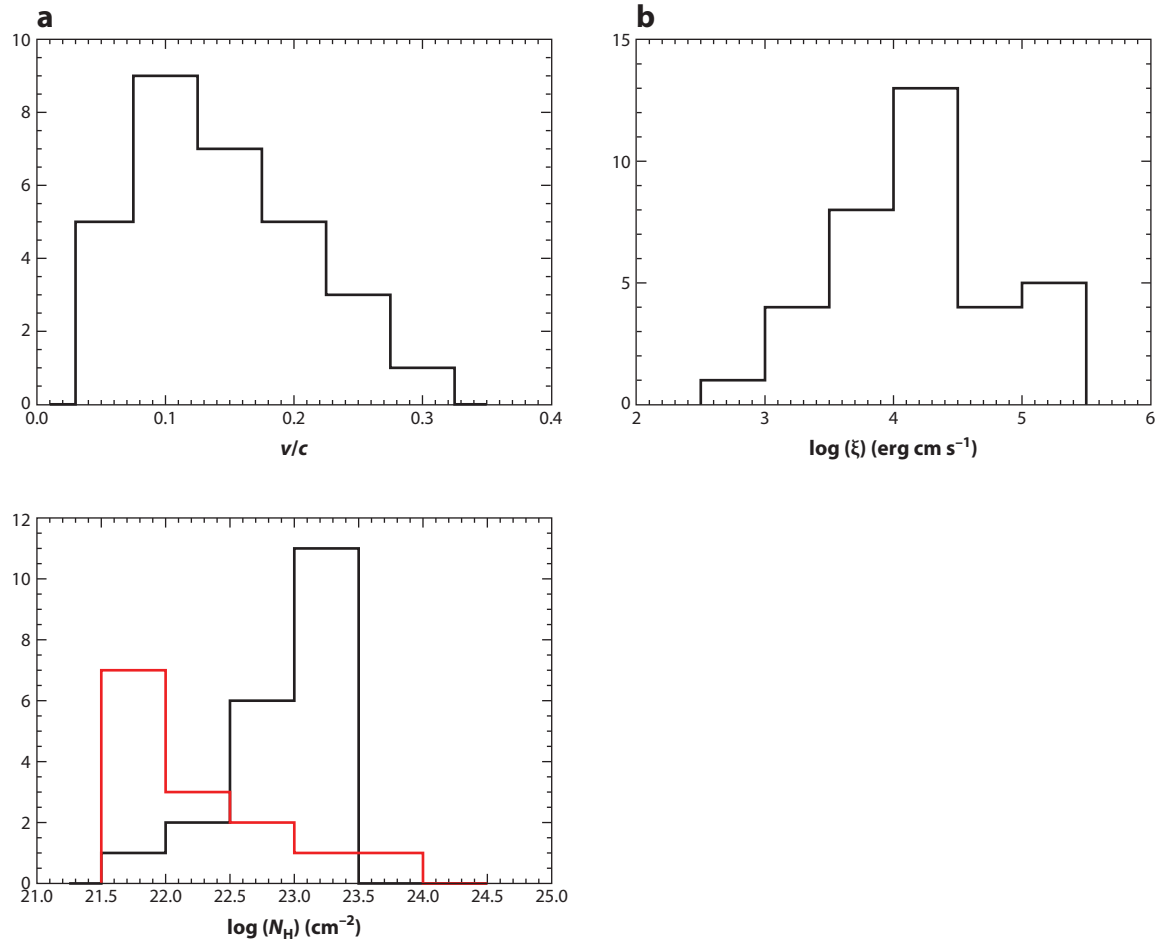

Figure 4

Distribution of (*a*) outflow velocities, (*b*) ionization parameter, and (*c*) column density obtained from modeling the individual spectra from extended observations of type 1 AGNs in the *XMM-Newton* and *Suzaku* data archives (Tombesi et al. 2011, Gofford et al. 2013). The red-lined histogram (in panel *c*) refers to lower limits in column density.

and turns out to be a powerful diagnostic of the flow history and dynamics. We return to the observability of UFOs in Section 3.3.

2.4. Evidence for a Shocked Flow

The mechanical energy in a fast wind, such as that in PG1211+143, was noted in Section 2.2 to be incompatible with the continued growth of the black hole and stellar bulge of the host galaxy, unless the coupling of wind energy to bulge gas is highly inefficient. A recent *XMM-Newton* observation of the narrow-line Seyfert galaxy NGC 4051 has provided the first evidence of a fast ionized wind being shocked, with subsequent strong cooling leading to most of the initial flow energy being lost before it can be communicated to the bulge gas. We outline a possible scenario for that event below.

NGC 4051 was found in the *XMM-Newton* archival search to have a high-velocity wind during an observation in 2002 when the source was in an unusually low state; the initial identification with FeXXVI Lyman-α by Tombesi et al. (2010) indicated a velocity of $\sim 0.15c$. In a full spectral fit (Tombesi et al. 2011), identification with FeXXV was preferred, with an increased velocity of $\sim 0.20c$. Significantly, in a 2001 observation of NGC 4051, when the X-ray flux was much higher, a strong outflow was detected at ~6,000 km s^{-1}, but no ultrafast wind was seen.

It seems that the detection of a UFO in NGC 4051 is unusually dependent on the source flux, with evidence for a high-velocity wind ($v \sim 0.13c$) again found only during periods when the ionizing continuum was low during a further *XMM-Newton* observation in 2009 (Pounds & Vaughan 2012). An additional factor may be the low redshift ($z = 0.00234$) of NGC 4051, which makes a high-velocity wind more difficult to detect with current observing facilities.

The 600-ks *XMM-Newton* observation of the Seyfert 1 galaxy NGC 4051 in 2009, extending over 6 weeks and 15 spacecraft orbits, broke new ground by finding an unusually rich absorption spectrum with multiple outflow velocities, in both RGS and EPIC spectra, up to ~9,000 km s^{-1} (Pounds & Vaughan 2011a). Interorbit variability is seen in both absorption and emission lines, with strong radiative recombination continua (RRC) and velocity-broadened resonance lines providing constraints on the dynamics and geometry of the putative postshock flow (Pounds & Vaughan 2011b, 2012).

2.5. A Self-Consistent Model for the Shocked Wind in NGC 4051

More complete modeling of both RGS and EPIC pn absorption spectra of NGC 4051 found a highly significant correlation between outflow velocity and ionization state (**Figure 5**), as expected from mass conservation in a postshock flow (King 2010, Pounds & King 2013). The additional analysis also found a range of column densities to be required by the individual XSTAR absorption components, suggesting an inhomogeneous shocked flow, perhaps with lower ionization gas clumps or filaments embedded in a more extended, lower-density, and more highly ionized flow.

Theoretical considerations suggested that a key factor in determining the structure of the postshock flow was likely to be the cooling time, as discussed in more detail in Section 4. In particular, the fate of a fast wind depends on the distance it travels before colliding with the ISM or slower-moving ejecta, with Compton cooling dominating for a shock occurring sufficiently close to the AGN continuum source.

Importantly, flux-linked variations in the ratio of FeXXV to FeXXVI absorption in the 2009 *XMM-Newton* observation (**Figure 6**) provided a measure of the Compton cooling time with the mean flow speed then determining the shell thickness of the hotter, more highly ionized flow component. The detection of strong RRC in the soft X-ray spectra furthermore suggested an increasing density in the decelerating postshock flow, with two-body cooling becoming increasingly important.

To pursue that idea, we note that at the (adiabatic) shock the free-free (thermal bremsstrahlung) and Compton cooling times are

$$t_{\rm ff} \simeq 3 \times 10^{11} \frac{T^{1/2}}{N} {\rm s} = 20 \frac{R_{16}^2}{M_7 \dot{m}} \ {\rm year} \tag{10}$$

and

$$t_{\rm C} = 10^{-4} \frac{R_{16}^2}{M_8} \ {\rm year}, \tag{11}$$

respectively (see King et al. 2011); here T and N are the postshock temperature and number density, respectively; R_{16} is the shock radius in units of 10^{16} cm; M_7 is the black hole mass in units of 10^7 $M_\odot$; and $\dot{m} \sim 1$ is the Eddington factor of the mass outflow rate.

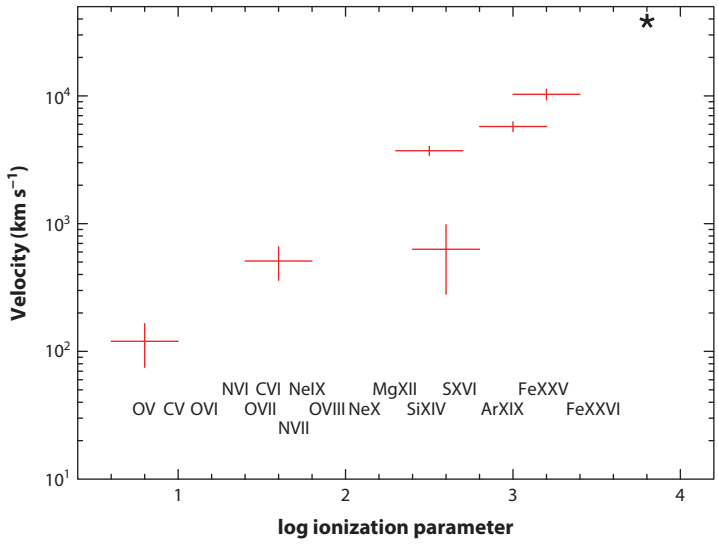

Figure 5

The outflow velocity and ionization parameters for six XSTAR photoionized absorbers used to fit the RGS and EPIC spectra of NGC 4051, together with a high-point representative of the preshock wind (*asterisk)*, show the linear correlation expected for a mass-conserved cooling flow. Reprinted from Pounds & King (2013) with permission.

After the adiabatic shock, the gas cools rapidly via inverse Compton cooling while its density rises as $N \propto T^{-1}$ (pressure is almost constant in an isothermal shock), and

$$t_{\rm ff} \propto \frac{T^{1/2}}{N} \propto T^{3/2}, \tag{12}$$

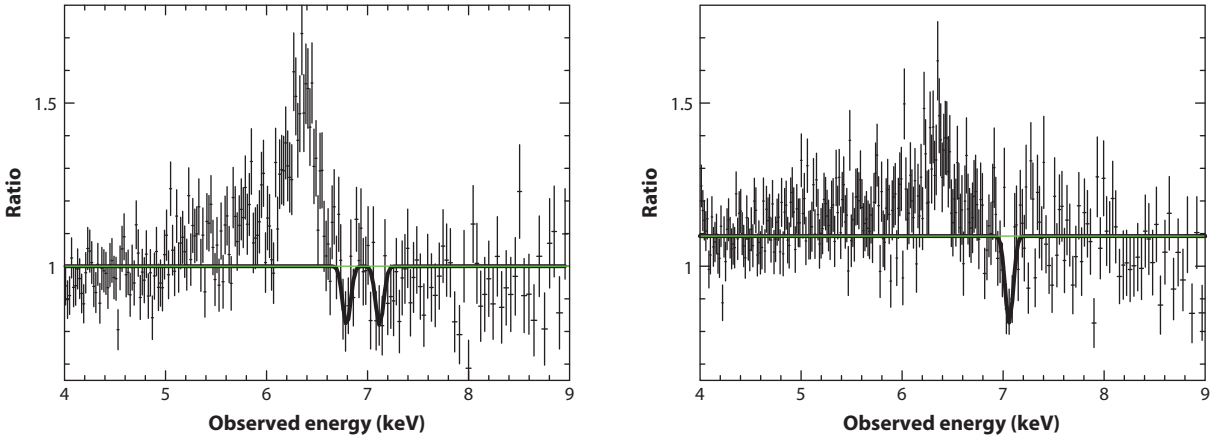

Figure 6

Fe K profiles from observations of NGC 4051 several days apart show an increased level of ionization coinciding with a hard X-ray flare (data taken from Pounds & Vaughan 2012). The ratio of resonance absorption lines of FeXXV and FeXXVI is a sensitive measure of the ionization state of the absorbing gas.

which means that the free-free cooling time decreases sharply while the Compton time does not change. Eventually free-free and other atomic two-body processes become faster than Compton when T has decreased sufficiently below the original shock temperature $T_s \sim 1.6 \times 10^{10}$ K. From Equations 10 and 11 above, this requires

$$\left(\frac{T}{T_s}\right)^{3/2} < 5 \times 10^{-5} \tag{13}$$

or

$$T < 2 \times 10^7 \text{ K}. \tag{14}$$

The temperature of ionization species forming around a few kiloelectronvolts is therefore likely to be determined by atomic cooling processes rather than Compton cooling. The strong RRC in NGC 4051 (Pounds & Vaughan 2011b, Pounds & King 2013) are direct evidence for that additional cooling, with the RRC flux yielding an emission measure for the related flow component. In particular, the onset of strong two-body cooling results in the lower-ionization, lower-velocity gas being confined to a relatively narrow region in the later stages of the postshock flow. The structure and scale of both high- and low-ionization flow regions can be deduced from the observations and modeling parameters.

For the highly ionized postshock flow, the Fe Ly-α to He-α ratio is governed by the ionizing continuum and recombination time. Significant variations in this ratio are found on inter-orbit timescales (Pounds & Vaughan 2012). An example is shown in **Figure 6**. For a mean temperature of ~1 keV and recombination coefficient of 4.6×10^{-12} cm^3 s^{-1} (Verner & Ferland 1996), the observed recombination timescale of $\sim 2 \times 10^5$ s corresponds to an average particle density of $\sim 4 \times 10^6$ cm^{-3}. Comparison with a relevant absorption column $N_H \sim 4 \times 10^{22}$ cm^{-2} from the XSTAR modeling indicates a column length scale of $\sim 10^{16}$ cm. Assuming a mean velocity of the highly ionized postshock flow of 6,000 km s^{-1}, the observed absorption length corresponds to a flow time of $\sim 1.7 \times 10^7$ s (0.6 years). Equation 11 finds a comparable cooling time for NGC 4051 at a shock radius of $R \sim 10^{17}$ cm.

For the low-ionization flow component, decay of strong RRC of NVII, CVI, and CV (Pounds & Vaughan 2011b, Pounds & King 2013) occurs over ~2–6 days. With an electron temperature from the mean RRC profile of ~5 eV and a recombination coefficient for CVI of $\sim 10^{-11}$ cm^3 s^{-1} (Verner & Ferland 1996), the observed RRC decay timescale corresponds to a (minimum) electron density of $\sim 2 \times 10^6$ cm^{-3}. A column density of 1.5×10^{21} cm^{-2} from modeling absorption in the main low-ionization flow component then corresponds to an absorbing path length of 7×10^{14} cm.

The RRC emission flux provides a consistency check on the above scaling. Assuming solar abundances and 30% of recombinations direct to the ground state, a CVI RRC flux of $\sim 10^{-5}$ photons cm^{-2} s^{-1} corresponds to an emission measure of $\sim 2 \times 10^{62}$ cm^{-3}, assuming a Tully-Fisher distance to NGC 4051 of 15.2 Mpc. With a mean particle density of $\sim 2 \times 10^6$ cm^{-3} the emission volume ($4\pi R^2 \Delta R$) is $\sim 5 \times 10^{49}$ cm^3. Assuming spherical shell geometry of the flow with fractional solid angle b, shell thickness $\Delta R \sim 7 \times 10^{14}$ cm, and shell radius $R \sim 10^{17}$ cm, the measured RRC flux is reproduced for $b \sim 0.5$.

Although this excellent agreement may be fortuitous given the approximate nature and averaging of several observed and modeled parameters, the mutual consistency of absorption and emission of the photoionized flow is encouraging. Given that only blueshifted RRC emission is seen, $b \sim 0.5$ is consistent with a wide-angle flow, visible only on the near side of the accretion disc.

Figure 7 illustrates the main features of the overall NGC 4051 outflow, a fast primary wind being shocked at a radial distance of order 0.1pc, within the zone of influence of an SMBH of

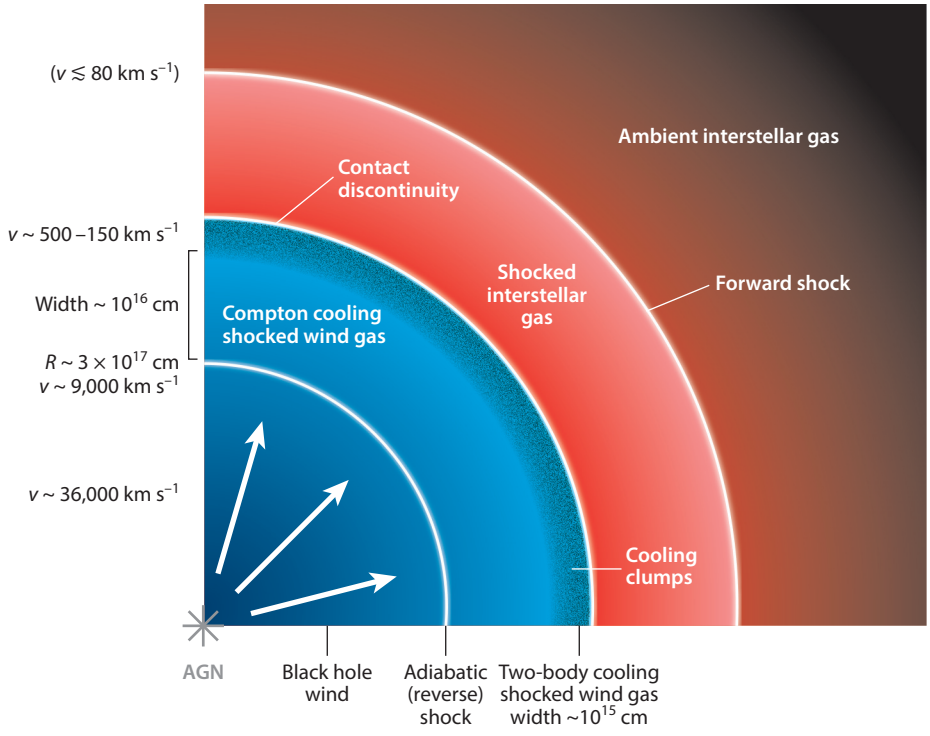

Figure 7

Schematic view (not to scale) of the shock pattern resulting from the impact of a black hole wind (*blue*) on the interstellar gas (*red*) of the host galaxy. The accreting supermassive black hole (SMBH) drives a fast wind (velocity $v \sim \eta c/\dot{m} \sim 0.1c$), whose ionization state makes it observable in X-ray absorption lines. It collides with the ambient gas in the host galaxy and is slowed in a strong shock. The inverse Compton effect from the quasar's radiation field rapidly cools the shocked gas, removing its thermal energy and strongly compressing and slowing it over a very narrow radial extent. In the most compressed gas, two-body cooling becomes important, and the flow rapidly cools and slows over an even narrower region (labeled "cooling clumps"). In NGC 4051 this region is detected in the soft X-ray spectrum, in which absorption (and emission) is dominated by the lighter metals. The cooled gas exerts the preshock ram pressure on the galaxy's interstellar gas and sweeps it up into a dense shell (snowplow). The shell's motion then drives a milder outward shock into the ambient ISM. This shock ultimately stalls unless the SMBH mass has reached the value M_σ satisfying the M–σ relation. Reprinted from Pounds & King (2013) with permission.

1.7×10^6 $M_\odot$. The initially hot gas then cools in the strong radiation field of the AGN, with a Compton cooling length determining the absorption columns of Fe and the other heavy metal ions. Two-body recombination provides additional cooling as the density rises downstream, eventually becoming dominant. Absorption (and emission) in the soft X-ray band is located primarily in this thinner outer layer of the postshock flow.

It is interesting to note that similar shocking of fast outflows provides a natural link between UFOs and the equally common WAs in AGNs (Tombesi et al. 2013). Although the onset of strong two-body cooling, resulting in the intermediate column densities being small, may explain why evidence for intermediate-flow velocities has awaited an unusually long observation of a low-mass AGN, the accumulated debris of shocked wind and ISM could be a major component of the WA. See Section 7.4.1 for a discussion.

2.6. Variability of Ultrafast Outflows

Although it is likely that powerful winds blow continuously in AGNs in rapid growth phases, it is important to note that the existing observations of UFOs are restricted to bright, low-redshift AGNs, $z \leq 0.1$, in which the X-ray fluxes are sufficient to yield high-quality spectra. Repeated observations of several bright AGNs frequently show changes in the EW of the primary Fe K absorption line.

Variability in the strength of blueshifted Fe K absorption over several years in PG1211+143 was first noted in a comparison of the initial *XMM-Newton* and *Chandra* observations (Reeves et al. 2008) and confirmed by repeated *XMM-Newton* observations (Pounds & Reeves 2009). Multiple observations of the luminous Seyfert 1 galaxy Mrk 509 (Cappi et al. 2009) found variations in both intensity and blueshift of Fe K absorption lines. The archival searches provide the most comprehensive variability data; repeated observations of several AGNs demonstrate that variability of absorption line EW over several years is common. More rapid variability in EW, over a few months, is reported in the *XMM-Newton* archive for Mrk 509, Mrk 79, and Mrk 841, with both velocity and EW change in $\leq$2 days for Mrk 766.

In addition, the hit rate of UFOs for multiply observed AGNs in the archival *XMM-Newton* data search (Tombesi et al. 2010) was relatively low, being 1 of 6 observations for NGC 4151, MCG-6-30-15 (0/5), Mrk 509 (3/5), NGC 4051 (1/2), Mrk 79 (1/3), Mrk 205 (1/3), and Mrk 290 (1/4). Overall, though 101 suitably extended observations yielded 36 narrow absorption line detections in the Fe K band, only 22 were observed at >7 keV and qualify as UFOs. Although the UFO hit rate of ~22% is a lower limit set by the sensitivity of available exposures it seems clear that the fast outflows currently being detected in low-redshift AGNs are far from continuous.

3. BLACK HOLE WINDS

3.1. The Eddington Accretion Ratio in AGNs

We have seen from Section 2 that a large fraction of observed AGNs show UFOs. Unless we view every AGN from a very particular angle (so implying a much larger total population), this must mean that these winds have large solid angles $4\pi b$ with $b \sim 1$; i.e., they are quasi-spherical. We recall that UFOs are observed to have total scalar momenta $\sim L_{\rm Edd}/c$, in which $L_{\rm Edd}$ is the Eddington luminosity of the SMBH. We can argue that, on quite general grounds, SMBH mass growth is likely to occur at accretion rates close to the value $\dot{M}_{\rm Edd} = L_{\rm Edd}/\eta c^2$, which would produce this luminosity. As we noted earlier, the Soltan (1982) relation shows that the largest SMBHs gained most of their mass by luminous accretion, i.e., during AGN phases. But the fraction of AGNs among all galaxies is small, strongly suggesting that when SMBHs grow, they are likely to do so as fast as possible. The maximum possible rate of accretion from a galaxy bulge with velocity dispersion σ is the dynamical value

$$\dot{M}_{\rm dyn} \simeq \frac{f_{\rm g}\sigma^3}{G}, \tag{15}$$

where $f_{\rm g}$ is the gas fraction. This rate applies when gas that was previously in gravitational equilibrium is disturbed and falls freely, because one can estimate that the gas mass was roughly $M_{\rm g} \sim \sigma^2 f_{\rm g} R/G$. Once this is destabilized it must fall inward on a dynamical timescale $t_{\rm dyn} \sim R/\sigma$. This gives the result in Equation 15, because $\dot{M}_{\rm dyn} \sim M_{\rm g}/t_{\rm dyn}$.

Numerically we have

$$\dot{M}_{\rm dyn} \simeq 280\sigma_{200}^3\ {\rm M}_\odot\ {\rm year}^{-1}, \tag{16}$$

where we have taken $f_g = 0.16$, the cosmological baryon fraction of all matter. We have

$$\dot{M}_{\rm Edd} = \frac{L_{\rm Edd}}{\eta c^2} = \frac{4\pi GM}{\kappa \eta c}, \tag{17}$$

where $L_{\rm Edd}$ is the Eddington luminosity and κ is the electron scattering opacity. With $\eta = 0.1$ and black hole masses M close to the observed M–σ relation (Equation 2), we find

$$\dot{M}_{\rm Edd} \simeq 4.4\, \sigma_{200}^4\ {\rm M}_\odot\ {\rm year}^{-1} \tag{18}$$

and an Eddington accretion ratio

$$\dot{m} < \frac{\dot{M}_{\rm dyn}}{\dot{M}_{\rm Edd}} \simeq \frac{64}{\sigma_{200}} \simeq \frac{54}{M_8^{1/4}}. \tag{19}$$

Thus even dynamical infall cannot produce extremely super-Eddington accretion rates on to SMBHs. But the rate (Equation 15) is already a generous overestimate, because it assumes that the infalling gas instantly loses all its angular momentum. Keeping even a tiny fraction of this instead forces the gas to orbit the black hole and form an accretion disc, which slows things down drastically. Gas spirals inward through a disc on the viscous timescale

$$t_{\rm visc} = \frac{1}{\alpha}\left(\frac{R}{H}\right)^2 \left(\frac{R^3}{GM}\right)^{1/2}, \tag{20}$$

where $\alpha \sim 0.1$ is the Shakura-Sunyaev viscosity parameter, whereas the disc aspect ratio H/R is almost constant with radius and typically close to 10^{-3} for an AGN accretion disc (e.g., Collin-Souffrin & Dumont 1990). Then $t_{\rm visc}$ approaches a Hubble time even for disc radii of only 1 pc. Gas must evidently be rather accurately channeled toward the SMBH to accrete at all, constituting a major problem for theories of how AGNs are fuelled.

Given all this, we see that though it is possible for AGN accretion rates to reach Eddington ratios $\dot{m} \sim 1$, significantly larger ones are unlikely unless the SMBH mass is far below the M–σ value appropriate to the galaxy bulge it inhabits. In other words, only relatively modest values $\dot{m} \sim 1$ of the Eddington ratio are likely in SMBH growth episodes.

Indirect evidence supporting this view comes from stellar-mass compact binary systems. The dynamical rate is relatively much larger here, as the equivalent of Equation 15 is $\dot{M} \sim v_{\rm orb}^3/G \sim M_2/P$, with $v_{\rm orb}$ the orbital velocity of a companion star of mass M_2 in a binary of period P. Accreting binaries can therefore have relatively far higher mass supply rates, usually on the thermal or nuclear timescale of the companion. These systems have highly super-Eddington apparent luminosities, probably as the result of geometric collimation (cf. King et al. 2001). But significantly there are no obvious AGN analogs of these ultraluminous X-ray sources, suggesting that Eddington ratios $\dot{m} \gg 1$ are very unusual or absent in AGNs.

3.2. Eddington Winds

Given this, we can crudely model the UFOs discussed in Section 2 as quasi-spherical winds from SMBHs accreting at modest Eddington ratios $\dot{m} = \dot{M}/\dot{M}_{\rm Edd} \sim 1$. Winds like this have electron scattering optical depth $\tau \sim 1$, measured inward from infinity to a distance of order the Schwarzschild radius $R_s = 2GM/c^2$ (cf. Equation 27 below). So on average every photon emitted by the AGN scatters about once before escaping to infinity. Because electron scattering is front-back symmetric, each photon on average gives up all its momentum to the wind, and so the total (scalar) wind momentum should be of order the photon momentum, or

$$\dot{M}_{\rm w} v \simeq \frac{L_{\rm Edd}}{c}, \tag{21}$$

where v is the wind's terminal velocity. The winds of hot stars obey relations like this. For accretion from a disc, as here, the classic paper by Shakura & Sunyaev (1973) finds a similar result at super-Eddington mass inflow rates. The excess accretion is expelled from the disc in a quasi-spherical wind.

Equation 17 now directly gives the wind terminal velocity as

$$v \simeq \frac{\eta}{\dot{m}} c \sim 0.1c\,. \tag{22}$$

From Equation 22 we get the instantaneous wind mechanical luminosity as

$$L_{\rm BH\,wind} = \dot{M}_{\rm w}\frac{v^2}{2} \simeq \frac{L_{\rm Edd}}{c}\frac{v}{2} \simeq \frac{\eta}{2}L_{\rm Edd} \simeq 0.05 L_{\rm Edd}. \tag{23}$$

This relation turns out to be highly significant (see Sections 5.3 and 7.3).

Ohsuga & Mineshige (2011) show in detail that winds with these properties (their models A and B) are a natural outcome of mildly super-Eddington accretion. In particular, their model A and B winds are predicted (cf. their figure 3) to have mechanical luminosities of $\sim 0.1 L_{\rm Edd}$, in rough agreement with Equation 23.

Compared with the original disparity $E_{\rm BH} = \eta M c^2 \sim 2 \times 10^3 E_{\rm gas}$ between black hole and bulge gas binding energies outlined in the Introduction, we now have a relation

$$E_{\rm BH\,wind} \simeq \frac{\eta^2}{2} M c^2 \sim 100 E_{\rm gas} \tag{24}$$

between the available black hole wind mechanical energy and the bulge binding energy. Although the mismatch is less severe, it still strongly suggests that the bulge gas would be massively disrupted if it experienced the full mechanical luminosity emitted by the black hole for a significant time. So the coupling of mechanical energy to the host ISM cannot be efficient all the time. This point is clearly made in Section 2 above, e.g., the penultimate sentence of Section 2.2. We show how this works in Section 4.

3.3. Observability

As we have seen in Section 2, observations frequently give the hydrogen column density $N_{\rm H}$ through a UFO wind from the X-ray absorption spectrum. We can show that this quantity determines whether a given UFO wind is observable or not. From the mass conservation equation

$$\dot{M}_{\rm w} = 4\pi b r^2 v \rho(r), \tag{25}$$

where $\rho(r)$ is the mass density, we find the equivalent hydrogen column density of the wind as

$$N_{\rm H} \simeq \int_{R_{\rm in}}^{\infty} \frac{\rho}{m_{\rm p}} {\rm d}r = \int_{R_{\rm in}}^{\infty} \frac{\dot{M}_{\rm w}}{4\pi r^2 b v} {\rm d}r = \frac{L_{\rm Edd}}{4\pi b m_{\rm p} R_{\rm in} c v^2}, \tag{26}$$

where $R_{\rm in}$ is its inner radius and $m_{\rm p}$ is the proton mass, and we have used Equation 21 at the last step. From the definition of $L_{\rm Edd}$, we find the wind electron scattering optical depth

$$\tau = N_{\rm H}\sigma_{\rm T} \simeq \frac{GM}{b v^2 R_{\rm in}}, \tag{27}$$

with $\sigma_{\rm T} \simeq \kappa m_{\rm p}$ the Thomson cross section. This shows self-consistently that the scattering optical depth τ of a continuous wind is $\sim$1 (cf. King & Pounds 2003; Equation 4) at the launch radius $R_{\rm launch} \simeq GM/bv^2 = (c^2/2bv^2)R_{\rm s} \simeq 50 R_{\rm s}$.

The measured values of N_H (Tombesi et al. 2011, Gofford et al. 2013; see **Figure 4**) are always smaller than the value $N_H \simeq 1/\sigma_T \simeq 10^{24}$ cm^{-2} for a continuous wind and actually lie in the range $N_{22} \sim 0.3-100$, where $N_{22} = N_H/10^{22}$ cm^{-2}. It is perhaps not surprising that observations do not show any UFO systems with $N_H \gtrsim 10^{24}$ cm^{-2}. These AGNs would be obscured at all photon energies by electron scattering and perhaps difficult to see. Although such systems may be common, we probably cannot detect them. To have a good chance of seeing a UFO system, we need a smaller N_H, so from Equation 27 the inner surface R_{in} of the wind must be larger than R_{launch}. This is only possible if all observed UFOs are episodic; i.e., we see them some time after the wind from the SMBH has switched off. In this sense UFOs are more like a series of sporadically launched quasi-spherical shells than a continuous outflow. The N_H value of each shell is dominated by the gas near its inner edge (cf. Equation 26), so we probably at most detect only the inner edge of the most recently launched shell. We can quantify this by setting $R_{in} = vt_{off}$, where t_{off} is the time since the launching of the most recent wind episode ended. Using Equation 27 gives

$$t_{off} = \frac{GM}{bv^3 N_H \sigma_T} \simeq \frac{3M_7}{bv_{0.1}^3 N_{22}} \text{month}, \tag{28}$$

where $v_{0.1} = v/0.1c$.

As seen from **Figure 4**, UFOs typically have $N_{22} \sim 0.3-100$, and most of the SMBH masses are $\sim 10^7$ $M_\odot$. Evidently the launches of most observed UFO winds halted weeks or months before the observation. At first sight this is surprising. The strength of the characteristic blueshifted absorption features defining UFOs is closely related to N_H. These features would be still stronger if there were UFOs with $30 < N_{22} < 100$, but none are seen. We note from Equation 28 that observing a UFO like this would require us to catch it within days of launch. Given the relatively sparse coverage of X-ray observations of AGNs, this is unlikely. So the apparent upper limit to the observed N_H may simply reflect a lack of observational coverage and implies that UFOs are shortlived.

The lower limit to N_H in the Tombesi et al. sample is also interesting. Once N_H is smaller than some critical value, any blueshifted absorption lines must become too weak to detect. The strongest are the resonance lines of hydrogen- and helium-like iron, which have absorption cross-sections $\sigma_{Fe} \simeq 10^{-18}$ cm^2. Given the abundance by number of iron as $Z_{Fe} = 4 \times 10^{-5}$ times that of hydrogen, the condition that one of these lines should have significant optical depth translates to $Z_{Fe} N_H \sigma_{Fe} > 1$ or $N_{22} > 2.5$. This is similar to the lowest observed values. From Equation 28 this means that current observations cannot detect UFO winds launched more than a few months in the past, because the blueshifted iron lines will be too weak. Even these observed UFOs should gradually decrease their N_H and become unobservable if followed for a few years. We see in Section 5 that the UFO wind typically travels $\sim 10M_7$ pc or more before colliding with the host galaxy's interstellar gas, which takes $t_{coll}/v \sim 300 M_7 v_{0.1}^{-1}$ year. Finally, a UFO may be unobservable simply because it is too strongly ionized, so that no significant N_H can be detected.

All this means that the state of the AGN seen in a UFO detection does not necessarily give a good idea of the conditions required to launch it. In particular, the AGN may be observed at a sub-Eddington luminosity, even though one might expect luminosities $\sim L_{Edd}$ to be needed for launching the UFO. This may be the reason why AGNs showing other signs of super-Eddington phenomena (e.g., narrow-line Seyfert 2 galaxies) are nevertheless seen to have sub-Eddington luminosities most of the time (e.g., NGC 4051; Denney et al. 2009). The rather short wind episodes are launched in very brief phases in which accretion is slightly super-Eddington, whereas the long-term average rate of mass gain may be significantly sub-Eddington.

In summary, it is likely that current UFO coverage is remarkably sparse. We cannot see a continuous wind at all. We can only see an episodic wind shell shortly after launch, and then only

for a tiny fraction $t_{\rm off}/t_{\rm coll} \sim 10^{-3} v_{0.1}^{-2} N_{22}^{-1}$ of its ~300–3,000-year journey to collision with the host ISM. So it seems that the vast majority of UFO wind episodes remain undetected. More AGNs must produce them than we observe, and the known UFO sources may have far more episodes than we detect.

All this has important consequences for how we interpret observations in discussing feedback. The most serious is that the most powerful form of feedback—from AGNs at the Eddington limit producing continuous winds—is probably not directly observable at all.

3.4. The Wind Ionization State and Broad-Absorption Line Quasi-Stellar Objects

The ionization parameter

$$\xi = \frac{L_i}{NR^2} \tag{29}$$

essentially fixes the ionization state of a black hole wind as seen in Section 2, and so determines which spectral lines are observed. Here $L_i = l_i L_{\rm Edd}$ is the ionizing luminosity, where $l_i < 1$ is a dimensionless parameter specified by the quasar spectrum, and $N = \rho/\mu m_{\rm p}$ is the number density of the UFO gas. We use Equations 22 and 25 to get

$$\xi = 3 \times 10^4 \eta_{0.1}^2 l_2 \dot{m}^{-2} = 3 \times 10^4 v_{0.1}^2 l_2, \tag{30}$$

where $l_2 = l_i/10^{-2}$, and $\eta_{0.1} = \eta/0.1$.

This relation shows how the wind momentum and mass rates determine the wind's ionization parameter and so its line spectrum as well as its speed v. Given a quasar spectrum L_ν, the ionization state must arrange that the threshold photon energy defining L_i and the corresponding ionization parameter ξ together satisfy Equation 30. This shows that the excitation must be high. A low-threshold photon energy (say in the IR) would imply a large value of l_2, but then Equation 30 gives a high value of ξ and so predicts the presence of very highly ionized species, physically incompatible with such low excitation.

For suitably chosen continuum spectra, Equation 30 has a range of solutions. A given spectrum may in principle allow more than one solution, the applicable one being specified by initial conditions. For a typical quasar spectrum, an obvious self-consistent solution of Equation 30 is $l_2 \simeq 1$, $\dot{m} \simeq 1, \xi \simeq 3 \times 10^4$. Here the quasar radiates the Eddington luminosity. We can also consider situations in which the quasar's luminosity has decreased after an Eddington episode but the wind is still flowing, with $\dot{m} \simeq 1$. Then the ionizing luminosity $10^{-2} l_2 L_{\rm Edd}$ in Equation 30 is smaller, implying a lower ξ. As an example, an AGN of luminosity $0.3 L_{\rm Edd}$ would have $\xi \sim 10^4$. This gives a photon energy threshold appropriate for helium-like or hydrogen-like iron (i.e., $h\nu_{\rm threshold} \sim 9$ keV). We conclude that Eddington winds from AGNs are likely to have velocities of ~$0.1c$ and show the presence of helium- or hydrogen-like iron, just as in the observations reported in Section 2. Zubovas & King (2013) show that this probably holds even for AGNs that are significantly sub-Eddington.

We can see from Equation 22 that a larger Eddington factor $\dot{m}$ is likely to produce a slower wind. From comparison with ultraluminous X-ray sources (see Section 3.1), we also expect the AGN radiation to be beamed away from a large fraction of the UFO, which should therefore be less ionized and, as a result, more easily detectable than the small fraction receiving the beamed radiation. These properties—slower, less ionized winds—characterize BAL QSO outflows, perhaps suggesting that systems with larger $\dot{m} > 1$ appear as BAL QSOs. Zubovas & King (2013) tentatively confirm this idea.

4. THE WIND SHOCK

4.1. Momentum- and Energy-Driven Flows

So far we have only studied the black hole wind. But we know that this wind must have a significant effect on the host galaxy when it impacts directly on its ISM. In this section, we model the wind and host ISM as roughly spherically symmetric and consider the effects of deviations from this simple picture later.

The pattern of the wind-ISM interaction is qualitatively identical to that of a stellar wind hitting the ISM around it (see, e.g., Dyson & Williams 1997). The black hole wind (**Figure 7**, shown in *blue*) is abruptly slowed in an inner (reverse) shock in which the temperature approaches $\sim 10^{11}$ K if ions and electrons reach equipartition (but see the discussion below). The shocked wind gas acts like a piston, sweeping up the host ISM (**Figure 7**, shown in *red*) at a contact discontinuity moving ahead of it. Because this swept-up gas moves supersonically into the ambient ISM, it drives an outer (forward) shock into it [see **Figures 7** and **8** (*top*)].

The dominant interaction here is the reverse shock slowing the black hole wind, which injects energy into the host ISM. The nature of this shock differs sharply depending on whether some form of cooling (typically radiation) removes significant energy from the hot shocked gas on a timescale shorter than its flow time. If the cooling is strong in this sense (momentum-driven flow), most of the preshock kinetic energy is lost (usually to radiation). The very rapid cooling means that the shocked wind gas is highly compressed, making the postshock region geometrically narrow (see **Figure 8**, *top*). This kind of narrow, strongly cooling region is often idealized as a discontinuity, known as an isothermal shock (cf. Dyson & Williams 1997). As momentum must be conserved, the postshock gas transmits just its ram pressure (Equation 21) to the host ISM. We shall see that this amounts to transfer of only a fraction $\sim \sigma/c \sim 10^{-3}$ of the mechanical luminosity $E_{\rm BH\ wind} \simeq 0.05 L_{\rm Edd}$ (cf. Equation 23) to the ISM. In other words, in the momentum-driven limit, only energy

$$E_{\rm mom} \sim \frac{\sigma}{c} E_{\rm BH\ wind} \sim \frac{\sigma}{c}\frac{\eta^2}{2} Mc^2 \sim 5 \times 10^{-5} Mc^2 \sim 0.1 E_{\rm gas} \tag{31}$$

is injected into the bulge ISM, that is, approximately 10% of the bulge gas binding energy $f_g M_b \sigma^2$ for black holes close to the M–σ relation (there is now an implicit factor σ_{200}^5/M_8 on the right-hand side). Thus momentum-driven flows do not threaten the bulge's integrity. Indeed we shall see that they never interact with most of it, so there is no danger that the black hole will drive away the gas and suppress accretion. A momentum-driven regime is a stable environment for black hole mass growth.

In the opposite limit in which cooling is negligible, the postshock gas retains all the mechanical luminosity

$$E_{\rm wind} \simeq 0.05 L_{\rm Edd} \simeq 100 E_{\rm gas} \tag{32}$$

(cf. Equation 24) thermalized in the shock and instead expands adiabatically into the ISM. The postshock gas is now geometrically extended (see **Figure 8**, *bottom*), unlike the momentum-driven (isothermal) case. This energy-driven flow is much more violent than momentum-driven flow. The estimate (Equation 32) is for a black-hole mass near the M–σ relation; a hole with mass that is a factor of 100 below this would already unbind the bulge in doubling its mass. Unless the shock interaction is markedly aspherical, a black hole in an energy-driven environment is unlikely to reach observed SMBH masses.

Given these starkly different outcomes we must decide under what conditions we have momentum- or energy-driven outflows. Simple estimates immediately show that ordinary atomic two-body processes have no significant effect in cooling the wind shock. But the wind shock is

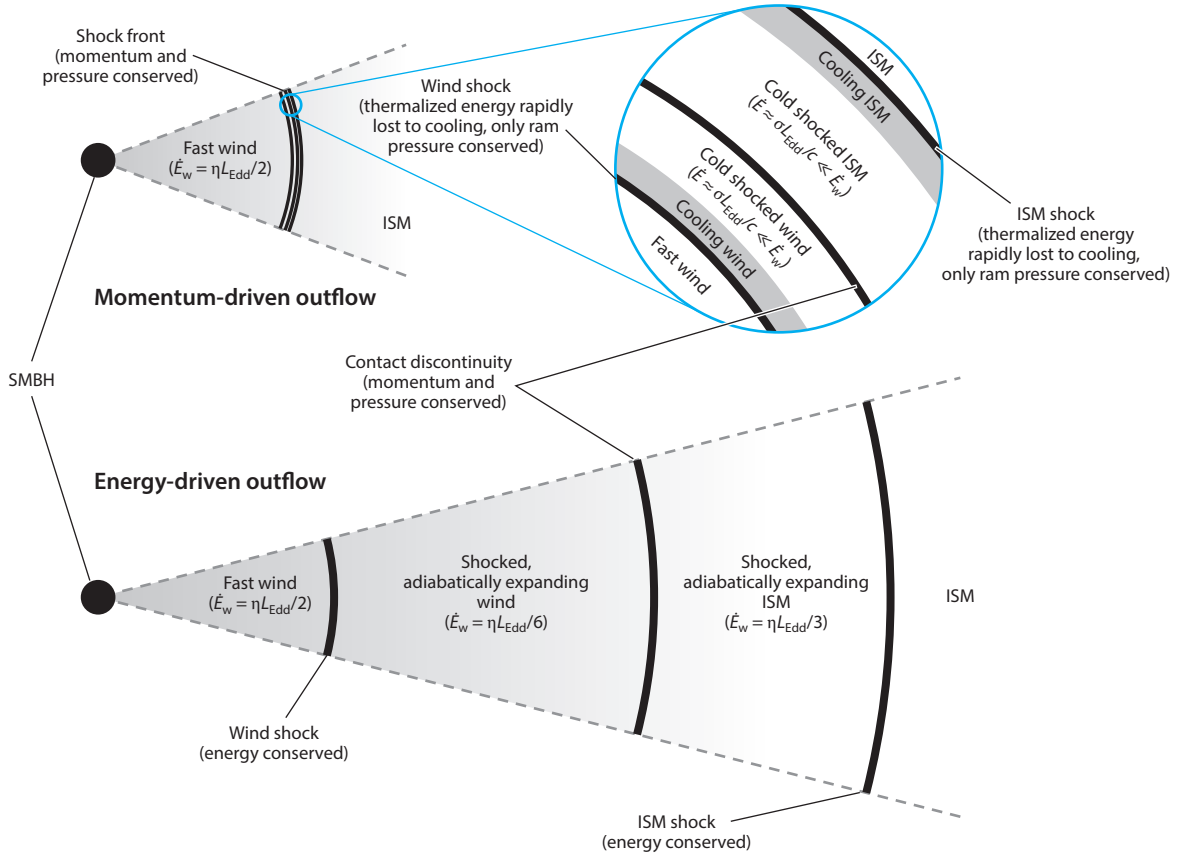

Figure 8

Schematic picture of (*top*) momentum-driven and (*bottom*) energy-driven outflows. In both cases a fast wind (velocity $\sim 0.1c$) impacts the interstellar gas of the host galaxy, producing an inner reverse shock that slows the wind and an outer forward shock that accelerates the swept-up gas. In the momentum-driven case, the shocks are very narrow and rapidly cool to become effectively isothermal; only the ram pressure is communicated to the outflow, leading to very low kinetic energy $\sim (\sigma/c)L_{Edd}$. This situation corresponds to the ultrafast outflows discussed in Section 2. In an energy-driven outflow, the shocked regions are much wider and do not cool; they expand adiabatically, communicating most of the kinetic energy of the wind to the outflow (in simple cases divided in a ratio of approximately 1:2 between the shocked wind and the swept-up gas). The outflow radial momentum flux is therefore greater than that of the wind. Momentum-driven conditions hold for shocks confined to within $\sim$1 kpc of the AGN and establish the $M-\sigma$ relation (Equation 42) (King 2003, 2005). Once the supermassive black hole (SMBH) mass attains the critical $M-\sigma$ value, the shocks move further from the AGN, and the outflow becomes energy driven. This produces the observed large-scale molecular outflows that probably sweep the galaxy clear of gas, as in the clear-out phase discussed in Section 5.3. Reprinted from Zubovas & King (2012a) with permission.

exposed to the radiation field of an Eddington-accreting SMBH. This has a characteristic temperature of no more than $\sim 10^7$ K, which is far lower than the wind's immediate postshock temperature of $\sim 10^{10}-10^{11}$ K. Electrons in the postshock gas lose energy to these photons through the inverse Compton effect (cf. Ciotti & Ostriker 1997) at a rate dependent on the radiation density. For wind shocks close to the SMBH, the accretion radiation field is intense enough that this effect cools the postshock wind gas in less than the momentum-driven flow time $\sim R/\sigma$ (see below), and we are self-consistently in the momentum-driven regime provided that the postshock gas is in

equipartition, i.e., that electron and ion temperatures remain effectively equal. (We consider this further in Section 4.2 below.)

For shocks at larger radii R the radiation energy density decreases as R^{-2}, increasing the cooling time as R^2. The flow time increases only as R, so for R greater than the critical cooling radius

$$R_C \sim 500 M_8^{1/2} \sigma_{200} \text{ pc}, \tag{33}$$

(King 2003, 2005; King et al. 2011; Zubovas & King 2012b) the cooling time is longer than the flow time, and the flow must be energy driven. Thus we have the general result that momentum-driven flows are confined to a small region $R < R_C$, whereas energy-driven flows must be large-scale. This is just as one would expect, given that a momentum-driven flow allows stable black hole mass growth, whereas an energy-driven one is likely to expel most of the bulge gas.

It is plausible then that the observed UFO winds can lead to momentum driving through strongly cooled shocks close to the SMBH. In this regime all but a small fraction of the mechanical luminosity (Equation 23) of the black hole wind is eventually radiated away as an inverse Compton continuum with characteristic photon energy ~1 keV. Pounds & Vaughan (2011a) report a possible detection of this spectral component in the Seyfert 1 galaxy NGC 4051. As required for consistency, the luminosity of this component is comparable with the expected mechanical luminosity (Equation 23) of the wind in that system.

Cooling shocks are called isothermal because the gas temperature rapidly returns to something like its preshock value. Momentum conservation requires that the gas is also strongly slowed and compressed as it cools. So the postshock velocity of the X-ray–emitting gas should correlate with its temperature (or roughly, ionization) while Compton cooling is dominant. Once this has compressed the gas sufficiently, two-body processes, such as free-free and bound-free emission, must begin to dominate, because they go as the square of the density, and their cooling times decrease with temperature also (Pounds & King 2013). Section 2.5 above shows that there is direct observational evidence for both of these effects in NGC 4051. So this object (uniquely) shows three signatures of a cooling shock: an inverse Compton continuum, an ionization-velocity correlation, and the appearance of two-body processes in the spectrum.

4.2. Shock Cooling

Cooling (or the lack of it) has a defining effect on the physics of the interaction between the black hole wind and the host ISM, so we must check the simple picture above. In particular the inverse Compton effect acts only on electrons, but the energy of the postshock gas is initially almost all in its ions. We assumed above that the electron and ion temperatures quickly come into equipartition after the shock, allowing the inverse Compton effect to drain energy from the ions.

This assumption can be questioned. Faucher-Giguère & Quataert (2012) show that if the only process that is coupling electrons and ions is Coulomb collisions, then there is a significant parameter space in which equipartition does not occur. However, they do not rule out substantial momentum-conserving phases. An important consideration here is that many processes other than direct Coulomb collisions may rapidly equilibrate electron and ion temperatures. Faucher-Giguère & Quataert (2012) attempt to put limits on the incidence of such collisionless coupling by appealing to observations of the solar wind, but this is an area of considerable physical uncertainty.

Another way of using observations to decide if shock cooling is effective is to look for the inverse Compton spectral component that directly reveals the cooling. Bourne & Nayakshin (2013) argue that the apparent lack of such a component in most AGN spectra rules out cooling shocks. But we recall from Section 3.3 that the coverage of UFOs is extremely sparse. Actually observing a collision, and the inverse Compton emission, is inevitably a very rare event. It appears that

observationally ruling out Compton shock cooling is so far inconclusive. In addition we stress that, as discussed in Sections 2.5 and 4.1, NGC 4051 does show three separate signatures of shock cooling.

5. THE M–σ RELATION

5.1. Reaching M–σ: The Momentum-Driven Phase

We are now equipped to discuss the impact of a UFO on the host interstellar gas. We already noted that the wind impact implies a pair of shocks on each side of the contact discontinuity between the wind and the host ISM. Initially the wind shock is close to the hole, and we assume that inverse Compton cooling from the AGN radiation field cools it rapidly and puts the flow in the momentum-driven regime. The region of gas between the wind shock and the contact discontinuity, where it impacts and sweeps up the host ISM, is very narrow (cf. **Figure 8**, *top*). The outer shock accelerating the ISM is also strongly cooled, so that the snowplow region of swept-up ISM is narrow as well. Now we can treat the whole region between the inner and outer shocks as a single narrow, outward-moving gas shell, whose mass grows as it sweeps up the host ISM (see **Figure 8**).

As a simple model of a bulge, we assume that the matter of the host galaxy is distributed with an isothermal profile of velocity dispersion σ, with mass density

$$\rho(r) = \frac{f_{\rm g}\sigma^2}{2\pi G r^2}, \tag{34}$$

so that the mass within radius R is

$$M(R) = \frac{2\sigma^2 R}{G}. \tag{35}$$

A distribution like this is expected if the bulge results from mergers. For a roughly constant gas fraction $f_{\rm g}$, the mass of the narrow swept-up gas shell at radius R is

$$M_{\rm g}(R) = \frac{2 f_{\rm g}\sigma^2 R}{G} \tag{36}$$

so that the shell has the equation of motion

$$\frac{\rm d}{{\rm d}t}[M_{\rm g}(R)\dot{R}] + \frac{GM_{\rm g}(R)[M + M(R)]}{R^2} = \frac{L_{\rm Edd}}{c}, \tag{37}$$

where M is the SMBH mass. From Equations 35 and 36 and the definition of $L_{\rm Edd}$ (Equation 17) this simplifies to

$$\frac{\rm d}{{\rm d}t}(R\dot{R}) + \frac{GM}{R} = -2\sigma^2\left(1 - \frac{M}{M_\sigma}\right), \tag{38}$$

where

$$M_\sigma = \frac{f_{\rm g}\kappa}{\pi G^2}\sigma^4. \tag{39}$$

Multiplying through by $R\dot{R}$ and integrating once gives the first integral

$$R^2\dot{R}^2 = -2GMR - 2\sigma^2\left[1 - \frac{M}{M_\sigma}\right]R^2 + \text{constant}. \tag{40}$$

For large R, we have

$$\dot{R}^2 \simeq -2\sigma^2\left[1 - \frac{M}{M_\sigma}\right], \tag{41}$$

which has no solution for $M < M_\sigma$. Physically this says that if the SMBH mass is below M_σ, the swept-up shell of interstellar gas cannot reach large radius because the Eddington thrust of the black hole wind is too small to lift its weight against the galaxy bulge potential. The SMBH cannot remove the gas from its surroundings and goes on accreting. Any gas shell it drives outward eventually becomes too massive and tends to fall back and probably fragment. This is likely to stimulate star formation in the shell remnants.

The precise value of M_σ depends on the average gas fraction f_g. For a protogalaxy forming at high redshift, we expect $f_g = \Omega_{\rm baryon}/\Omega_{\rm matter} \simeq 0.16$ (Spergel et al. 2003). Galaxies forming at later times may have larger f_g if they have gained a lot of gas or smaller f_g if they have been largely swept clear of gas or have turned a lot of their gas into stars. With the gas fraction f_g fixed at the cosmological value $f_c = 0.16$, the expression

$$M_\sigma = \frac{f_g \kappa}{\pi G^2}\sigma^4 \simeq 3.2 \times 10^8\ {\rm M}_\odot \sigma_{200}^4 \tag{42}$$

is remarkably close to the observed relation of Equation 2, even though it contains no free parameter. (We shall see in Section 5.5 why observations tend to give an exponent of σ slightly larger than the value of four derived here.) This agreement strongly suggests that SMBH growth stops at this point, although we must do some more work to show this (see Sections 5.2, 5.3, and 5.5 below).

The derivation here used the simplest possible description of a galaxy spheroid as an isothermal sphere (cf. Equation 35). We should ask whether things change significantly if the galaxy bulge is more complicated than this. If the potential is spherically symmetric but the cumulative mass $M(R)$ is not simply linear in R, we still get a first integral of the equation of motion (Equation 37) simply by multiplying through by $M(R)\dot{R}$, giving the condition for a swept-up, momentum-driven shell to reach large radii. Relations very like Equation 42 emerge in each case, so we expect qualitatively similar behavior. McQuillin & McLaughlin (2012) show this explicitly for three widely used density distributions (Hernquist 1990; Navarro et al. 1996, 1997; and Dehnen & McLaughlin 2005). The results are, in practice, scarcely distinguishable from Equation 42.

Whatever the bulge geometry, the black hole always communicates its presence only through the ram pressure of its wind, so we are always dealing with strongly radial forces in the solid angles exposed to this wind (this is not true of gas pressure, as we shall see in Section 5.3). It is likely that the orientation of the accretion disc with respect to the host galaxy changes with each new episode of accretion (so-called chaotic accretion; King & Pringle 2006, King et al. 2008), which tends to isotropize the long-term effect of momentum feedback. Together with the sudden huge increase in the spatial scale as the critical black hole mass is reached (see the next Section), this may explain why the simple spherically symmetric prediction (Equation 42) seems to give a surprisingly accurate estimate of the critical mass.

5.2. What Happens When $M = M_\sigma$?

The result (Equation 42) is so close to observations that it strongly suggests that feedback somehow cuts off the growth of the black hole at a mass very close to this value. Some feeding may continue from gas in the immediate vicinity of the hole that is too dense to be affected by the ram pressure of the black hole wind. A thin accretion disc has this property, for example, but cannot have a gas mass larger than $\sim (H/R)M \ll M$ without fragmenting and forming stars.

But we still have to explain precisely how the gas is expelled. For example, one may worry that although momentum driving can push the ISM away and inhibit central accretion, some kind of infall and SMBH accretion may restart shortly after momentum driving is switched off, perhaps leading to alternating stages of quiescence and growth and eventually to masses far above M_σ.

Observations show that black hole accretion occurs preferentially in gas-rich galaxies (cf. Vito et al. 2014), so it seems that the black hole must largely clear the galaxy bulge of gas to terminate its growth. We shall see later (Section 7) that if no other process than momentum driving operated, this requirement would indeed lead to black hole masses significantly larger than M_σ, which is in conflict with observation.

This last point means that the way black hole growth influences the host galaxy must change radically when $M = M_\sigma$. It is straightforward to see why it should. We saw above that for $M < M_\sigma$ the Eddington thrust cannot push the wind shocks to large R. As a result the wind shock remains efficiently Compton cooled, enforcing momentum driving. It follows that the SMBH cannot stifle its own growth if $M < M_\sigma$. But all this changes once the SMBH exceeds the critical mass (Equation 42). Now even for a very small increment $[O(R_{\rm inf}/R_C) \sim 10^{-2}]$ of M above M_σ, a momentum-driven shell can reach the critical radius $R_{\rm C}$. Crucially, this means that the wind shocks are no longer efficiently cooled. They become energy driven. The shocked wind gas can now use all its energy to push the interstellar gas as it expands into the host bulge. This motion becomes explosive and rapidly reaches kiloparsec length-scales, comparable with the size of the bulge itself, rather than the much smaller (parsec) scales of the momentum-driven phase.

So the real significance of the M–σ relation is that it marks the point at which outflows undergo a global transition from being momentum to energy driven.

5.3. Clearing Out a Galaxy: The Energy-Driven Phase

We know from Equation 24 that an energy-driven outflow has more than enough energy to remove the interstellar gas entirely. This presumably suppresses further SMBH growth. Here we examine how this works in detail.

Once $M > M_\sigma$, the outflow geometry changes completely (see **Figure 8**). The shocked wind region is no longer narrow (as in **Figure 8**, *top*) but large and expanding because of its strong thermal pressure (**Figure 8**, *bottom*). The shocked wind's thermal expansion pushes its shock inward where it must hover at the cooling radius $R_{\rm C}$ (Zubovas et al. 2013). If it tries to move within $R_{\rm C}$, momentum driving instantly pushes it out again (remember $M > M_\sigma$).

The shocked wind rapidly evens out its internal pressure as it expands at its sound speed $\sim 0.06c$, so we take this pressure P as uniform over this region (but changing with time). The contact discontinuity at the outer edge of the shocked wind sweeps up the surrounding shocked ISM as before, but now has the equation of motion

$$\frac{\mathrm{d}}{\mathrm{d}t}[M_{\rm g}(R)\dot{R}] + \frac{GM_{\rm g}(R)M(R)}{R^2} = 4\pi R^2 P, \tag{43}$$

where the pressure P is much larger than the ram pressure ρv^2 appearing in Equation 37. In the second term on the left-hand side, we have neglected the contribution $GM_{\rm g}M/R^2$ of the black hole gravity, as $R \gg R_{\rm C} > R_{\rm inf}$. To make the problem determinate, we need the energy equation. (This did not appear explicitly in the momentum-driven case because it was equivalent to the defining condition that all the wind energy not associated with the ram pressure was rapidly lost to radiation.) Here the energy equation constrains the pressure P by specifying the rate at which energy is fed into the shocked gas, minus the rate of $P\mathrm{d}V$ working on the ambient gas and against gravity:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{4\pi R^3}{3}\cdot\frac{3}{2}P\right] = \frac{\eta}{2}L_{\rm Edd} - P\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{4\pi}{3}R^3\right] - 4f_{\rm g}\frac{\sigma^4}{G}. \tag{44}$$

We take a specific heat ratio $\gamma = 5/3$, use Equation 23 for the energy input from the outflow, and apply Equation 35 to simplify the gravity term $GM(R)M(R)/R^2$. Now we use Equation 43

to eliminate P from Equation 44 and replace the gravity terms as before using the isothermal expression for $M(R)$. We take the AGN luminosity as $lL_{\rm Edd}$ to allow for small deviations from the Eddington value. This gives

$$\frac{\eta}{2} l L_{\rm Edd} = \dot{R}\frac{\mathrm{d}}{\mathrm{d}t}[M(R)\dot{R}] + 8 f_{\rm g}\frac{\sigma^4}{G}\dot{R} + \frac{\mathrm{d}}{\mathrm{d}t}\left\{\frac{R}{2}\dot{R}\frac{\mathrm{d}}{\mathrm{d}t}[M(R)\dot{R}] + 2 f_{\rm g}\frac{\sigma^4}{G}R\right\}, \tag{45}$$

and therefore

$$\frac{\eta}{2} l L_{\rm Edd} = \frac{2 f_{\rm g}\sigma^2}{G}\left\{\frac{1}{2}R^2\ddot{R} + 3R\dot{R}\ddot{R} + \frac{3}{2}\dot{R}^3\right\} + 10 f_{\rm g}\frac{\sigma^4}{G}\dot{R}. \tag{46}$$

This describes the motion of the interface (contact discontinuity in **Figure 8**, *bottom*) between wind and interstellar gas in the energy-driven case, replacing Equation 37 in the momentum-driven case.

The energy-driven regime applies as soon as the SMBH mass reaches M_σ, and we shall see that the host ISM is now quickly removed. We assume $M = M_\sigma$ in $L_{\rm Edd}$ and see that Equation 46 has a solution $R = v_{\rm e}t$ with

$$2\eta l c = 3\frac{v_{\rm e}^3}{\sigma^2} + 10 v_{\rm e}. \tag{47}$$

The assumption $v_{\rm e} \ll \sigma$ leads to a contradiction ($v_{\rm e} \simeq 0.01c \gg \sigma$), so

$$v_{\rm e} \simeq \left[\frac{2\eta l \sigma^2 c}{3}\right]^{1/3} \simeq 925 l^{1/3}\sigma_{200}^{2/3}\,{\rm km\,s^{-1}}. \tag{48}$$

This solution is an attractor. **Figure 9** shows that all solutions quickly converge to it, regardless of initial conditions. Physically, its meaning is that if shock cooling is ineffective, the extra gas pressure accelerates the previously momentum-driven gas shell to this new higher velocity. **Figure 9** also confirms that if the driving by the AGN switches off when the contact discontinuity is at radius R_0, it decelerates as predicted by the analytic solution of Equation 46 with $L_{\rm Edd} = 0$ found by King et al. (2011):

$$\dot{R}^2 = 3\left(v_{\rm e}^2 + \frac{10}{3}\sigma^2\right)\left(\frac{1}{x^2} - \frac{2}{3x^3}\right) - \frac{10}{3}\sigma^2 \tag{49}$$

where $x = R/R_0 \geq 1$. Noting that $v_{\rm e}$ depends only weakly (as $v_{\rm e} \sim l^{1/3}$) on the luminosity, these results show that fluctuations—or even the intermittent disappearance—of the AGN luminosity have almost no effect on the outflow once it has started, because the flow still has a large reservoir of thermal energy available for driving. In particular an outflow can persist long after the central AGN has turned off, and the real agency driving an observed outflow may have been an AGN even if it is currently observed to be weak or entirely absent.

The solutions of Equations 48 and 49 describe the motion of the contact discontinuity in which the shocked wind encounters swept-up interstellar gas (see **Figures 8** and **9**). This interface is strongly Rayleigh-Taylor (RT) unstable, because the shocked wind gas is highly expanded and has much lower mass density than the swept-up interstellar gas outside it, meaning that we have a light fluid underneath a heavy one. The RT instability leads to strong overturning motions even on small scales, so it is difficult to handle numerically. Deductions concerning the mean velocity and energy of the outflow, and its average spatial scale $R(t)$, are likely to be believable and agree closely with observations (see below), but we should be very cautious about results depending strongly on the detailed nature of the interface between the shocked wind and the swept-up ISM. The RT instability is probably the reason that the high-speed ($\sim$1,000 km s^{-1}) outflows with prodigious mass rates we predict here are generally seen with much of the outflowing gas in molecular form. Apparently the interstellar gas entering the forward shock is efficiently cooled

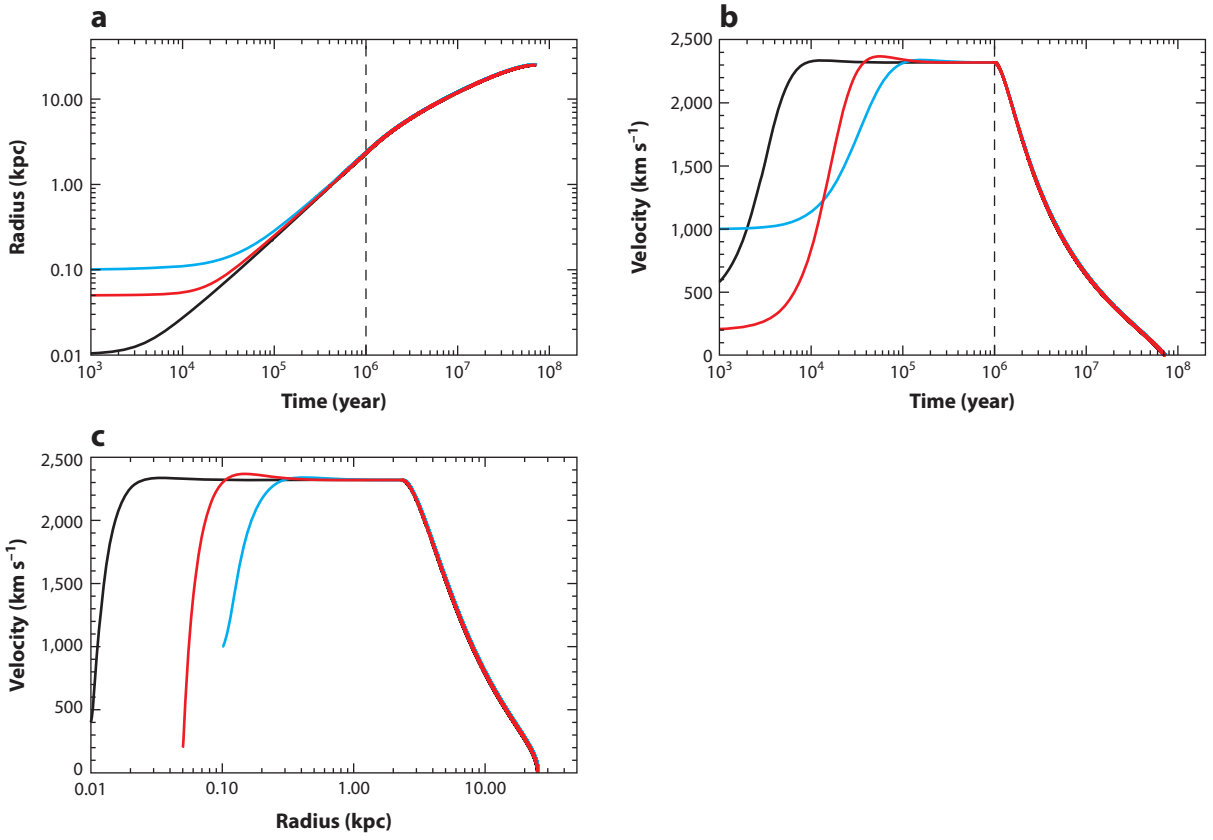

Figure 9

Evolution of an energy-driven shock pattern for the case $\sigma = 200$ km s^{-1}, $f_g = 10^{-2}$ computed numerically from the full Equation 46. (*a*) Radius versus time, (*b*) velocity versus time, (*c*) velocity versus radius. The curves refer to different initial conditions: *black line*, $R_0 = 10$ pc and $v_0 = 400$ km s^{-1}; *blue line*, $R_0 = 100$ pc and $v_0 = 1 \times 10^3$ km s^{-1}; and *red line*, $R_0 = 50$ pc and $v_0 = 200$ km s^{-1}. All these solutions converge to the attractor (Equation 48). The vertical dashed line marks the time $t = 10^6$ year when (for this case) the quasar driving is switched off. All solutions then follow the analytic solution (Equation 49). A case in which the quasar remains on for a Salpeter time $\sim 4 \times 10^7$ year would sweep the galaxy clear of gas. Reprinted from King et al. (2011) with permission.

by two-body radiation processes. A preliminary analysis (Zubovas & King 2014) suggests that the interstellar gas overtaken by the forward shock is likely to have a multiphase structure. Most of it cools all the way from the shock temperature $\sim 10^7$ K back to low temperatures, ending in largely molecular form, even though it is entrained in an outflow with the ~1,000-km s^{-1} velocity of the forward shock. But cooling is affected by the topology of the gas flow and the total area of interfaces of different gas phases. A full numerical calculation of this is currently impossible, so for the time being we can only make comparison with simple estimates.

The mass outflow rate is fixed by how fast the outer shock overtakes the ISM and entrains new interstellar gas ahead of the contact discontinuity. The ISM ahead of the shock is at rest, so this runs into it at a speed giving a velocity jump by a factor $(\gamma+1)/(\gamma-1)$ in the shock frame (where γ is the specific heat ratio: see, e.g., Dyson & Williams 1997 for a derivation). This fixes its velocity as

$$v_{\rm out} = \frac{\gamma+1}{2}\dot{R} \simeq 1.23 \times 10^3 \sigma_{200}^{2/3} \left(\frac{l f_c}{f_g}\right)^{1/3} \ {\rm km\ s^{-1}} \qquad (50)$$

(where we have used $\gamma = 5/3$ in the last form, and $f_c \simeq 0.16$ is the cosmological value of f_g). This implies a shock temperature of order 10^7 K for the forward (ISM) shock (as opposed to $\sim 10^{10-11}$ K for the wind shock). Because the outer shock and the contact discontinuity were very close together as energy-driven flow took over from momentum-driven flow (see **Figure 8**), this means that the outer shock is at

$$R_{\rm out}(t) = \frac{\gamma+1}{2} R(t) = \frac{\gamma+1}{2} v_e t. \tag{51}$$

This gives the mass outflow rate as

$$\dot{M}_{\rm out} = \frac{{\rm d}M(R_{\rm out})}{{\rm d}t} = \frac{(\gamma+1) f_g \sigma^2}{G} \dot{R}. \tag{52}$$

For comparison the mass rate of the black hole wind, assuming $M = M_\sigma$, is

$$\dot{M}_w \equiv \dot{m}\dot{M}_{\rm Edd} = \frac{4 f_c \dot{m} \sigma^4}{\eta c G}. \tag{53}$$

This is much smaller than the outflow rate $\dot{M}_{\rm out}$ it drives, so we define a mass-loading factor as the ratio of the mass flow rate in the shocked ISM to that in the wind:

$$f_L \equiv \frac{\dot{M}_{\rm out}}{\dot{M}_w} = \frac{\eta(\gamma+1)}{4\dot{m}} \frac{f_g}{f_c} \frac{\dot{R}c}{\sigma^2}. \tag{54}$$

Then we have

$$\dot{M}_{\rm out} = f_L \dot{M}_w = \frac{\eta(\gamma+1)}{4} \frac{f_g}{f_c} \frac{\dot{R}c}{\sigma^2} \dot{M}_{\rm Edd}. \tag{55}$$

If the AGN radiates at luminosity $\sim L_{\rm Edd}$, we have $\dot{R} = v_e$, and Equation 48 gives

$$f_L = \left(\frac{2\eta c}{3\sigma}\right)^{4/3} \left(\frac{f_g}{f_c}\right)^{2/3} \frac{l^{1/3}}{\dot{m}} \simeq 460 \sigma_{200}^{-2/3} \frac{l^{1/3}}{\dot{m}} \tag{56}$$

and

$$\dot{M}_{\rm out} \simeq 4.06 \times 10^3 \sigma_{200}^{10/3} l^{1/3}\ {\rm M_\odot\ year^{-1}} \tag{57}$$

for typical parameters, $f_g = f_c$ and $\gamma = 5/3$. The total gas mass in the bulge is roughly $M_g \sim 10^3 f_g M_\sigma$ (from Equation 1). Clearly if the outflow persists for a time $t_{\rm clear} \sim M_g/\dot{M}_{\rm out} \sim 1 \times 10^7 \sigma_{200}^{2/3} l^{-1/3}$ year it will sweep away a large fraction of the galaxy's gas. The precise outflow duration needed for this depends on both the type and the environment of the galaxy, in practice leading to three parallel but slightly offset M–σ relations (see Section 5.5 below).

Equations 50 and 57 give

$$\frac{1}{2}\dot{M}_w v^2 \simeq \frac{1}{2}\dot{M}_{\rm out} v_{\rm out}^2. \tag{58}$$

Thus most of the wind kinetic energy ultimately goes into the mechanical energy of the outflow, as we would expect for energy driving. The continuity relations across the contact discontinuity show that if the quasar is still active, the shocked wind retains one-third of the total incident wind kinetic energy $\dot{M}_w v^2/2$, giving two-thirds to the swept-up gas represented by $\dot{M}_{\rm out}$.

Equation 58 means that the swept-up gas must have a scalar momentum rate greater than the Eddington value $L_{\rm Edd}/c$, because we can rewrite it as

$$\frac{\dot{P}_w^2}{2\dot{M}_w} \simeq \frac{\dot{P}_{\rm out}^2}{2\dot{M}_{\rm out}}, \tag{59}$$

where the $\dot{P}$s are the momentum fluxes. With $\dot{P}_{\rm w} = L_{\rm Edd}/c$, we have

$$\dot{P}_{\rm out} = \dot{P}_{\rm w}\left(\frac{\dot{M}_{\rm out}}{\dot{M}_{\rm w}}\right)^{1/2} = \frac{L_{\rm Edd}}{c} f_{\rm L}^{1/2} \simeq 20\frac{L_{\rm Edd}}{c}\sigma_{200}^{-1/3} l^{1/6}. \quad (60)$$

Observations of molecular outflows consistently show $\dot{M}_{\rm out} v_{\rm out} > L_{\rm Edd}/c$, and in particular Cicone et al. (2014) find that momentum rates $20L/c$ are common. This is an inevitable consequence of mass loading ($f_{\rm L} > 1$). These high momentum rates are important, as they are probably the way that the galaxy resists the accretion that cosmological simulations suggest still continues at large scales (Costa et al. 2014).

Recent IR observations show abundant evidence for molecular outflows with speeds and mass rates similar to that in Equations 50 and 57. Feruglio et al. (2010), Rupke & Veilleux (2011), and Sturm et al. (2011) find large-scale (kiloparsec) flows with $v_{\rm out} \sim 1 \times 10^3$ km s^{-1} and $\dot{M}_{\rm out} \sim 1 \times 10^3$ M$_\odot$ year^{-1}) in the nearby quasar Mrk 231. Other galaxies show similar phenomena (cf. Tacconi et al. 2002; Lonsdale et al. 2003; Veilleux et al. 2009; Riffel & Storchi-Bergmann 2011a,b; Sturm et al. 2011; see also Zubovas & King 2012a, their tables 1 and 2, for a detailed comparison with the theoretical predictions). In each case it appears that AGN feedback is the driving agency. There is general agreement for Mrk 231, for example, that the mass outflow rate $\dot{M}_{\rm out}$ and the kinetic energy rate $\dot{E}_{\rm out} = \dot{M}_{\rm out} v_{\rm out}^2/2$ are too large to be driven by star formation but comparable with values predicted for AGN feedback.

It appears that energy-driven outflows from SMBH that have just reached their M–σ masses should be able to sweep galaxy spheroids clear of gas. A robust observational test of this is the expected mechanical luminosity (cf. Equation 23)

$$L_{\rm mech} \sim \frac{\eta}{2} L_{\rm Edd} \simeq 0.05L, \quad (61)$$

where $L = lL_{\rm Edd}$ is the observed AGN luminosity. This is investigated by Cicone et al. (2014). As their figure 12 shows, observation does largely confirm the relation in Equation 61. If the AGNs are close to their Eddington luminosities (so that $L \propto M \propto \sigma^4$ and $l \simeq 1$), the clear-out rate $\propto \sigma^{10/3}$ (Equation 57) should scale linearly with the driving luminosity L. Cicone et al.'s figure 9 shows evidence for this correlation, with normalization close to that predicted.

Very recently, Tombesi et al. (2015) announced the clear detection of a powerful UFO with a velocity of $0.25c$ in the X-ray spectrum of IRAS F11119+3257, a nearby ($z = 0.189$) ultraluminous infrared galaxy hosting a powerful molecular outflow. This observation offers clear confirmation of the connection suggested here between disc wind driving on small scales and energy driving on large scales. The observations of this source are in remarkable quantitative agreement with the predictions of Equations 50, 57, 60, and 61.

5.4. Effects of a Galaxy Disc: Stimulated Star Formation and Outflow Morphology

We have so far discussed galaxy spheroids in isolation. This is in line with the observational evidence (see Kormendy & Ho 2013 for a review) that the SMBH scaling relations apply only to this component of a galaxy and are essentially unaffected by the presence of a galaxy disc. In particular, we suggest that the critical M–σ black hole mass is set by small-scale momentum-driven outflows interacting with only a very small central part of the bulge. But the energy-driven outflows we considered in the previous subsection are global. They expand to far greater scales and, unless the galaxy is an elliptical, must inevitably encounter its disc as they expand. In a gas-rich

galaxy the gas in the innermost disc at radius R_0 must be close to self-gravitating. Assuming that the potential is roughly isothermal, it is straightforward to show that this implies a gas density $\sim\rho_{\rm d} \sim 2\sigma^2/R_0$, i.e., greater than the bulge gas density by the factor $\sim 1/f_{\rm g} \sim 6$. We see from Equation 46 that higher gas densities mean lower spherical outflow velocities as they meet greater resistance. So when an initially spherical outflow encounters a high-density gas disc, the outflow goes around it, over its upper and lower plane faces. But the pressure in the outflow is at least initially far higher than in the disc. We can read off the pressure at the contact discontinuity from Equation 43 as

$$P_{\rm CD} = \frac{f_{\rm g}\sigma^2(2\sigma^2 + v_{\rm e}^2)}{\pi G R^2} \simeq \frac{f_{\rm g}\sigma^2 v_{\rm e}^2}{\pi G R^2} \tag{62}$$

and estimate the pressure at the forward shock into the ISM as

$$P_{\rm fs} = \frac{4}{3}\rho(R)v_{\rm e}^2 = \frac{2 f_{\rm g}\sigma^2 v_{\rm e}^2}{3\pi G R^2} \simeq \frac{2}{3} P_{\rm CD}. \tag{63}$$

By contrast, the midplane pressure in a disc close to self-gravitating is

$$P_{\rm disc} \sim \rho c_{\rm s}^2 \sim \rho\sigma^2 \sim 2\frac{\sigma^4}{G R_{\rm d}^2}, \tag{64}$$

where we have assumed the sound speed $c_{\rm s} \sim \sigma$ and the self-gravity condition $G\rho \sim \Omega^2$ with $\Omega = \sqrt{2}\sigma/R_{\rm d}$ and the Kepler frequency at disc radius $R_{\rm d}$. Thus when the outflow shock arrives at $R = R_{\rm d}$, its pressure is a factor $\sim (v_{\rm e}/\sigma)^2 \sim 25$ larger than the disc's, and this remains true until the outflow shock has traveled out to radii $R > R_{\rm d} v_{\rm e}/\sigma \sim 5 R_{\rm d}$. Any such compression must trigger a burst of star formation in the disc (cf. Thompson et al. 2005, their appendix B), and here it rises to values

$$\dot{\Sigma}_* \sim 2 \times 10^3 \epsilon_{-3} \sigma_{200}^{10/3} l^{2/3} R_{\rm kpc}^{-2}\ {\rm M}_\odot\ {\rm kpc}^{-2} \tag{65}$$

(Zubovas et al. 2013), where $\epsilon_* = 10^{-3}\epsilon_{-3}$ is the efficiency of massive stars in converting mass into radiation, and we have substituted for $v_{\rm e}$ using Equation 48. Zubovas et al. (2013) show that this leads to a starburst of total luminosity

$$L_* \simeq 5 \times 10^{47} L_{46}^{5/6} l^{-1/6}\ {\rm erg\ s^{-1}}, \tag{66}$$

where L_{46} is the AGN luminosity in units of 10^{46} erg s^{-1}. Such systems would appear as ultraluminous infrared galaxies.

This suggests that in a galaxy having both a bulge and a disc, the clear-out phase leaves the galaxy bulge without gas but may be accompanied by a starburst in the disc. Recent observations of dusty QSOs appear to show this, with the black hole mass already on the M–$M_{\rm b}$ relation (Equation 1) and so fully grown (Bongiorno et al. 2014). However, in an elliptical, clear-out must leave the galaxy "red and dead."

Because a galaxy disc is a major obstacle to an outflow, it follows that it may be able to divert a quasi-spherical outflow into a bipolar shape. This is particularly true in cases in which the SMBH mass grows only a little, such as in a minor accretion event. Zubovas et al. (2011) suggest that the gamma-ray emitting bubbles disposed symmetrically about the plane of the Milky Way (Su et al. 2010) may be the remnants of a relatively recent and rather weak event like this.

5.5. The Three M–σ Relations

So far in this section, we have seen that the arrival of the black hole mass at the M–σ relation means that its feedback makes a radical change from momentum driving to energy driving. The

energy-driven phase that clears the gas from a galaxy bulge is short and violent. But it is clear that the black hole must inject a nonnegligible amount of energy to eject the gas, and this requires accretion energy, i.e., some black hole mass growth. Evidently if the mass increment ΔM needed for this is $\gg M_\sigma$, then we will have failed to explain the M–σ relation.

The mass increment ΔM is influenced by two factors. First, it must require significantly less SMBH growth to remove the gas from a spiral galaxy with a relatively small bulge than, for example, an elliptical, in which the much larger bulge mass means that energy driving by the central SMBH wind must continue for longer in order to expel the remaining gas. Zubovas & King (2012b) find that energy driving, and therefore SMBH mass growth above M_σ, must continue only for approximately 4 Myr (about 0.1 Salpeter times) in a typical spiral, but for approximately 2 Salpeter times in an elliptical. So the final SMBH mass in a spiral is close to M_σ, but in an elliptical it can reach

$$M_{\rm final} \sim e^2 M_\sigma \sim 7.5 M_\sigma . \tag{67}$$

The second factor affecting M, $M_{\rm b}$ is the galaxy environment. Cluster ellipticals can gain gas as they orbit through the intracluster gas. Some brightest cluster galaxies, which are near the center of the cluster potential, are known to contain unusually massive SMBHs (McConnell et al. 2011). Taking account of the extra black hole mass growth required to remove the bulge, and the mass a galaxy may gain from its surroundings, implies three parallel but slightly offset M–σ relations for spirals, field, and cluster ellipticals (see **Figure 10**). In principle there is also a relation for cluster spirals, but these are rare. We see from the figure that the spread in offsets means that an observed sample drawn from galaxies of all three types would tend to produce a slope slightly steeper than the individual ones for each type, perhaps accounting for the slight discrepancy between the observed overall slope $\alpha = 4.4 \pm 0.3$ and the theoretical value of four. All three types of galaxies obey a similar M–$M_{\rm b}$ relation (Equation 1) within the errors, as growth of the SMBH above M_σ goes together with higher $M_{\rm b}$.

6. THE SUPERMASSIVE BLACK HOLE–BULGE MASS RELATION

6.1. Feedback and the M–M_b Relation

In the Introduction, we noted the observed proportionality (Equation 1) between M and $M_{\rm b}$ as well as the M–σ relation. So far, we have concentrated almost entirely on the second of these relations and suggested that it arises because the black hole feedback itself directly limits the mass reservoir available for black hole growth. Quite independently of details, almost every discussion of this relation adopts this view (see Section 7 below).

But the character of the M–$M_{\rm b}$ relation must be very different. Because we are assuming that feedback ensures that the black hole mass M is set by σ, we cannot argue that M is independently set by $M_{\rm b}$. However, reversing the argument to suggest that the black hole mass M sets $M_{\rm b}$ is also implausible, because $M_{\rm b}$ is in the form of stars.

So there can be no directly causal connection between the black hole mass M and the stellar bulge mass $M_{\rm b}$. (Indeed one view—see Section 6.3 below—asserts that the connection is purely statistical.) Instead, their relation must arise because whatever determines $M_{\rm b}$ makes it proportional to σ^4. Empirically, we already know that this is approximately true, at least for elliptical galaxies, which are the largest spheroids of all, because these are observed to obey the Faber & Jackson (1976) relation

$$L_* \sim 2 \times 10^{10}\ {\rm L}_\odot\ \sigma_{200}^4 . \tag{68}$$

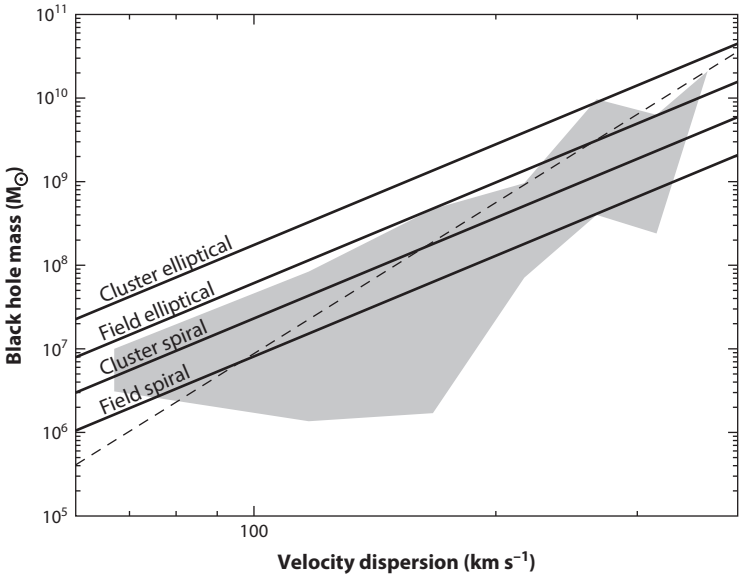

Figure 10

The four (in reality three, as cluster spirals are rare) M–σ relations (*solid lines*) and their combined effect on observational fits (*dashed line*). All solid lines have slopes $M \propto \sigma^4$, and the dashed line has $M \propto \sigma^6$. The gray area is the approximate locus of data points of McConnell et al. (2011, their figure 3). Reprinted from Zubovas & King (2012b) with permission.

Here L_* is the total stellar luminosity and mass of an elliptical, so for mass-to-light ratios of $\sim$5, we immediately get the stellar mass as

$$M_* \sim 1 \times 10^{11}\ \mathrm{M}_\odot \sigma_{200}^4 \sim 10^3 M. \tag{69}$$

There is now general agreement that this relation, like the M–σ relation, may result from feedback inhibiting and ultimately suppress the process that produces it. The difference is that here the feedback is from stars, and what ultimately must be suppressed is star formation. Several papers make this point, starting with Murray et al. (2005). Power et al. (2011) show that this approach gives a bulge stellar mass

$$M_\mathrm{b} \sim \frac{0.14 f_\mathrm{g} t_\mathrm{H} \sigma^4}{\epsilon_* c G}, \tag{70}$$

where $\epsilon_* \simeq 2 \times 10^{-3}$ measures the total luminous energy yield from a main-sequence star in terms of its rest-mass energy $M_* c^2$, and t_H is the Hubble time. Comparing with Equation 39, we get

$$M \simeq M_\sigma \sim \frac{1.8 \kappa \epsilon_* c}{\pi G t_\mathrm{H}} M_\mathrm{b} \sim 10^{-3} M_\mathrm{b}, \tag{71}$$

which is similar to observational estimates (cf. Equation 1). Both the M–σ and the M–M_b relations hold for elliptical galaxies, so Equation 70 automatically reproduces the Faber & Jackson (1976) relation for typical mass-to-light ratios. In this view, the similarity of the SMBH and stellar (Faber-Jackson) M, $M_\mathrm{b} \propto \sigma^4$ relations (Equations 2 and 69) follows directly because both result from momentum feedback and the ratio $M/M_\mathrm{b} \sim 10^{-3}$ reflects the relative efficiencies of the black hole and stellar versions.

6.2. The M–σ Relation for Nucleated Galaxies

A similar argument (McLaughlin et al. 2006) shows that for nucleated galaxies (i.e., those whose central regions are dominated by nuclear star clusters, with no detectable sign of the presence of an SMBH), there should be an offset M–σ relation between the mass of the cluster and the velocity dispersion, i.e.,

$$M_c \simeq 20 M_\sigma \simeq 6 \times 10^9 \sigma_{200}^4 \mathrm{M}_\odot. \quad (72)$$

Typically these galaxies are small, with $\sigma < 120$ km s^{-1}. The factor ~20 offset in cluster mass for a given σ arises because momentum driving by an ensemble of cluster stars is approximately 20 times less efficient per unit mass than from a black hole accreting at the Eddington rate.

6.3. Mergers and the M–M_b Relation

Jahnke & Macciò (2011) offer a radically different interpretation of the M–M_b relation. Building on earlier work by Peng (2007), they assume that black hole and bulge masses are built up by repeated mergers of smaller galaxies with uncorrelated M and M_b. They follow this evolution using dark matter halo merger trees and, as a result of the central limit theorem, find that M is roughly proportional to M_b, with the scatter decreasing for larger masses, in which there have been more mergers. They conclude that the SMBH–bulge scaling relation may have an explanation that is largely or even entirely noncausal.

But it is hard to accept that there is no more physics in the SMBH scaling relations than this. First, the actual ratio M/M_b is left undetermined by this procedure. Second, to get from the M–M_b relation to M–σ requires one to assume something like the $M_b \propto \sigma^4$ relation (Equation 69) implied by Faber-Jackson, so physics presumably must enter here too (cf. Section 6.2 above). Finally, it would seem a remarkable coincidence that the outcome of this indirect process by chance produces an M–σ relation exactly equivalent to requiring that the SMBH Eddington thrust should just balance the weight of the bulge gas.

7. MOMENTUM, ENERGY, OR RADIATION?

The study of AGN outflows and their effects on the host galaxy has two main aims. A viable picture must explain both the scaling relations and, simultaneously, the fact that galaxy spheroids appear ultimately to be largely swept clear of gas by high-speed molecular outflows that have significantly greater scalar momenta $\dot{M} v$ than the radiative value L/c of the central AGN (the clear-out problem). The discussion given above offers plausible physical grounds that the shock interaction characterizing the black hole wind feedback changes from being momentum driven, acting on small spatial scales near the black hole, to energy driven, instead acting globally on the whole galaxy bulge and producing a high-energy clear-out of its gas. The M–σ relation marks the point at which this transition occurs in a given galaxy. We argue below (Section 8) that observations support this picture of local-global transition in several ways, but before accepting this conclusion we should consider other possibilities.

First, the switch from momentum to energy driving depends on the details of gas cooling. It is sometimes argued (e.g., Silk & Nusser 2010) that strong cooling of the ambient ISM enforces momentum driving by the central SMBH throughout. In fact cooling the ambient gas is not relevant to the question of energy or momentum driving; as we have seen, it is the cooling of the shocked wind gas that decides this. But as we emphasized in Section 4, at least some of the physics of the suggested momentum-energy switch is still beyond a full numerical treatment with realistic

parameters. It is sensible then to check our treatment above by considering the momentum-driven and energy-driven cases in isolation and then the effect of direct radiation pressure.

7.1. Wind Momentum Driving

We first simply assume that a black hole wind acts on its surroundings by pure momentum driving alone, at all radii. For generality, we let the preshock wind have speed v_w and take its mechanical luminosity $\dot{M}_w v_w^2/2$ as a fixed fraction a of L_{Edd}; i.e., we do not explicitly assume that the wind has the Eddington momentum, as seems to hold for UFOs. Then the momentum feedback first becomes important at a critical black hole mass M_{crit} roughly given by equating the wind thrust $\dot{M}_w v_w = 2aL_{Edd}/v$ to the weight

$$W = \frac{GM(R)M_g(R)}{R^2} = \frac{4f_g\sigma^4}{G} \tag{73}$$

of the overlying gas in an isothermal potential (cf. Equation 37). With $L_{Edd} = 4GM_{crit}c/\kappa$, we get

$$M_{crit} = \frac{v_w}{2ac} M_\sigma. \tag{74}$$

By definition $a = \dot{M}_w v_w^2/2L_{Edd}$ and $L_{Edd} = \eta \dot{M}_{Edd} c^2$, so

$$M_{crit} = \frac{v_w}{2ac} M_\sigma = \frac{\eta c}{v_w} \frac{\dot{M}_{Edd}}{\dot{M}_w} M_\sigma, \tag{75}$$

(cf. Fabian 1999). We see that for general wind parameters the critical mass differs from M_σ. We find $M_{crit} = M_\sigma$ only if $v_w = \eta c \dot{M}_{Edd}/\dot{M}_w$, which is Equation 22. This immediately implies that the wind momentum is Eddington, i.e., $\dot{M}_w v_w = \eta \dot{M}_{Edd} c = L_{Edd}/c$. In other words, assuming pure momentum driving gives the critical mass as M_σ if and only if the driving wind has the Eddington momentum, i.e., has the properties observed for UFOs.

But pure momentum driving is unable to drive off the bulge gas without a significant increase of the black hole mass above M_{crit}. Equation 41 shows that the outflow velocity is far smaller than observed unless M reaches values $\sim 14\ M_\sigma$, and there is a similar deficit in the predicted mass outflow rates (cf. Equations 52 and 57). Several authors have reached similar conclusions (cf. Silk & Nusser 2010, McQuillin & McLaughlin 2012). Moreover, if galaxy bulges accrete at the rates suggested by cosmological simulations, it seems unlikely that any hypothetical momentum-driven outflows would have enough thrust to prevent infall and so could not turn off star formation definitively (cf. Costa et al. 2014). We conclude that pure momentum driving, even given the lack of a likely shock cooling process, probably does not give a realistic picture of the interaction between SMBHs and their hosts.

7.2. Wind Energy Driving

The direct opposite case from that considered in Section 7.1 is pure energy driving by winds, in which radiative cooling is assumed to be negligible throughout. This was often the implicit assumption in early treatments (e.g. Haehnelt et al. 1998, Silk & Rees 1998). The equation of motion for this case is Equation 46. This shows that gas is always driven out at constant speed for any SMBH mass, however small. Setting $R = v_e t$ and using the definition of L_{Edd} (cf. Equation 17) gives the speed v_e as

$$v_e^3 = \frac{\pi G^2 c \eta M}{3 f_g \kappa \sigma^2}. \tag{76}$$

This expresses the fact that the adiabatically expanding shocked gas pushes the interstellar gas away as a hot atmosphere for any SMBH mass. One can easily find the corresponding mass outflow rate by setting $\dot{M}_{\rm out}v_{\rm e}^2/2 \sim L_w$, because we know that the outflow mechanical luminosity is a significant fraction of the driving wind mechanical luminosity $L_{\rm w} = \eta l L_{\rm Edd}/2$.

To define a critical SMBH mass for energy-driven outflow, one must impose a further condition. This is usually taken as $v_{\rm e} \sim \sigma$, defining some kind of escape velocity. But it is not obvious that this is appropriate. The outflow is driven by pressure, so the ballistic escape velocity is not relevant. Even if the AGN switches off, the residual gas pressure still drives outflow for a long time (cf. **Figure 9**). If we nevertheless impose this condition, we find a critical mass

$$M_{\rm energy} = \frac{3f_{\rm g}\kappa}{\pi G^2 \eta c}\sigma^5 = \frac{3\sigma}{\eta c}M_\sigma = 0.02M_\sigma = 6\times 10^6\,{\rm M}_\odot\,\sigma_{200}^5, \tag{77}$$

which is a factor $3\sigma/\eta c \sim 1/50$ too small in comparison with observations.

Silk & Rees (1998) considered the growth of a protogalaxy (i.e., gas with $f_{\rm g} \sim 1$) around a seed SMBH that formed earlier, but their argument applies to the coevolution of the SMBH and host also, provided we take $f_{\rm g} \sim 0.1$. They assume the wind sweeps mass into a shell with speed $v_{\rm s}$, and implicitly neglect pressure work, and the fact that energy is shared between the shocked wind and the swept-up outflow. This would imply the relation

$$L_{\rm w} = 2\pi r^2 f_{\rm g}\rho(r)v_{\rm s}^3, \tag{78}$$

as each new shell of mass $4\pi r^2\rho(r)v_{\rm s}$ now simply acquires kinetic energy $v_{\rm s}^2/2$ as it is swept up. Using the isothermal relation (Equation 34) and requiring $v_{\rm s} \sim \sigma$ gives

$$M_{\rm SR} \simeq \frac{f_{\rm g}\kappa}{4\pi G^2 f_{\rm w} c}\sigma^5 \simeq 5\times 10^4\left(\frac{f_{\rm g}}{0.16}\right){\rm M}_\odot\,\sigma_{200}^5, \tag{79}$$

where $f_{\rm w} = L_{\rm w}/L_{\rm Edd}$. The neglect of pressure work overestimates the wind-driving efficiency, so this mass is even smaller than that given by Equation 77. It is clear that wind energy driving of this type does not correctly reproduce the M–σ relation, giving a critical mass too low by factors of 50–100.

A more promising approach has recently been outlined by Nayakshin (2014), Zubovas & Nayakshin (2014), and Bourne et al. (2014), who consider the effects of strong inhomogeneity of the bulge gas. They assume first that inverse Compton shock cooling may not be effective because of two-temperature effects (but see Section 4.2 above). Second, they suggest that sufficiently dense clouds of interstellar gas would feel a net outward force $\sim \rho v^2$ per unit area when overtaken by a free-streaming UFO wind of preshock density ρ and speed v, thus mimicking a momentum-driven case. The density of these clouds is a factor $1/f_{\rm g} \sim 6$ below the star-formation threshold. If most of the ISM gas mass is in the form of such clouds, SMBH feeding must stop when the outward force overcomes gravity, which gives a relation like Equation 39 up to some numerical factor. This idea throws up several gas-dynamical problems. First, a cogent treatment must explain the origin and survival of clouds at densities close to but just below the star-formation threshold, which must contain most of the interstellar gas. The clouds must be completely immersed in the wind, so the net outward force on them is a surface drag, which is dimensionally also $\sim \rho v^2$ per unit area. Estimating this surface drag requires knowledge of how the cloud-wind interfaces evolve on very small scales. Because these are formidable theoretical tasks, we should ask for observational tests. The main difference from the quasi-spherical momentum-driven case is that instead of being radiated away, most of the energy of the UFO wind now continuously drives the tenuous intercloud part of the ISM out of the galaxy at high speed. If this tenuous gas is a fraction $f_{\rm t}$ of the total gas content, Equations 50 and 57 show that, for SMBH masses not too far from M_σ, this outflow should have speed $v_{\rm out} \sim 1.23\times 10^3 f_{\rm t}^{-1/3}$ km s^{-1} and mass-loss rate $M_{\rm out} \sim 4\times 10^3 f_{\rm t}\,{\rm M}_\odot$ year^{-1}

and so be potentially observable for many AGN spheroids. Because most of the gas feels only the ram pressure of the black hole wind, this type of flow does not naturally lead to gas ejection once the M_σ mass is reached but requires SMBH growth to higher masses (see the discussion in the last paragraph of Section 7.1). From the work of Section 5.4 one may also expect a continuously elevated star-formation rate in the central parts of their galaxy discs also, which is not in general observed.

7.3. Cosmological Simulations

Cosmological simulations often produce an empirical M–σ relation as part of much larger structure formation calculations. Limits on numerical resolution inevitably require a much more broad-brush approach than adopted here. The effect of the SMBH on its surroundings is usually modeled by distributing energy over the gas of the numerical galaxy at a certain rate. This injected mechanical luminosity is then iterated until the right relation appears. This empirical approach (e.g., di Matteo et al. 2005) seems to require a mechanical luminosity $\sim 0.05 L_{\rm Edd}$ to produce the observed M–σ relation. This is precisely what we expect (cf. Equation 23) for a black hole wind with the Eddington momentum $\dot{M}_{\rm out} v = L_{\rm Edd}/c$.

But the success of this procedure is puzzling. If the ambient gas absorbed the full numerically injected mechanical luminosity $0.05 L_{\rm Edd}$, the resulting outflows would give the energy-driven (Equation 32) or Silk-Rees mass (Equation 79) above, which is too small compared with observations. The fact that cosmological simulations instead actually iterate roughly to the observed M–σ value (Equation 42) suggests that they somehow arrange that the injected energy only couples to the gas at the very inefficient rate that occurs in momentum driving or possibly that the numerical gas distribution is highly inhomogeneous. The real physics determining this in both cases operates at length scales far below the resolution of any conceivable cosmological simulation, so the inefficiency must be implicit in some of the subgrid physics that all such simulations have to assume (cf. Costa et al. 2014, their appendix B).

7.4. Radiation Driving

We remarked in the Introduction that in principle direct radiation pressure is the strongest perturbation that a black hole makes on its surroundings, but its effects are more limited than this suggests. As we already indicated, the reason is that in many situations radiation decouples from matter before it has transferred significant energy or momentum. This is particularly likely for radiation emitted by an AGN in the center of a galaxy bulge. The gas density (cf. Equation 34) is sharply peaked toward the center, and so is its tendency to absorb or scatter the radiation from the accreting black hole.

7.4.1. Electron scattering opacity. The electron scattering optical depth from gas outside a radius R is

$$\tau(R) = \int_R^\infty \kappa \rho(r) {\rm d}r = \frac{\kappa f_{\rm g} \sigma^2}{2\pi G R}, \tag{80}$$

which is mostly concentrated near the inner radius R. This means that gas initially close to the AGN is probably swept into a thin shell by its radiation and thus at radius R has optical depth

$$\tau_{\rm sh}(R) \simeq \frac{\kappa f_{\rm g} \sigma^2}{\pi G R}, \tag{81}$$

which is very similar to the undisturbed gas outside R (cf. Equation 80). Gas distributed in this way has large optical depth near the black hole when its inner edge R is small (i.e., less than the value $R_{\rm tr}$ specified in Equation 82 below). Then the accumulating accretion luminosity L of the AGN

is initially largely trapped and isotropized by electron scattering, producing a blackbody radiation field whose pressure grows as the central AGN radiates. This growing pressure pushes against the weight $W = 4f_g\sigma^4/G$ (Equation 73) of the swept-up gas shell at radius R. This is exactly like the material energy driving we discussed in Section 5.3, except here the photon gas has $\gamma = 4/3$ rather than $\gamma = 5/3$ there. Clearly the effectiveness of this radiation driving is eventually limited because the shell's optical depth falls off like $1/R$ as it expands. The force exerted by the radiation drops as it begins to leak out of the cavity, until for some value $\tau_{\rm tot}(R) \sim 1$ it cannot drive the shell further.

This shows that the sweeping up of gas by radiation pressure must stop at a transparency radius

$$R_{\rm tr} \sim \frac{\kappa f_g \sigma^2}{\pi G} \simeq 50 \left(\frac{f_g}{0.16}\right) \sigma_{200}^2 \ {\rm pc}, \tag{82}$$

where (up to a logarithmic factor) the optical depth $\tau_{\rm tot}$ is of order 1 so that the radiation just escapes, acting as a safety valve for the otherwise growing radiation pressure. This process is discussed in detail by King & Pounds (2014), who suggest that the stalled gas at $R_{\rm tr}$ may be the origin of the WA phenomenon (cf. Tombesi et al. 2013). The radius $R_{\rm tr}$ is so small that very little accretion energy is needed to blow interstellar gas to establish this structure and to adjust it as the galaxy grows and changes σ.

7.4.2. Dust. At larger radii much of the cold diffuse matter in the galaxy bulge may be in the form of dust. The absorption coefficient of dust depends strongly on wavelength and is far higher than electron scattering in the UV, but decreases sharply in the IR (e.g., Draine & Lee 1984). The energy of a UV photon absorbed by a dust grain may be reemitted almost isotropically as many IR photons, which then escape freely. The net effect is that dusty gas feels only the initial momentum of the incident UV photon, whereas most of the incident energy escapes. Then a spherical shell around an AGN would experience a radial force $\simeq L/c$, where L is the UV luminosity, as long as it remained optically thick to this kind of dust absorption. This is dynamically similar to wind-powered flows in the momentum-driven limit, and this type of radiation-powered flow is often also called momentum driven, even though the physical mechanism is very different.

An important distinction between the wind- and radiation-powered cases is that ambient gas in the path of a wind cannot avoid feeling its effects, whereas this is not true for radiation, as the gas may be optically thin. Galaxies are generally optically thin to photons in various wavelength ranges, and a radiation-driven shell may stall at finite radius because its optical depth τ becomes so small that the radiation force decouples, as we saw in the electron-scattering case. Ishibashi & Fabian (2012, 2013, 2014) appeal to this property to suggest that star formation in massive galaxies proceeds from inside to outside as radiation–momentum driven shells of dusty gas are pushed out and then stall at the dust transparency radius $R_{\rm dust} \simeq (\kappa_d/\kappa)R_{\rm tr}$. For large dust opacities $\kappa_d \sim 1 \times 10^3\kappa$, this can give $R_{\rm dust} \sim 50$ kpc. In contrast, galaxies are probably never optically thin to winds, and the density of a black hole wind is always diluted as $1/R^2$, so it inevitably shocks against a swept-up shell of interstellar gas at large R.

The mathematical similarity (cf. Equation 37) between wind-powered and radiation-powered momentum driving allows an empirical estimate of an M–σ relation for the latter if we assume that observed AGNs define the relation and that their observed luminosities correspond to $L/L_{\rm Edd} \sim 0.1$–1. This gives $M_{\rm crit} = (L_{\rm Edd}/L)M_\sigma \sim 1$–$10M_\sigma$ (Murray et al. 2005). Optical depth effects may narrow this range closer to the observed one (DeBuhr et al. 2011). This suggests that radiation driving may be compatible with the M–σ relation, but a momentum-driven outflow like this can never simultaneously reproduce the high-velocity molecular outflows characterizing the clear-out phase. In particular its momentum is $L/c < L_{\rm Edd}/c$, which is considerably smaller than the observed $\sim 20L_{\rm Edd}/c$ of such flows (see Section 5.3). In other words, we have the usual difficulty

that momentum driving can accommodate the M–σ relation, but not simultaneously solve the clear-out problem.

One way of possibly overcoming this (e.g., Faucher-Giguère et al. 2012; Faucher-Giguère & Quataert 2012) is to assume (cf. Roth et al. 2012) that instead of degrading incident high-energy photons to lower-energy ones that escape freely, the effect of dust absorption is to retain much of the incident radiant energy. Then if the dust is distributed spherically and is in a steady state, the radiation force on it is $\tau L/c$, where τ is the radial optical depth of the dust (cf. Roth et al. 2012). This form of radiation driving of optically thick dust can in principle produce outflows whose scalar momenta are boosted above that of the driving luminosity L/c by a factor $\sim\tau$ because photons may be reabsorbed several times. For $\tau \gg 1$ the radiation field is effectively trapped and presumably approaches a blackbody form (cf. the discussion of the electron scattering case above), limiting the boost.

Evidently for radiation driving of dust to solve simultaneously both the SMBH scaling and clear-out problems require a sharp transition in the properties of the dust opacity at the critical M–σ mass. This must change the outflows from effectively momentum driven (incident photons are absorbed but their energy escapes as softer photons) to energy driven (incident photons trapped) at this point, in a switch analogous to the turnoff of Compton cooling in the wind-driven case. There have so far been no suggestions of how this may happen, but the physics of dust opacity is sufficiently complex that this is perhaps not surprising.

8. SMALL- VERSUS LARGE-SCALE FEEDBACK

We have shown that UFO winds are very common in AGNs, despite quite restrictive conditions on their observability. They provide an obvious way for the central SMBH to communicate its presence to its host. This in turn suggests ways of understanding both the SMBH–galaxy scaling relations and the need to expel gas from the galaxy spheroid to terminate star formation. It is clear that this AGN feedback must operate at times on small scales and at others on large scales. Our discussion of feedback points to a natural association between momentum driving and small scales and between energy driving and large scales. Small-scale phenomena naturally explained by wind momentum driving include:

1. Super-solar elemental abundances in observed AGN spectra. Wind momentum driving automatically sweeps up and compresses the same gas many times before the black hole mass reaches M_σ. Generations of massive stars forming out of the same swept-up gas can repeatedly enrich the gas close to the SMBH with nuclear-processed material before both the M_σ mass is reached and momentum driving changes to energy driving.
2. Dark matter cusp removal. The same repeated sweeping-up of a gas mass comparable with the SMBH mass, followed by fallback, has a strong tendency to weaken dark matter cusps. Because the baryonic mass involved is much larger, this is a more powerful version of the mechanism invoked by Pontzen & Governato (2012) (see also Garrison-Kimmel et al. 2013), who considered supernovae near the SMBH.
3. Quiescence of AGN hosts. Most AGN hosts do not show dramatically elevated star formation in the central regions of their galaxy discs or, so far, much evidence for high-speed ($\sim$1,000 km s^{-1}) and massive (several 100 $M_\odot$ year^{-1}) outflows on large scales. This is compatible with wind driving by momentum but not energy.

Large-scale phenomena suggesting the action of energy driving include:

4. Metals in the circumgalactic medium. These must be made in galaxies and only later expelled to make the circumgalactic medium. This suggests that expulsion through energy driving

acts only after stellar evolution has had time to enrich a significant fraction of the galaxy bulge gas.

5. Mechanical luminosities of galaxy-scale molecular outflows. These are observed to be close to 5% of their central AGN luminosities L, just as expected for energy driving, with momenta close to the predicted $20L/c$.
6. Suppression of cosmological infall. Energy-driven outflows at large radii probably prevent galaxies accreting indefinitely (Costa et al. 2014).

This list seems to favor a combination of momentum and energy driving, with some kind of switch between them. The suppression of inverse Compton shock cooling at the point when the black hole mass reaches M_σ appears promising, but requires further work on how observable the cooling is, as does the possibility of two-fluid effects. Radiation driving on dust could produce similar behavior, although the physics controlling the required switch between momentum and energy driving is so far unexplained.

It is worth stressing that even a detailed understanding of the dual role of AGN feedback in establishing both the SMBH–host scaling relations and the quenching of star formation would solve only half of the problem. For a full picture of how black holes and galaxies influence each other, we need to know what physical mechanism can produce a supply of gas with so little angular momentum that much of it can accrete onto the central SMBH within a few Salpeter times (see Section 3.1 and Equation 20). We saw in Section 1.2 that the hole's gravity is far too weak to influence the galaxy on the mass scale needed for this. Only feedback can do this, perhaps suggesting that SMBH feedback may ultimately be the underlying cause of SMBH feeding (cf. Dehnen & King 2013).

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

We thank Walter Dehnen, Dean McLaughlin, Sergei Nayakshin, Chris Nixon, Jim Pringle, Chris Power, James Reeves, Simon Vaughan, Mark Wilkinson, and Kastytis Zubovas for help, collaboration, and advice on the subjects reviewed here. We have benefitted hugely from discussions with many people over the years, including Mitch Begelman, Martin Elvis, Andy Fabian, Claude-André Faucher-Giguère, Reinhard Genzel, Martin Haehnelt, Luis Ho, Knud Jahnke, Roberto Maiolino, David Merritt, Ramesh Narayan, Ken Ohsuga, Brad Peterson, Eliot Quataert, Martin Rees, Joop Schaye, Joe Silk, Francesco Tombesi, and Sylvain Veilleux.

LITERATURE CITED

Batcheldor D. 2010. *Ap. J. Lett.* 711:L108
Blustin AJ, Page MJ, Fuerst SV, Branduardi-Raymont G, Ashton CE. 2005. *Astron. Astrophys.* 431:111
Bongiorno A, Maiolino R, Brusa M, et al. 2014. *MNRAS* 443:2077
Bourne A, Nayakshin S. 2013. *MNRAS* 436:2436
Bourne MA, Nayakshin S, Hobbs A. 2014. *MNRAS* 441:3055
Cappi M. 2006. *Astron. Nachr.* 327:1012
Cappi M, Tombesi F, Bianchi S, et al. 2009. *Astron. Astrophys.* 504:401
Chartas G, Brandt WN, Gallagher SC, Garmire G. 2002. *Ap. J.* 569:179

Cicone C, Maiolino R, Sturm E, et al. 2014. *Astron. Astrophys.* 562:A21
Ciotti L, Ostriker JP. 1997. *Ap. J. Lett.* 487:L105
Collin-Souffrin S, Dumont AM. 1990. *Astron. Astrophys.* 229:292
Costa T, Sijacki, D, Haehnelt MG. 2014. *MNRAS* 444:2355
Dadina M, Cappi M, Malaguti G, Ponti G, de Rosa A. 2005. *Astron. Astrophys.* 442:461
DeBuhr J, Quataert E, Ma C-P. 2011. *MNRAS* 412:1341
Dehnen W, King A. 2013. *Ap. J. Lett.* 777:L28
Dehnen W, McLaughlin DE. 2005. *MNRAS* 363:1057
Denney KD, Watson LC, Peterson BM, et al. 2009. *Ap. J.* 702:1353
Di Matteo T, Springel V, Hernquist L. 2005. *Nature* 433:604
Draine BT, Lee HM. 1984. *Ap. J.* 285:89
Dyson JE, Williams DA. 1997. *The Physics of the Interstellar Medium.* Bristol: Inst. Phys. 2nd ed.
Faber SM, Jackson RE. 1976. *Ap. J.* 204:668
Fabian AC. 1999. *MNRAS Lett.* 308:L39
Fabian AC. 2012. *Annu. Rev. Astron. Astrophys.* 50:455
Faucher-Giguère C-A, Quataert E. 2012. *MNRAS* 425:605
Faucher-Giguère C-A, Quataert E, Murray N. 2012. *MNRAS* 420:1347
Ferrarese L, Merritt D. 2000. *Ap. J. Lett.* 539:L9
Feruglio C, Maiolino R, Piconcelli E, et al. 2010. *Astron. Astrophys.* 518:L155
Garrison-Kimmel S, Rocha M, Boylan-Kolchin M, Bullock JS, Lally J. 2013. *MNRAS* 433:3539
Gebhardt K, Bender R, Bower G, et al. 2000. *Ap. J. Lett.* 539:L13
Gofford J, Reeves JN, Tombesi F. 2013. *MNRAS* 430:60
Haehnelt MG, Natarajan P, Rees MJ. 1998. *MNRAS* 300:817
Halpern JP. 1984. *Ap. J.* 1984:90
Häring N, Rix H-W. 2004. *Ap. J. Lett.* 604:L89
Hernquist L. 1990. *Ap. J.* 356:359
Ishibashi W, Fabian AC. 2012. *MNRAS* 427:2998
Ishibashi W, Fabian AC. 2013. *MNRAS* 431:2350
Ishibashi W, Fabian AC. 2014. *MNRAS* 441:1474
Jahnke K, Macciò AV. 2011. *Ap. J.* 734:92
Kallman T, Liedahl D, Osterheld A, Goldstein W, Kahn S. 1996. *Ap. J.* 465:994
Kaspi S, Behar E. 2006. *Ap. J.* 636:674
Kaspi S, Smith PS, Netzer H, et al. 2000. *Ap. J.* 533:631
King AR. 2003. *Ap. J. Lett.* 596:L27
King AR. 2005. *Ap. J. Lett.* 635:L121
King AR. 2010. *MNRAS* 402:1516
King AR, Davies MB, Ward MJ, Fabbiano G, Elvis M. 2001. *Ap. J. Lett.* 552:L109
King AR, Pounds KA. 2003. *MNRAS* 345:657
King AR, Pounds KA. 2014. *MNRAS Lett.* 437:L81
King AR, Pringle JE. 2006. *MNRAS Lett.* 373:L90
King AR, Pringle JE, Hofmann JA. 2008. *MNRAS* 385:1621
King AR, Zubovas K, Power C. 2011. *MNRAS Lett.* 415:L6
Kormendy J, Ho LC. 2013. *Annu. Rev. Astron. Astrophys.* 51:511
Lonsdale CJ, Smith HE, Diamond PJ. 2003. *Ap. J.* 592:804
Markowitz A, Reeves JN, Braito V. 2006. *Ap. J.* 646:783
Marziani P, Sulentic JW, Dultzin-Hacyan D, Clavani M, Moles M. 1996. *Ap. J. Suppl.* 104:37
McConnell NJ, Ma C-P, Gebhardt K, et al. 2011. *Nature* 480:215
McKernan B, Yaqoob T, Reynolds CS. 2007. *MNRAS* 379:1359
McLaughlin DE, King AR, Nayakshin S. 2006. *Ap. J. Lett.* 650:L37
McQuillin RC, McLaughlin DE. 2012. *MNRAS* 423:2162
Murray N, Quataert E, Thompson TA. 2005. *Ap. J.* 618:569
Navarro JF, Frenk CS, White SDM. 1996. *Ap. J.* 462:563
Navarro JF, Frenk CS, White SDM. 1997. *Ap. J.* 490:493

Nayakshin S. 2014. *MNRAS* 437:2404
O'Brien PT, Reeves JN, Simpson C, Ward MJ. 2005. *MNRAS* 360:25
Ohsuga K, Mineshige S. 2011. *Ap. J.* 736:2
Peng CY. 2007. *Ap. J.* 671:1098
Pontzen A, Governato F. 2012. *MNRAS* 421:3464
Pounds KA, King AR. 2013. *MNRAS* 433:1369
Pounds KA, Page KL. 2006. *MNRAS* 372:1275
Pounds KA, Reeves JN. 2007. *MNRAS* 374:823
Pounds KA, Reeves JN. 2009. *MNRAS* 397:249
Pounds KA, Reeves JN, King AR, et al. 2003. *MNRAS* 345:705
Pounds KA, Reeves JN, Page KL, O'Brien PT. 2004. *Ap. J.* 616:696
Pounds KA, Vaughan S. 2011a. *MNRAS* 413:1251
Pounds KA, Vaughan S. 2011b. *MNRAS* 415:2379
Pounds KA, Vaughan S. 2012. *MNRAS* 423:165
Power C, Zubovas K, Nayakshin S, King AR. 2011. *MNRAS Lett.* 413:L110
Reeves JN, Done C, Pounds KA, et al. 2008. *MNRAS Lett.* 385:L108
Reeves JN, O'Brien PT, Ward MJ. 2003. *Ap. J.* 593:65
Reynolds C, Fabian A. 1995. *MNRAS* 273:1167
Riffel RA, Storchi-Bergmann T. 2011a. *MNRAS* 411:469
Riffel RA, Storchi-Bergmann T. 2011b. *MNRAS* 417:2752
Roth N, Kasen D, Hopkins PF, Quataert E. 2012. *Ap. J.* 759:36
Rupke DSN, Veilleux S. 2011. *Ap. J. Lett.* 729:L27
Shakura NI, Sunyaev RA. 1973. *Astron. Astrophys.* 24:337
Silk J, Nusser A. 2010. *Ap. J.* 725:556
Silk J, Rees MJ. 1998. *Astron. Astrophys.* 331:L1
Soltan A. 1982. *MNRAS* 200:115
Spergel DN, Verde L, Peiris HV, et al. 2003. *Ap. J. Suppl.* 148:175
Sturm E, González-Alfonso E, Veilleux S, et al. 2011. *Ap. J. Lett.* 733:L16
Su M, Slatyer TR, Finkbeiner DP. 2010. *Ap. J.* 724:1044
Tacconi LJ, Genzel R, Lutz D, et al. 2002. *Ap. J.* 580:73
Thompson TA, Quataert E, Murray N. 2005. *Ap. J.* 630:167
Tombesi F, Cappi M, Reeves JN, et al. 2010. *Astron. Astrophys.* 521:A57
Tombesi F, Cappi M, Reeves JN, et al. 2011. *Ap. J.* 742:44
Tombesi F, Cappi M, Reeves JN, et al. 2013. *MNRAS* 430:1102
Tombesi F, Meléndez M, Veilleux S, et al. 2015. *Nature* 519:436–38
Turner MJL, Abbey A, Arnaud M, et al. 2001. *Astron. Astrophys.* 365:L27
Vaughan S, Uttley P. 2008. *MNRAS* 390:421
Veilleux S, Rupke DSN, Kim DC, et al. 2009. *Ap. J. Suppl.* 182:628
Verner DA, Ferland GJ. 1996. *Ap. J. Suppl.* 103:467
Vito F, Maiolino R, Santini P, et al. 2014. *MNRAS* 441:1059
Zubovas K, King A. 2013. *Ap. J.* 769:51
Zubovas K, King AR. 2012a. *Ap. J. Lett.* 745:L34
Zubovas K, King AR. 2012b. *MNRAS* 426:2751
Zubovas K, King AR. 2014. *MNRAS* 439:400
Zubovas K, King AR, Nayakshin S. 2011. *MNRAS Lett.* 415:L21
Zubovas K, Nayakshin S. 2014. *MNRAS* 440:2625
Zubovas K, Nayakshin S, King A, Wilkinson M. 2013. *MNRAS* 433:3079