Co-Evolution of Central Black Holes and Galaxies
Proceedings IAU Symposium No. 267, 2009
B. M. Peterson, R. S. Somerville, & T. Storchi-Bergmann, eds.

doi:10.1017/S1743921310006101

What Do Statistics Reveal About the $M_{\rm BH}$–$M_{\rm bulge}$ Correlation and Co-Evolution?

Chien Y. Peng[1]

[1]Herzberg Institute of Astrophysics, National Research Council of Canada,
5071 West Saanich Road, Victoria, British Columbia, V9E 2E7, Canada
Email: cyp@nrc-cnrc.gc.ca

Abstract. Observational data show that the correlation between the masses of supermassive black holes $M_{\rm BH}$ and galaxy bulge masses $M_{\rm bulge}$ follows a nearly linear trend, and that the correlation is strongest with the bulge rather than the total stellar mass $M_{\rm gal}$. With increasing redshift, the ratio $\Gamma = M_{\rm BH}/M_{\rm bulge}$ relative to $z = 0$ also seems to be larger for $M_{\rm BH} \gtrsim 10^{8.5}\ M_{\odot}$. This study looks more closely at statistics to see what effect it has on creating, and observing, the $M_{\rm BH}$–$M_{\rm bulge}$ correlation. It is possible to show that if galaxy merging statistics can drive the correlation, minor mergers are responsible for causing a convergence to linearity most evident *at high masses*, whereas major mergers have a central limit convergence that more strongly *reduces the scatter*. This statistical reasoning is agnostic about galaxy morphology. Therefore, combining statistical prediction (more major mergers $\Longrightarrow$ tighter correlation) with observations (bulges = tightest correlation), would lead one to conclude that more major mergers (*throughout an entire merger tree*, not just the primary branch) give rise to more prominent bulges. Lastly, with regard to controversial findings that Γ increases with redshift, this study shows why the luminosity function (LF) bias argument, taken correctly at face value, actually strengthens, rather than weakens, the findings. However, correcting for LF bias is unwarranted because the BH mass scale for quasars is bootstrapped to the $M_{\rm BH}$–σ_* correlation in normal galaxies at $z = 0$, and quasar–quasar comparisons are mostly internally consistent. In Monte-Carlo simulations, high Γ galaxies are indeed present: they are statistical outliers (i.e., "under-merged") that take longer to converge to linearity via minor mergers. Additional evidence that the galaxies are undermassive at $z \gtrsim 2$ for their $M_{\rm BH}$ is that the quasar hosts are very compact for their expected mass.

Keywords. galaxies: active, galaxies: nuclei, methods: numerical, methods: statistical

1. Introduction

The discoveries of fundamental correlations between $M_{\rm BH}$ with stellar velocity dispersion (Ferrarese & Merritt 2000; Gebhardt *et al.* 2000) and $M_{\rm bulge}$ (Kormendy & Richstone 1995; Magorrian *et al.* 1998) have greatly influenced our view of the role of black hole activity in galaxy evolution (e.g., Kauffmann & Haehnelt 2000; Granato *et al.* 2004; Di Matteo *et al.* 2005). The $M_{\rm BH}$-$M_{\rm bulge}$ correlation is remarkable in that it is almost linear, has a scatter of only 0.3 dex, and holds true over five orders of magnitude in $M_{\rm BH}$ dynamic range; locally, the ratio of $M_{\rm bulge}/M_{\rm BH} \approx 800$ (Marconi & Hunt 2003; Häring & Rix 2004).

How did the correlations come about and how do selection biases affect our observations of the correlations? Direct cause and effect are not only possible, there are numerous theoretical proposals. While quasar feedback is one of the most widely investigated and favored explanations (Di Matteo *et al.* 2005; Robertson *et al.* 2006; Hopkins *et al.* 2007a), galaxy mergers may perhaps share the role. This study therefore isolates the role

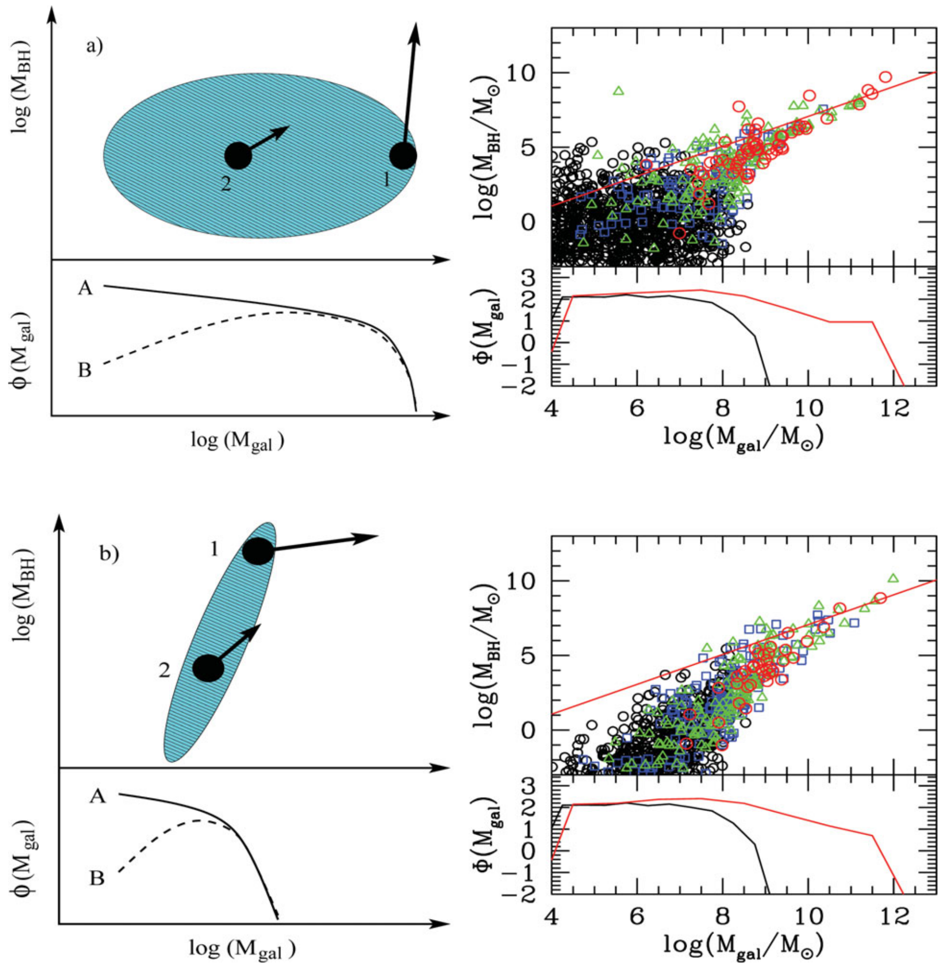


Figure 1. How mergers cause $M_{\rm BH}$ and $M_{\rm gal}$ to correlate. *Top left:* No correlation initially. Objects at location 1 increase $M_{\rm gal}$ more quickly than $M_{\rm BH}$ upon merging, compared to objects at location 2. The correlation steepens over time due to minor mergers. *Bottom left:* Steep correlation initially. The opposite situation occurs, i.e., $M_{\rm gal}$ grows more quickly than $M_{\rm BH}$ for objects at 1 compared to 2. Symmetry between these two scenarios produces a linear relation, asymptotically. *Right:* Monte-Carlo simulations with the initial conditions shown on the left.

of merger statistics to examine how it might affect the growth of the $M_{\rm BH}$–$M_{\rm bulge}$ correlation. It also examines more closely how luminosity selection biases affect the inference of the $M_{\rm BH}$–$M_{\rm bulge}$ correlation and its evolution since $z \approx 4$, as deduced from quasars.

2. Merging Statistics: How the $M_{\rm BH}$–$M_{\rm bulge}$ Correlation Can Result and Why It Is Important to Also Consider the $M_{\rm BH}$–$M_{\rm total}$ Relation

How galaxy merging affects the BH vs. bulge correlations has been considered in several studies (e.g., Islam *et al.* 2003; Ciotti & van Albada 2001) from a purely statistical standpoint, and using specific initial conditions and assumptions (e.g. no scatter or starting with a prior correlation). Going further, Peng (2007) shows that the two most salient features of the $M_{\rm BH}$–$M_{\rm bulge}$ correlation — linearity and strong correlation with

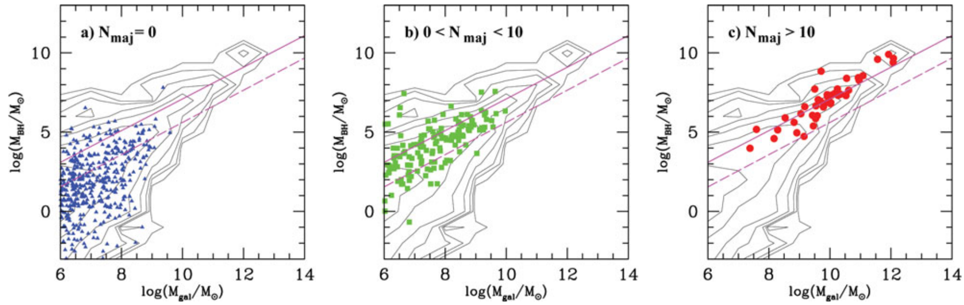


Figure 2. Effects of major mergers based on Monte-Carlo simulations. The solid line corresponds to a linear correlation; it is not a fit. The contours show the distribution of points at the end of the Monte Carlo simulation. *Left:* No major mergers $N_{\rm maj} = 0$. The dashed line shows roughly the mean of the points with slope fixed to a linear correlation. *Middle:* Cumulative major mergers $0 < N_{\rm maj} < 10$ over all progenitors in a merger tree of a galaxy. *Right:* Cumulative major mergers $N_{\rm maj} \geqslant 10$ over all progenitor galaxies.

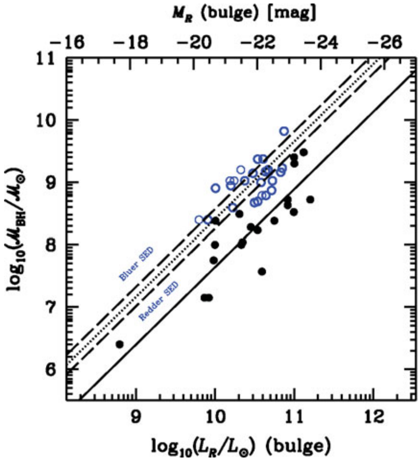
bulges — can be attained without having to make assumptions about the initial conditions. The heuristic toy model proposed by Peng (2007), shown in Figure 1, explains that a linear quality of the $M_{\rm BH}$–$M_{\rm bulge}$ relation comes about because of minor mergers. Minor mergers affect the $M_{\rm BH}$–$M_{\rm gal}$ relation in a way that over some number of events would attain a cosmic average ratio in $M_{\rm BH}/M_{\rm gal}$. That there is necessarily a correlation can be reasoned from a symmetry argument; that the correlation trends toward *linearity* can be understood by noticing that the only way minor mergers can no longer change the $M_{\rm BH}/M_{\rm gal}$ ratio is when it has the same value everywhere along the mass sequence.

Major mergers, however, play a different role: they do not significantly change the ratio of $M_{\rm BH}/M_{\rm gal}$ because the galaxies involved in merging have both similar $\langle M_{\rm BH}\rangle$ and $\langle M_{\rm gal}\rangle$ by the definition of major mergers. As explained in more detail by Peng (2007), in the limit where major mergers are occurring between identical mass galaxies, the BH masses sum according to the central limit theorem given as

$$\sigma\left(\log\mu_{\rm merge}\right) = \frac{\sigma\langle M_{\rm BH,1} + M_{\rm BH,2}\rangle}{\langle M_{\rm BH,1} + M_{\rm BH,2}\rangle}. \tag{2.1}$$

Thus, because the sum $\langle M_{\rm BH,1} + M_{\rm BH,2}\rangle$ increases more quickly than the dispersion, $\sigma\langle M_{\rm BH,1} + M_{\rm BH,2}\rangle$, the scatter in the $M_{\rm BH}$ distribution after merging, i.e., $\sigma\left(\log\mu_{\rm merge}\right)$, decreases with each major merger. The effect of central limit convergence due to major mergers can be quite dramatic. Figure 2 shows results from one possible Monte Carlo simulation, in which there is no correlation initially, and has two orders of magnitude of scatter. Minor mergers (Figure 2a) alone do not reduce the scatter significantly by the end of the simulation. However, major mergers cause a rapid decrease in scatter in only a few events. Note that the relevant accounting of major mergers is the cumulative sum *over the entire merger tree*, i.e., over all progenitors, their progenitors, etc., rather than the more common approach of tracking the main branch.

It is worth noting that statistical reasoning does not predict morphology from first principles. Therefore statistical reasoning can explain the observations of a tight $M_{\rm BH}$–$M_{\rm bulge}$ correlation if and only if massive bulges were preferentially formed through more major mergers than disky galaxies, summed over all progenitor histories. Even though the notion that major mergers lead to formation of bulges is now widely regarded to be true, it is interesting that it can be reasoned purely from statistical principles and the known existence of a tight $M_{\rm BH}$–$M_{\rm bulge}$ correlation. Furthermore, the $M_{\rm BH}$–$M_{\rm bulge}$ relation may be a special case of the $M_{\rm BH}$–$M_{\rm total}$ relation, despite the latter having a larger scatter. Thus, to understand the co-evolution of galaxies with $M_{\rm BH}$, one ought to consider both the $M_{\rm BH}$–$M_{\rm bulge}$ and $M_{\rm BH}$–$M_{\rm total}$ relations.



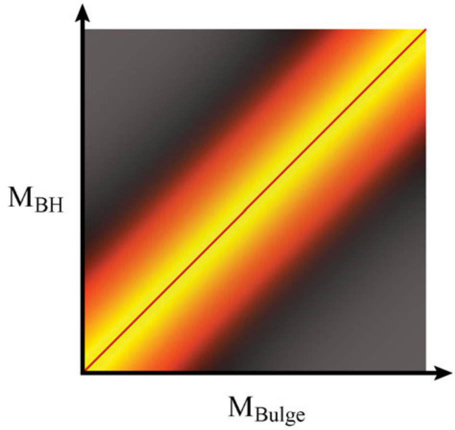


Figure 3. Observed $z \gtrsim 1.7$ quasar data (open circles) from Peng *et al.* (2006a). Solid circles are $z = 0$ normal galaxies. The absolute luminosities are how the galaxies would appear after accounting for luminosity fading.

Figure 4. The intrinsic correlation. What it means for the intrinsic correlation between $M_{\rm BH}$ and $M_{\rm bulge}$ to be linear. This distribution is also called "the prior" and the conditional $P(M_{\rm BH}|M_{\rm bulge})$. To simplify discussion, we assume $P(M_{\rm BH}|M_{\rm bulge}){=}P(M_{\rm bulge}|M_{\rm BH})$ as shown; doing so does not affect the conclusions qualitatively.

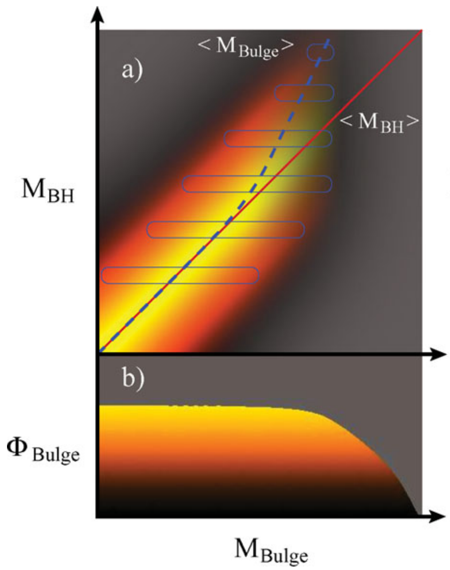
3. Luminosity-Function Bias of Galaxies and Quasars and Other Biases

In recent years, there have been a number of efforts to study the $M_{\rm BH}$–$M_{\rm bulge}$ correlation beyond the local universe using quasars, radio galaxies, and other means (e.g., McLure *et al.* 2006; Peng *et al.* 2006a,b; Woo *et al.* 2006; Treu *et al.* 2007). For the most part, the studies find that the central BHs were larger at $z \gtrsim 2$ in the past for a given bulge stellar mass by a factor Γ of $3 \lesssim \Gamma \lesssim 6$ (Peng *et al.* 2006a,b), shown in Figure 3. These findings have been called into question by other studies on the basis that the luminosity function (LF) selection was not explicitly accounted for (e.g., Lauer *et al.* 2007). However, the criticisms have not been very germane, both because the BH mass scale in quasars is not on an absolute scale, and the LF bias goes in the opposite direction in quasars than claimed, as discussed below. The issues are subtle and have led to substantial confusion.

3.1. *Revisiting the Luminosity Function Selection Bias to See Why It Affects the* $M_{\rm BH}$*–*$M_{\rm bulge}$ *Correlation in Galaxies Differently From Quasars.*

Figures 4–6 illustrate schematically the idea of the LF selection bias. Figure 4 shows the prior that there is an intrinsic, perfectly linear, correlation between $M_{\rm BH}$ and $M_{\rm bulge}$. This intrinsic correlation $P(M_{\rm BH}|M_{\rm bulge})$ is also known as the conditional. Even though a linear relation does not require $P(M_{\rm BH}|M_{\rm bulge})$ to be the same as $P(M_{\rm bulge}|M_{\rm BH})$, in the discussion below, doing so does not affect the directional sense of the conclusions.

The *luminosity function bias* was pointed out at least as early as Adelberger & Steidel (2005), and more recently by Fine *et al.* (2006), Salviander *et al.* (2006), and Lauer



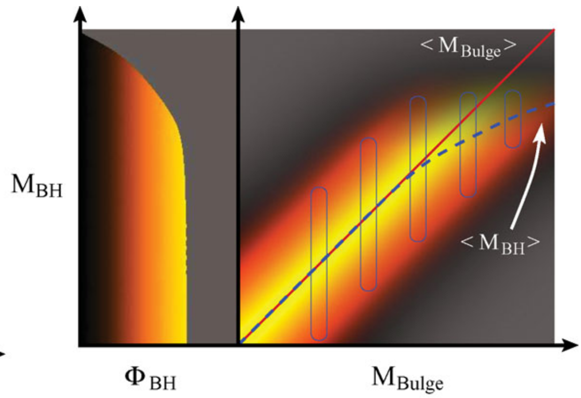


Figure 5. Measuring BHs in normal galaxies. (a) Given the intrinsic prior of Figure 4, what is actually observed when objects are first drawn from the bulge mass (i.e., luminosity) function $\Phi_{\rm Bulge}$ (Panel b), followed by perfect $M_{\rm BH}$ measurement. Note that $\langle M_{\rm bulge}\rangle$ (with $M_{\rm BH}$ as the independent variable, as represented by horizontal rectangles as visual aid) trends upward (dashed line) even though $\langle M_{\rm BH}\rangle$ (solid line, with $M_{\rm bulge}$ as the independent variable) is not biased.

Figure 6. Measuring galaxies around luminous quasars. Selecting quasars from the quasar mass (i.e., luminosity) function $\Phi_{\rm BH}$, given that the intrinsic correlation $P(M_{\rm bulge}|M_{\rm BH})$ between $M_{\rm BH}$ and $M_{\rm bulge}$ is linear (Figure 4). The tapering of the correlation at high $M_{\rm BH}$ is due to there being fewer luminous quasars in the universe, as illustrated by the mass function $\Phi(M_{\rm BH})$ to the left. Note that $\langle M_{\rm BH}\rangle$ (at a given $M_{\rm bulge}$) trends to the right (dashed line), but $\langle M_{\rm bulge}\rangle$ (solid line) is not biased.

et al. (2007). In essence, the act of selecting a sample of galaxies to observe leaves an imprint of the LF on the correlation of $M_{\rm BH}$ vs. $M_{\rm bulge}$. To obtain the $M_{\rm BH}$–$M_{\rm bulge}$ correlation in normal galaxies, the observing sequence is to first select bulges or galaxies from $\Phi(M_{\rm bulge})$, the galaxy bulge mass function shown schematically in Figure 5b, followed by measuring the BH through stellar dynamics or other means. The latter probability — measuring a BH of mass $M_{\rm BH}$ *after selecting on* $M_{\rm bulge}$ — is the conditional $P(M_{\rm BH}|M_{\rm bulge})$ of Figure 4. The observational sequence: $\Phi(M_{\rm bulge}) \times P(M_{\rm BH}|M_{\rm bulge})$ therefore establishes the observed correlation $P(M_{\rm BH}, M_{\rm bulge})$ shown in Figure 5a. The effect of selecting on $\Phi(M_{\rm bulge})$ tapers off the underlying correlation at the right side indiscriminantly of $M_{\rm BH}$. This LF imprint is present even if every galaxy and BH can be detected and measured precisely. It is *not* a Malmquist bias, and the effect prevents us from directly observing the intrinsic correlation. Figure 5a is the same as Figure 2 in Lauer *et al.* (2007). The tapering by $\Phi(M_{\rm bulge})$ causes the $\langle M_{\rm bulge}\rangle$ for a given $M_{\rm BH}$ to deviate from the intrinsic trend, as illustrated by the dashed line in Figure 5a. It leads to the notion that "at a given BH mass, there are more low-mass host galaxies than high-mass." This is the main reason behind the argument that high-z data in Figure 3 are biased.

However, what is subtle and widely misconstrued is the fact that the distribution of Figure 5a applies to normal galaxies but not to quasars. In quasars, the reverse observational sequence occurs, i.e., *measuring host galaxies around BHs*, as opposed to *measuring BHs in normal galaxies.* When selecting on quasars, there is an agreement that one does not draw them from the galaxy luminosity function, but instead must select from $\Phi(L_{\rm QSO})$, the quasar luminosity function. Moreover, because $M_{\rm BH}$ in quasars scales like ${L_{\rm QSO}}^{0.5}$ (Kaspi *et al.* 2000) and quasars appear to radiate at a fixed fraction of the Eddington ratio (Kollmeier *et al.* 2006), selecting on $L_{\rm QSO}$ is essentially drawing on $\Phi(M_{\rm BH})$, as shown in Figure 6. After selecting on quasars, the host galaxy masses $M_{\rm bulge}$ are then drawn from the conditional probability of finding $M_{\rm bulge}$ around a BH of mass $M_{\rm BH}$, i.e., $P(M_{\rm bulge}|M_{\rm BH})$. Therefore, the observational sequence for quasars is given by the product $P(M_{\rm BH}, M_{\rm bulge}) = \Phi(M_{\rm BH}) \times P(M_{\rm bulge}|M_{\rm BH})$. Comparing this to normal galaxies, the labels of $M_{\rm BH}$ and $M_{\rm bulge}$ are simply switched. The act of doing so leads to Figure 6, whereby the $\Phi(M_{\rm BH})$ selection attenuates the intrinsic correlation of Figure 4 on the upper ($M_{\rm BH}$) side. In other words, in quasars it is $\langle M_{\rm BH}\rangle$, not $\langle M_{\rm bulge}\rangle$, that is lower than intrinsic. Therefore, the fact that high-z data in Figure 3 lie on the opposite side of the expected trend is a testament to a positive evolution in Γ if the BH mass scale is absolute. However, it is not, as discussed below, which means this effect is only secondary.

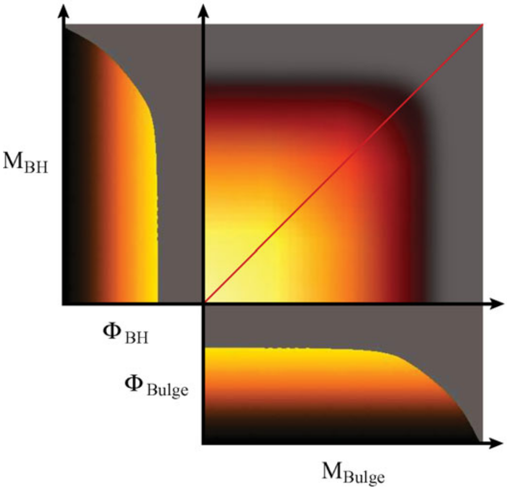
It appears that one reason there is widespread misconception on this subject is a tendency to want to apply the intuitively obvious notion that there are more low-mass than high-mass galaxies at a place where doing so is inappropriate. In other words, after selecting a quasar from the BH mass function $\Phi(M_{\rm BH})$, the tendency is to believe the host galaxies are drawn from galaxy mass function as the conditional, i.e., $P(M_{\rm bulge}|M_{\rm BH}) = \Phi(M_{\rm bulge})$, instead of the intrinsic correlation of Figure 4. On that common notion, the joint product $P(M_{\rm BH}, M_{\rm bulge}) = \Phi(M_{\rm bulge}) \times \Phi(M_{\rm BH})$ heuristically produces a distribution given by Figure 7. Clearly, observations do not support this because the joint product produces no correlation between $M_{\rm BH}$ and $M_{\rm bulge}$. Note that Lauer *et al.* (2007) did not make this particular error; their conditional probability comes from the linear correlation of Figure 4, not $\Phi(M_{\rm BH})$. Statistically, the only way for the conditional $P(M_{\rm bulge}|M_{\rm BH}) = \Phi(M_{\rm bulge})$ is if $M_{\rm BH}$ and $M_{\rm bulge}$ are intrinsically unrelated.

3.2. *The Effects of Malmquist and Quasar-to-Host Galaxy Contrast Biases on the $M_{\rm BH}$–$M_{\rm bulge}$ Correlation in Quasars.*

Malmquist bias is another common factor used to argue against findings that the ratio of $M_{\rm BH}/M_{\rm bulge}$ is higher at high-z than now. However, Figure 8 illustrates schematically that Malmquist bias only attenuates the underside of the distribution. It does not affect the trend at the massive end. It is also qualitatively very different from high-z observations of Figure 3 because, as seen in Figure 8, the attenuation is uniform at a constant $M_{\rm BH}$; it does not cause the points to lie systematically to the left of the correlation line.

Lastly, measuring host galaxies around quasars is affected by the fact that only luminous host galaxies can be detected from beneath luminous quasars. Under normal circumstances without gravitational lensing, host galaxies of quasars are extremely difficult to detect when the quasar:host ratio is larger than 10:1 at a seeing of $0\farcs1$. This selection bias tapers the correlation along a diagonal line illustrated schematically in Figure 8; the angle of the diagonal depends on the magnification ratio if the quasar sample is from gravitational lenses, as in Figure 3. Nevertheless, the observational pressure is to shift the $\langle M_{\rm BH}\rangle$ and $\langle M_{\rm bulge}\rangle$ averages to the right of the intrinsic correlation.

Comparing Figure 8 with Figure 3 therefore qualitatively illustrates that finding a larger $M_{\rm BH}/M_{\rm bulge}$ ratio in high-z quasars is not due to known luminosity selection effects. Qualitatively, observational pressures greatly favor detections to the right of the correlation line where the quasar luminosity contrast is low. The missing objects to the *right* of the intrinsic correlation may be caused by quasar surveys that fail to classify



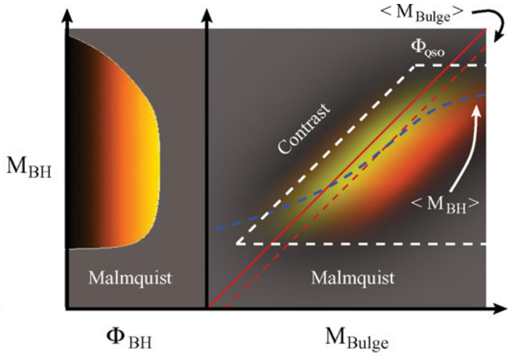


Figure 7. The joint distribution $P(M_{\rm BH}, M_{\rm bulge})=\Phi(M_{\rm bulge})\times\Phi(M_{\rm BH})$. A common, but incorrect, notion is that, for every quasar observed from the BH mass function ($\Phi_{\rm BH}$), the host galaxies can be drawn from the bulge mass function $\Phi_{\rm Bulge} = P(M_{\rm bulge}|M_{\rm BH})$. Rather than producing the $M_{\rm BH}$–$M_{\rm bulge}$ correlation, the joint product results in no correlation.

Figure 8. Other selection biases affecting quasar observations. Observing the $M_{\rm BH}$–$M_{\rm bulge}$ correlation in quasars has the additional biases shown, given that the intrinsic correlation $P(M_{\rm bulge}|M_{\rm BH})$ between $M_{\rm BH}$ and $M_{\rm bulge}$ is linear (Figure 4). The tapering of the correlation at high $M_{\rm BH}$ is again due to selection on $\Phi(M_{\rm BH})$. Malmquist bias selects against faint quasars, independent of $M_{\rm bulge}$ (lower dashed line). On the other hand, faint host galaxies sitting beneath luminous quasars are hard to detect, giving rise to a diagonal selection bias. The exact angle of the diagonal bias depends on the degree of host galaxy magnification. Both $\langle M_{\rm BH}\rangle$ and $\langle M_{\rm bulge}\rangle$ are now shifted to the right of nominal center (solid line) due to contrast bias. Comparing this expected distribution with Figure 3 shows a lack of LF bias in high-z observations.

low contrast, thus redder, objects as being quasar candidates. However, given that even redder and lower contrast systems make it into the Peng *et al.* (2006a) quasar sample at $z = 1$, this effect is judged at face value to probably not be the main culprit.

Note that, hypothetically, it is possible for studies using *other selection functions* besides the ones mentioned to distill a sample of low-luminosity quasars that are then found to the right of the correlation. That would not necessarily contradict current conclusions using quasars. Instead, that hypothetical sample can have properties that distinguish its objects physically from the host galaxies of luminous quasars. Selection functions that draw on different physical attributes may find objects in a different parameter space of the same underlying $M_{\rm BH}$–$M_{\rm total}$ correlation. This might explain the different conclusions seen between quasars and sub-mm galaxies hosting active nuclei (e.g., Alexander *et al.* 2008). To talk about evolution, it is therefore necessary to compare objects selected based on the *same physical and observational* selection functions. In that respect, the quasar–quasar comparisons of high-z and low-z are currently the most *internally consistent* sample to address the issue of the $M_{\rm BH}$–$M_{\rm bulge}$ evolution. Lastly it is important to note that where selection biases strongly partition observable parameter spaces, it is important to not only consider the mean of some trend, but also the distribution *as a whole.* Removing biases in distributions from known selection functions is both possible and feasible.

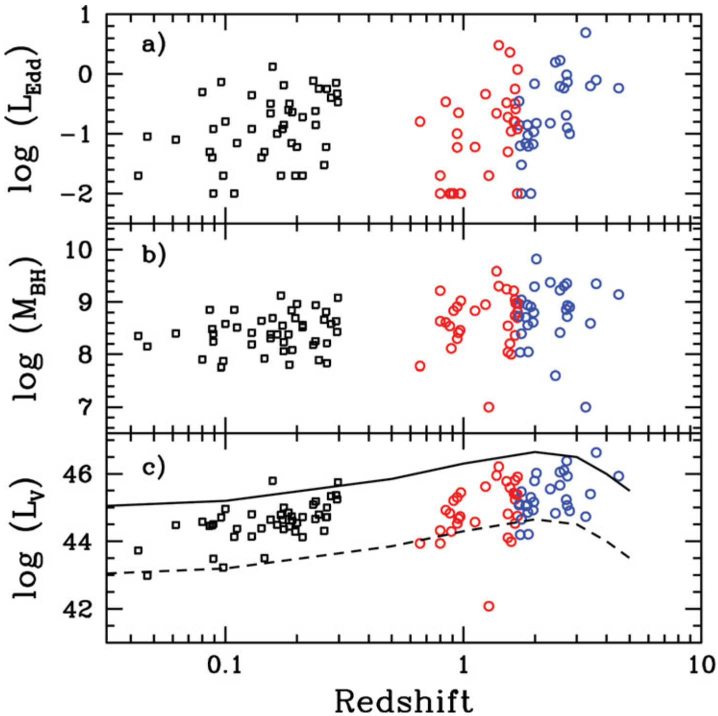


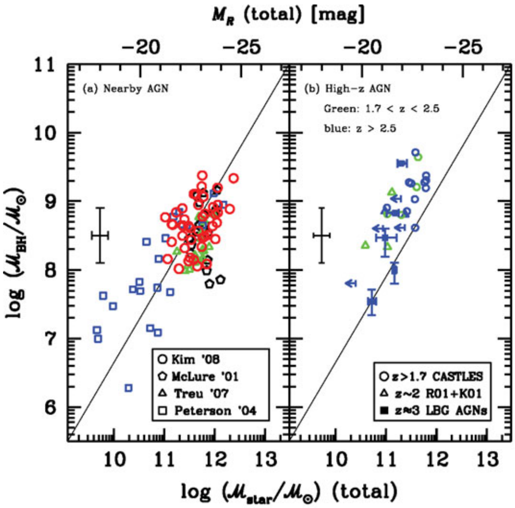
Figure 9. Comparison of quasar properties at high and low z. *Top:* Quasar radiating efficiency in units of $L_{\rm QSO}/L_{\rm Edd}$. *Middle:* $M_{\rm BH}$ in the quasar sample (in $M_\odot$). *Bottom:* Quasar luminosity in ergs s^{-1}. The solid line and dashed line are reference lines showing the trend in the evolution of the quasar break luminosity $L^*_{\rm bol}$ (with arbitrary normalization) from Hopkins *et al.* (2007b). Note that it is the ratio of $\Delta L = \log(L_V/L^*_{\rm bol})$ that affects the degree of bias, so high-z quasars are not relatively biased despite $L_{\rm QSO}$ being somewhat higher. Note that high-z quasars are not atypically large in all these observables compared to low-z quasars from Kim *et al.* (2008).

3.3. *The Black Hole Mass Scale in Quasars Is Tied to Normal Galaxies Through the $M_{\rm BH}$–σ_* Correlation*

In the context of the evolution in $M_{\rm BH}/M_{\rm bulge}$ ratio Γ, the discussions above on the luminosity function bias are mostly academic because the BH mass scale in quasars is tied to normal galaxies through the $M_{\rm BH}$–σ_* correlation (Onken *et al.* 2004). The bias due to the LF selection is normalized out to first order.

To second order, there are other concerns when comparing the high-z sample with low-z, such as the relative luminosities of the quasars, the Eddington ratio, and the possibility that the high-z BHs are unusually massive. These concerns are addressed by Figure 9, which shows that the high-z sample is not too different from the low-z sample in those respects. The one caveat is that, even though the systematic bias in LF selection above is normalized out to first order, there remain residual biases relative to some reference point of the $M_{\rm BH}$–$M_{\rm bulge}$ correlation. Objects more, or less, luminous compared to that reference point may lie systematically away from the correlation, keeping in mind this is at most a second-order effect. Taking that pivot point to be around the break of the $\Phi(M_{\rm BH})$, one can see in Figure 9c that the high-z quasar luminosities in the Peng *et al.* (2006a) sample track the evolution of the LF break (taken from Hopkins *et al.* 2007b, with arbitrary normalization) of the quasars fairly closely both at low and high redshifts.

The fact that the $M_{\rm BH}$ scale in quasars is normalized to normal galaxies means that claims of $M_{\rm BH}$–$M_{\rm bulge}$ evolution is only meaningful if low-z quasars do not show the same



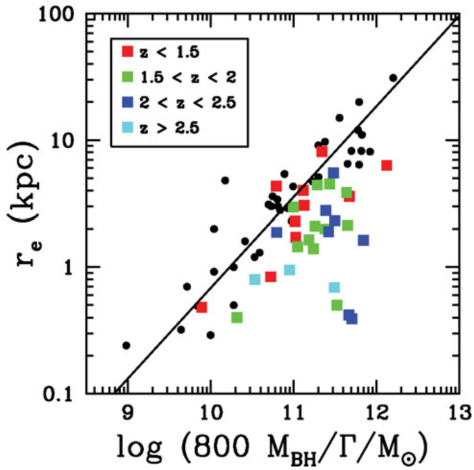


Figure 10. Comparing the $M_{\rm BH}$–$M_{\rm bulge}$ relation for low-z and high-z quasars hosts. *Left:* Low-z quasars. There is no offset relative to normal galaxies (solid line) because the BH mass scale in quasars is normalized to agree. *Right:* However, high-z quasar hosts exhibit an offset.

Figure 11. The radius r_e vs. "bulge mass" relation of high-z quasar hosts compared with low-z elliptical galaxies. The solid circles are normal elliptical galaxies at $z = 0$ with dynamical $M_{\rm BH}$ measurements. The square data points come from gravitationally lensed quasar host measurements of Peng *et al.* (2006a). The product $800 M_{\rm BH}/\Gamma$ is the expected bulge mass at the observed epoch as inferred from Figure 3 or the right hand side of Figure 10. Quasar hosts at $z \gtrsim 2$ are very compact.

offset. Figure 10a shows that the low-z quasars scatter around the normal galaxy correlation (solid line), which indicates that the bootstrapping does not leave large residual biases. In contrast, the high-z sample in Figure 10b clearly lies off the correlation, despite the $M_{\rm BH}$, luminosity, and Eddington ratios being quite similar to the low-z sample.

4. Quasar Host Galaxies at $z \gtrsim 2$ Are Under-Sized for Their Mass

Additional interesting evidence that high-z quasar hosts have a larger $\Gamma = M_{\rm BH}/M_{\rm bulge}$ relative to $z = 0$ (which can also be thought of as a mass deficit in the bulge) comes from comparing the size–$M_{\rm bulge}$ relation at the observed epoch with galaxies today, as shown in Figure 11. In that Figure, the host galaxy mass is inferred from the luminosity of the host galaxy. However, it is useful to recast the mass in terms of Γ, so as to emphasize how the controversial mass deficit parameter affects the size-$M_{\rm bulge}$ correlation in high-z quasars. Doing so, the host galaxy bulge mass is $M_{\rm bulge}(z) = 800 \times M_{\rm BH}/\Gamma$. This equation comes from the fact that normal galaxy bulges at $z = 0$ have $\Gamma = 1$ and $M_{\rm bulge}(z = 0) = 800 \times M_{\rm BH}$. The correlation of r_e with $M_{\rm bulge}$ is revealing because unknown luminosity selection biases are effectively normalized away by accounting for Γ. Figure 11 shows that the host galaxies at $z \sim 1$ seem to lie on the size–$M_{\rm bulge}$ correlation, whereas higher-redshift host galaxies appear to be much more compact per unit mass. By $z \sim 2$, the host galaxies appear to be too small by a factor of 2–3 compared to normal galaxies of the same mass today (Peng 2004). One way to weaken the conclusions is for Γ to be even larger than the controversial claim, which permits these objects to lie on the modern day size–mass correlation. The fact that massive galaxies at high z appear to be extremely compact has been observed by a number of studies, including Trujillo *et al.* (2006), van Dokkum *et al.* (2008), and Stockton *et al.* (2008) among others, and may point to the same evolutionary pathways between quasar host galaxies and distant red galaxies.

5. Conclusions

The statistics of galaxy merging may shed some light on to the controversial findings of evolution in the $M_{\rm BH}/M_{\rm gal}$ ratio. In the Monte-Carlo simulations of Peng (2007), high-mass objects often tend to lie to the left of the asymptotic linear correlation. This happens because such objects were large outliers in the initial distribution and thereby take a longer time to evolve onto the asymptotic relation. Another potential explanation for the larger $M_{\rm BH}/M_{\rm gal}$ ratio is that the quasar phase may signify recent BH growth, so by observing luminous quasars we catch them in a special state on the $M_{\rm BH}$–$M_{\rm bulge}$ correlation. This is consistent with the findings of Hopkins *et al.* (2007a) who explain large offsets as being due to gas rich mergers that both feed the central BH and possess a larger mass fraction in gas. As explained by merger statistics, the temporary up-tick in the BH mass can subsequently merge back onto the asymptotic linear correlation through minor mergers. Indeed, this is seen in Monte-Carlo simulations where the BHs were artificially boosted in mass followed by regular mergers. Combining statistical simulations with observations that high-z quasar host galaxies are very compact, and the fact that major mergers do not change the $M_{\rm BH}/M_{\rm bulge}$ ratio, seems to consistently point to minor mergers being important for transforming quasar hosts morphologically from their compact state at $z \sim 2$ into massive, extended, elliptical galaxies today.

Acknowledgements

I thank Jenny Greene, Chris Kochanek, and Luis Ho for providing very thoughtful comments and additional references.

References

Adelberger, K. L. & Steidel, C. C. 2005, *ApJ*, 627, L1
Alexander, D. M., *et al.* 2008, *AJ*, 135, 1968
Ciotti, L. & van Albada, T. S. 2001, *ApJ*, 552, L13
Di Matteo, T., Springel, V., & Hernquist, L. 2005, *Nature*, 433, 604
Ferrarese, L. & Merritt, D. 2000, *ApJ*, 539, L9
Fine, S., *et al.* 2006, *MNRAS*, 373, 613
Gebhardt, K., *et al.* 2000, *ApJ*, 539, L13
Granato, G. L., De Zotti, G., Silva, L., Bressan, A., & Danese, L. 2004, *ApJ*, 600, 580
Häring, N. & Rix, H.-W. 2004, *ApJ*, 604, L89
Hopkins, P. F., Hernquist, L., Cox, T. J., Robertson, B., & Krause, E. 2007a, *ApJ*, 669, 67
Hopkins, P. F., Richards, G. T., & Hernquist, L. 2007b, *ApJ*, 654, 731
Islam, R. R., Taylor, J. E., & Silk, J. 2003, *MNRAS*, 340, 647
Kaspi, S., Smith, P. S., Netzer, H., Maoz, D., Jannuzi, B. T., & Giveon, U. 2000, *ApJ*, 533, 631
Kauffmann, G., & Haehnelt, M. 2000, *MNRAS*, 311, 576
Kim, M., *et al.* 2008, *ApJ*, 687, 767
Kollmeier, J. A., *et al.* 2006, *ApJ*, 648, 128
Kormendy, J. & Richstone, D. 1995, *ARAA*, 33, 581
Lauer, T. R., Tremaine, S., Richstone, D., & Faber, S. M. 2007, *ApJ*, 670, 249
Magorrian, J., *et al.* 1998, *AJ*, 115, 2285
Marconi, A. & Hunt, L. K. 2003, *ApJ*, 589, L21
McLure, R. J., *et al.* 2006, *MNRAS*, 368, 1395

Onken, C. A., *et al.* 2004, *ApJ*, 615, 645
Peng, C. Y. 2004, PhD thesis, The University of Arizona
—. 2007, *ApJ*, 671, 1098
Peng, C. Y., Impey, C. D., Ho, L. C., Barton, E. J., & Rix, H.-W. 2006a, *ApJ*, 640, 114
Peng, C. Y., *et al.* 2006b, *ApJ*, 649, 616
Robertson, B., *et al.* 2006, *ApJ*, 641, 90
Salviander, S., *et al.* 2006, *New Astron. Revs.*, 50, 803
Stockton, A., McGrath, E., Canalizo, G., Iye, M., & Maihara, T. 2008, *ApJ*, 672, 146
Treu, T., Woo, J.-H., Malkan, M. A., & Blandford, R. D. 2007, *ApJ*, 667, 117
Trujillo, I., *et al.* 2006, *ApJ*, 650, 18
van Dokkum, P. G., *et al.* 2008, *ApJ*, 677, L5
Woo, J.-H., Treu, T., Malkan, M. A., & Blandford, R. D. 2006, *ApJ*, 645, 900