The Astrophysical Journal Supplement Series, 163:1–49, 2006 March

A UNIFIED, MERGER-DRIVEN MODEL OF THE ORIGIN OF STARBURSTS, QUASARS, THE COSMIC X-RAY BACKGROUND, SUPERMASSIVE BLACK HOLES, AND GALAXY SPHEROIDS

Philip F. Hopkins,[1] Lars Hernquist,[1] Thomas J. Cox,[1] Tiziana Di Matteo,[2] Brant Robertson,[1] and Volker Springel[3]
Received 2005 June 14; accepted 2005 October 26

ABSTRACT

We present an evolutionary model for starbursts, quasars, and spheroidal galaxies in which mergers between gas-rich galaxies drive nuclear inflows of gas, producing starbursts and feeding the buried growth of supermassive black holes (BHs) until feedback expels gas and renders a briefly visible optical quasar. The quasar lifetime and obscuring column density depend on both the instantaneous and peak quasar luminosity, and we determine this dependence using a large set of galaxy merger simulations varying galaxy properties, orbital geometry, and gas physics. We use these fits to deconvolve observed quasar luminosity functions and obtain the evolution of the formation rate of quasars with peak luminosity, $\dot{n}(L_{\rm peak}, z)$. Quasars spend extended periods at luminosities well below peak, so $\dot{n}(L_{\rm peak})$ has a maximum corresponding to the "break" in the observed luminosity function. From $\dot{n}(L_{\rm peak})$ and our simulations, we obtain self-consistent hard and soft X-ray and optical luminosity functions and predict many observables at multiple redshifts, including column density distributions of optical and X-ray samples, the luminosity function of broad-line quasars in X-ray samples and broad-line fraction versus luminosity, active BH mass functions, the distribution of Eddington ratios, the mass function of relic BHs and total BH mass density, and the cosmic X-ray background. In every case, our predictions agree well with observed estimates, without invoking ad hoc assumptions about source properties or distributions. We provide a library of Monte Carlo realizations of our models for comparison with observations.

Subject headings: cosmology: theory — galaxies: active — galaxies: evolution — galaxies: nuclei — quasars: general

Online material: color figures

1. INTRODUCTION

The measurement of anisotropies in the cosmic microwave background (e.g., Spergel et al. 2003), combined with observations of high-redshift supernovae (e.g., Riess et al. 1998, 2000; Perlmutter et al. 1999), has established a "standard model" of the universe, in which the energy density is dominated by an unknown form driving accelerated cosmic expansion and most of the mass is nonbaryonic, in a ratio of roughly 5:1 to ordinary matter. On small scales, it is believed that structure formed through gravitational instability. In the currently favored cold dark matter (CDM) paradigm, objects grow hierarchically, with smaller ones forming first and then merging into successively larger bodies. As baryons fall into dark matter potential wells, the gas is shocked and then cools radiatively to form stars and galaxies, in a "bottom-up" progression (White & Rees 1978).

Even with the many successes of this picture, the processes underlying galaxy formation and evolution are poorly understood. For example, there has yet to be an ab initio calculation, starting from an initial state prescribed by the standard model, resulting in a population of objects that reproduces observed galaxies. However, from the same initial conditions, computer simulations have yielded a new, successful interpretation of the Lyα forest in which absorption in caused by density fluctuations in the intergalactic medium (e.g., Cen et al. 1994; Zhang et al. 1995; Hernquist et al. 1996), over many orders of magnitude in column density (e.g., Katz et al. 1996b), explicitly related to growth of structure in a CDM universe (e.g., Croft et al. 1998, 1999, 2002; McDonald et al. 2000, 2005; Hui et al. 2001; Viel et al. 2003, 2004). This suggests that the difficulties with understanding galaxy formation and evolution lie not in the initial conditions or with the description of dark matter, but rather with the physics that has been used to model the baryons.

Observations have revealed regularities in the structure of galaxies that point to some of this "missing" physics. Supermassive black holes appear to reside at the centers of most galaxies (e.g., Kormendy & Richstone 1995; Richstone et al. 1998; Kormendy & Gebhardt 2001), and the masses of these black holes are correlated with either the mass (Magorrian et al. 1998; McLure & Dunlop 2002; Marconi & Hunt 2003) or velocity dispersion (i.e., the $M_{\rm BH}$-σ relation; Ferrarese & Merritt 2000; Gebhardt et al. 2000; Tremaine et al. 2002) of spheroids, demonstrating a direct link between the origin of galaxies and supermassive black holes. Simulations that follow the self-regulated growth of black holes in galaxy mergers (Di Matteo et al. 2005; Springel et al. 2005a) have shown that the energy released through this process has a global impact on the structure of the merger remnant. If this conclusion applies to spheroid formation in general, the simulations demonstrate that models of the origin and evolution of galaxies must account for black hole growth and feedback in a fully *self-consistent* manner.

Analytical and semianalytical modeling (Silk & Rees 1998; Fabian 1999; Wyithe & Loeb 2002, 2003; Begelman & Nath 2005) suggests that beyond a certain threshold, feedback energy from black holes can expel gas from the centers of galaxies, shutting down accretion onto them and limiting their masses. However, these calculations usually ignore the impact of this process on star formation and therefore do not explain the link between

[1] Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138.
[2] Carnegie Mellon University, Department of Physics, 5000 Forbes Avenue, Pittsburgh, PA 15213.
[3] Max-Planck-Institut für Astrophysik, Karl-Schwarzchild-Strasse 1, 85740 Garching bei München, Germany.

black hole growth and spheroid formation, and furthermore make simplifying assumptions about the dynamics of such accretion. For example, the duration of black hole growth is a free parameter that is fixed either using observational estimates or assuming it to be similar to, for example, the dynamical time of the host galaxy or the *e*-folding time for Eddington-limited black hole growth $t_S = M_{BH}/\dot{M} = 4.5 \times 10^7 l^{-1}(\epsilon_r/0.1)$ yr for accretion with radiative efficiency $\epsilon_r = L/\dot{M}c^2 \sim 0.1$ and $l = L/L_{Edd} \lesssim 1$ (Salpeter 1964). Moreover, these studies have adopted idealized models of quasar light curves, usually corresponding to growth at a constant Eddington ratio or on-off, "light bulb," scenarios. As we discuss below, less restrictive modeling suggests that this phase is actually more complex.

Efforts to model quasar accretion and feedback more self-consistently (e.g., Ciotti & Ostriker 1997, 2001; Granato et al. 2004) by treating the hydrodynamic response of gas to black hole growth have generally been restricted to idealized geometries, such as spherical symmetry, employing simple models for star formation and galaxy-scale quasar fueling. However, these works have made it possible to estimate duty cycles of quasars and shown that the objects left behind have characteristics similar to those observed, with quasar feedback being a critical element in reproducing these features (e.g., Sazonov et al. 2005; Kawata & Gibson 2005; Cirasuolo et al. 2005; for a review, see Ostriker & Ciotti 2005).

Springel et al. (2005b) have incorporated black hole growth and feedback into simulations of galaxy mergers and included a multiphase model of star formation and pressurization of the interstellar gas by supernovae (Springel & Hernquist 2003) to examine implications of these processes for galaxy formation and evolution. Di Matteo et al. (2005) and Springel et al. (2005a, 2005b) have shown that gas inflows excited by gravitational torques during a merger both trigger starbursts and fuel rapid black hole growth. The growth of the black hole is determined by the gas supply and terminates as gas is expelled by feedback, halting accretion, leaving a dead quasar in an ordinary galaxy. The self-regulated nature of black hole growth in mergers explains observed correlations between black hole mass and properties of normal galaxies (Di Matteo et al. 2005), as well as the color distribution of ellipticals (Springel et al. 2005a). These results lend support to the view that mergers have played an important role in structuring galaxies, as advocated especially by Toomre & Toomre (1972) and Toomre (1977). (For reviews, see, e.g., Barnes & Hernquist 1992; Barnes 1998; Schweizer 1998.)

Subsequent analysis by Hopkins et al. (2005a, 2005b, 2005c, 2005d) has shown that the merger simulations can account for quasar phenomena as a phase of black hole growth. Unlike what has been assumed in, for example, semianalytical studies of quasars, the simulations predict complicated evolution of quasar lifetimes, fueling rates for black hole accretion, obscuration, and quasar light curves. The light curves were studied by Hopkins et al. (2005a, 2005d), who showed that the self-termination process gives observable lifetimes of $\sim 10^7$ yr for bright optical quasars and predicts a large population of obscured sources as a natural stage of quasar evolution, as implied by observations (for a review, see Brandt & Hasinger 2005). Hopkins et al. (2005a) analyzed simulations over a range of galaxy masses and found that the quasar light curves and lifetimes are always qualitatively similar, with both the intrinsic and observed quasar lifetimes being decreasing functions of luminosity, with longer lifetimes at all luminosities for higher mass (higher peak luminosity) systems. The dependence of the lifetime on luminosity led Hopkins et al. (2005b) to suggest a new interpretation of the quasar luminosity function, in which the steep bright end consists of quasars radiating near the Eddington limit and is directly related to the distribution of intrinsic peak luminosities (or final black hole masses), as has been assumed previously (e.g., Small & Blandford 1992; Haiman & Loeb 1998; Haiman & Menou 2000; Kauffmann & Haehnelt 2000; Somerville et al. 2001; Tully et al. 2002; Wyithe & Loeb 2003; Volonteri et al. 2003; Haiman et al. 2004; Croton et al. 2006), but where the shallow, faint end of the luminosity function describes black holes growing toward or declining from peak phases of quasar activity, with Eddington ratios generally between $l \sim 0.01$ and 1. The "break" in the luminosity function corresponds directly to the *peak* in the distribution of intrinsic quasar properties. As argued by Hopkins et al. (2005b, 2005c), this new interpretation of the luminosity function can self-consistently explain various properties of both the quasar and galaxy populations, connecting the origin of galaxy spheroids, supermassive black holes, and quasars.

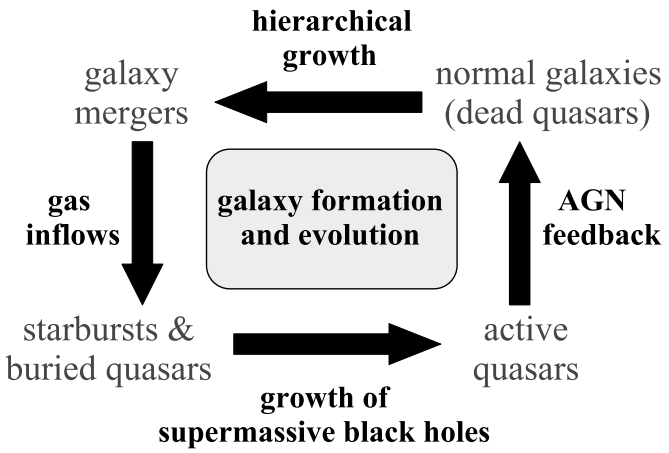

FIG. 1.—Schematic representation of a "cosmic cycle" for galaxy formation and evolution regulated by black hole growth in mergers.

Motivated by these results, and earlier work by many others that we summarize below, in this paper we consider a picture for galaxy formation and evolution, illustrated schematically as a "cosmic cycle" in Figure 1, in which starbursts, quasars, and the simultaneous formation of spheroids and supermassive black holes represent connected phases in the lives of galaxies. Mergers are expected to occur regularly in a hierarchical universe, particularly at high redshifts. Those between gas-rich galaxies drive nuclear inflows of gas, triggering starbursts and fueling the growth of supermassive black holes. During most of this phase, quasar activity is obscured, but once a black hole dominates the energetics of the central region, feedback expels gas and dust, making the black hole visible briefly as a bright quasar. Eventually, as the gas is further heated and expelled, quasar activity can no longer be maintained and the merger remnant relaxes to a normal galaxy with a spheroid and a supermassive black hole. In some cases, depending on the gas content of the progenitors, the remnant may also have a disk (Springel & Hernquist 2005; Robertson et al. 2005a). The remnant will then evolve passively and would be available as a seed to repeat the above cycle. As the universe evolves and more gas is consumed, the mergers involving gas-rich galaxies will shift toward lower masses, explaining the decline in the population of the brightest quasars from $z \sim 2$ to the present, and the remnants that are gas-poor will redden quickly owing to the termination of star formation by black hole feedback (Springel et al. 2005a), so that they resemble elliptical galaxies surrounded by hot X-ray emitting halos (e.g., Cox et al. 2005).

There is considerable observational support for this scenario, which has led the development of this picture for the coevolution

of galaxies and quasars over recent decades. Infrared (IR) luminous galaxies are thought to be powered in part by starbursts (e.g., Soifer et al. 1984a, 1984b; Sanders et al. 1986, 1988a, 1988b; for a review, see, e.g., Soifer et al. 1987), and the most intense examples locally, ultraluminous infrared galaxies (ULIRGs), are invariably associated with mergers (e.g., Allen et al. 1985; Joseph & Wright 1985; Armus et al. 1987; Kleinmann et al. 1988; Melnick & Mirabel 1990; for reviews, see Sanders & Mirabel 1996; Jogee 2006). Radio observations show that ULIRGs have large, central concentrations of dense gas (e.g., Scoville et al. 1986; Sargent et al. 1987, 1989), providing a fuel supply to feed black hole growth. Indeed, some ULIRGs have "warm" IR spectral energy distributions (SEDs), suggesting that they harbor buried quasars (e.g., Sanders et al. 1988c), an interpretation strengthened by X-ray observations demonstrating the presence of two nonthermal point sources in NGC 6240 (Komossa et al. 2003), which are thought to be supermassive black holes that are heavily obscured at visual wavelengths (e.g., Gerssen et al. 2004; Max et al. 2005, Alexander et al. 2005a, 2005b). These lines of evidence, together with the overlap between bolometric luminosities of ULIRGs and quasars, indicate that quasars are the descendents of an infrared luminous phase of galaxy evolution caused by mergers (Sanders et al. 1988a), an interpretation supported by observations of quasar hosts (e.g., Stockton 1978; Heckman et al. 1984; Stockton & MacKenty 1987; Stockton & Ridgway 1991; Hutchings & Neff 1992; Bahcall et al. 1994, 1995, 1997; Canalizo & Stockton 2001).

However, many of the physical processes that connect the phases of evolution in Figure 1 are not well understood. Early simulations showed that mergers produce objects resembling galaxy spheroids (e.g., Barnes 1988, 1992; Hernquist 1992, 1993a) and that if the progenitors are gas-rich, gravitational torques funnel gas to the center of the remnant (e.g., Barnes & Hernquist 1991, 1996), producing a starburst (e.g., Mihos & Hernquist 1996), but these works did not explore the relationship of these events to black hole growth and quasar activity. While a combination of arguments based on time variability and energetics suggests that quasars are produced by the accretion of gas onto supermassive black holes in the centers of galaxies (e.g., Salpeter 1964; Zel'dovich & Novikov 1964; Lynden-Bell 1969), the mechanism that provides the trigger to fuel quasars therefore remains uncertain. Furthermore, there have been no comprehensive models that describe the transition between ULIRGs and quasars that can simultaneously account for observed correlations such as the $M_{\rm BH}$-σ relation.

Here we study these relationships using numerical simulations of galaxy mergers that account for the consequences of black hole growth. In our simulations, black holes accrete and grow throughout a merger event, producing complex, time-varying quasar activity. Quasars reach a peak luminosity $L_{\rm peak}$ during the "blowout" phase of evolution, where feedback energy from black hole growth begins to drive away the gas, eventually slowing accretion. Prior to and following this brief period of peak activity, quasars radiate at instantaneous luminosities L, with $L < L_{\rm peak}$. However, we show that even with this complex behavior, the global characteristics that determine the observed properties of quasars, i.e., lifetimes, light curves, and obscuration, can be expressed as functions of L and $L_{\rm peak}$, allowing us to make predictions for quasar populations that agree well with observations, supporting the scenario sketched in Figure 1.

In § 2 we discuss our methodology and show how the quasar lifetimes and obscuration from our simulations can be expressed as functions of the instantaneous and peak luminosities of quasars. We also define a set of commonly adopted models for the quasar lifetime and obscuration against which we compare our predictions throughout. In § 3 we apply our models to the quasar luminosity function, using the observed luminosity function to determine the distribution of quasar peak luminosities, and show that this allows us to simultaneously reproduce the hard X-ray, soft X-ray, and optical quasar luminosity functions at all redshifts $z \lesssim 3$, and the distribution of column densities in both optical and X-ray samples. In § 4 we determine the time in our simulations when quasars will be observable as broad-line objects and use this to predict the broad-line luminosity function and fraction of broad-line objects in quasar samples as a function of luminosity, as well as the mass function of low-redshift, active broad-line quasars. In § 5 we estimate the distribution of Eddington ratios in our simulations as a function of luminosity and infer Eddington ratios in observed samples at different redshifts. In § 6 we use our modeling to predict both the mass distribution and total density of present-day relic supermassive black holes, and describe their evolution with redshift. In § 7 we similarly apply this model to predict the integrated cosmic X-ray background spectrum, accounting for the observed spectrum from $\sim$1 to 100 keV. In § 8 we discuss the primary qualitative implications of our results and propose falsifiable tests of our picture. Finally, in § 9 we conclude and suggest directions for future work.

Throughout, we adopt a $\Omega_M = 0.3$, $\Omega_\Lambda = 0.7$, and $H_0 = 70$ km s^{-1} Mpc^{-1} ($h = 0.7$) cosmology.

2. THE MODEL: METHODOLOGY

Our model of quasar evolution has several elements, which we summarize here and describe in greater detail below.

1. In what follows, a "quasar" is taken to mean the course of black hole activity in a *single* merger event. We use the term "quasar lifetime" to refer to the time spent by such a quasar at a given luminosity or fraction of the quasar peak luminosity, integrated over all black hole activity in a single merger event. This is not meant to suggest that this would constitute the entire accretion history of a black hole; a given black hole may have multiple "lifetimes" triggered by different mergers, with each merger in principle fueling a distinct "quasar" with its own lifetime. There is no a priori luminosity threshold for quasar activity; the time history can include various epochs at low luminosities and accretion rates.

2. We model the galaxy mergers using hydrodynamic simulations, varying the orbital parameters of the encounter, the internal properties of the merging galaxies, prescriptions for the gas physics, initial "seed" black hole masses of the merging systems, and numerical resolution of the simulations. The black hole accretion rate is determined from the surrounding gas (smoothed over the scale of our spatial resolution, reaching 20 pc in the best cases), i.e., the density and sound speed of the gas, and its motion relative to the black hole, using Eddington-limited, Bondi-Hoyle-Lyttleton accretion theory. The black hole radiates with a canonical efficiency $\epsilon_r = 0.1$ corresponding to a standard Shakura & Sunyaev (1973) thin disk, and we assume that $\sim$5% of this radiated luminosity is deposited as thermal energy into the surrounding gas, weighted by the SPH smoothing kernel (which has a $\sim r^{-2}$ profile) over the scale of the spatial resolution. This scale is such that we cannot resolve the complex accretion flow immediately around the black hole, but we adopt this prescription because (1) it reproduces the observed slope and normalization in the $M_{\rm BH}$-σ relation (Di Matteo et al. 2005), (2) it follows from observations, based on estimates of the energy contained in highly absorbed UV portion of the quasar SED (e.g., Elvis et al. 1994; Telfer et al. 2002), (3) it follows from theoretical considerations of momentum coupling to dust grains in the dense gas very near the quasar (Murray et al. 2005) and hydrodynamic

simulations of small-scale radiative heating from quasar accretion (Ciotti & Ostriker 2001), and (4) even if the feedback is initially highly collimated, a driven wind or shock in a dense region such as the center of the merging galaxies will rapidly isotropize, as long as it is decelerated by gravity and the surrounding medium, allowing the high sound speed within the shock to equalize angle-dependent pressure variations (e.g., Koo & McKee 1990), and furthermore initial local distortions will be washed away in favor of triaxial structure determined by the large-scale density gradients (Bisnovatyi-Kogan & Silich 1991), as occurs in our simulations.

3. For each of our merger simulations, we compute the bolometric black hole luminosity and column density along $\sim$1000 lines of sight to the black hole(s) (evenly spaced in solid angle), as a function of time from the beginning of the simulation until the system has relaxed for $\sim$1 Gyr after the merger.

4. We bin different merger simulations by $L_{\rm peak}$, the peak bolometric luminosity of the black hole in the simulation, and the conditional distributions of luminosity, $P(L|L_{\rm peak})$, and column density, $P(N_{\rm H}|L, L_{\rm peak})$, are computed using all simulations that fall into a given bin in $L_{\rm peak}$. The final black hole mass (black hole mass at the end of the individual merger; subsequent mergers and quasar episodes could further increase the black hole mass) is approximately $M_{\rm BH}^{f} \approx M_{\rm Edd}(L_{\rm peak})$ (but not exactly; see § 2.4), so we obtain similar results if we bin instead by $M_{\rm BH}^{f}$. Our calculation of $M_{\rm BH}^{f}(L_{\rm peak})$ allows us to express our conditional distributions of luminosity and column density in terms of either peak luminosity or final black hole mass. Critically, we find that expressed in terms of $L_{\rm peak}$ or $M_{\rm BH}^{f}$, there is no systematic dependence in the quasar evolution on the varied merger simulation properties; this allows us to calculate a large number of observables in terms of $L_{\rm peak}$ or $M_{\rm BH}^{f}$ without the large systematic uncertainties inherent in attempting to directly estimate, for example, quasar light curves in terms of host galaxy mass, gas fraction, multiphase pressurization of the interstellar medium, orbital parameters and merger stage, and other variables.

5. The observed quasar luminosity function is the convolution of the time a given quasar spends at some observed luminosity with the rate at which such quasars are created. Knowing the distributions $P(L|L_{\rm peak})$ and $P(N_{\rm H}|L, L_{\rm peak})$, we can calculate the time spent by a quasar with some $L_{\rm peak}$ at an observed luminosity in a given wave band. We use this to fit to observational estimates of the bolometric quasar luminosity function $\phi(L)$, deconvolving these quantities to determine the function $\dot{n}(L_{\rm peak})$, i.e., the rate at which quasars of a given peak luminosity must be created or activated (triggered in mergers) in order to reproduce the observed bolometric luminosity function.

6. Given these inputs, we determine the joint distribution in instantaneous luminosity and black hole mass, column density distribution, peak luminosity, and final black hole mass as a function of redshift, i.e., $n(L,\ L_{\nu},\ M_{\rm BH},\ N_{\rm H},\ L_{\rm peak},\ M_{\rm BH}^{f}|z)$, at all redshifts where the observed quasar luminosity function can provide the necessary constraint. From this joint distribution, we can compute, for example, luminosity functions in other wave bands, conditional column density distributions, active black hole mass functions and Eddington ratio distributions, and relic black hole mass functions and cosmic backgrounds. We can compare each of these results to those determined using simpler models for either the quasar lifetime or column density distributions; in § 2.5 we describe a canonical set of such models, to which we compare throughout this paper.

2.1. *The Simulations*

The simulations were performed with GADGET-2 (Springel 2005), a new version of the parallel TreeSPH code GADGET (Springel et al. 2001). GADGET-2 is based on a fully conservative formulation (Springel & Hernquist 2002) of smoothed particle hydrodynamics (SPH), which maintains simultaneous energy and entropy conservation when smoothing lengths evolve adaptively (for a discussion, see, e.g., Hernquist 1993b; O'Shea et al. 2005). Our simulations account for radiative cooling and for heating by a UV background (as in Katz et al. 1996a; Davé et al. 1999), and incorporate a subresolution model of a multiphase interstellar medium (ISM) to describe star formation and supernova feedback (Springel & Hernquist 2003). Feedback from supernovae is captured in this subresolution model through an effective equation of state for star-forming gas, enabling us to stably evolve disks with arbitrary gas fractions (see, e.g., Springel et al. 2005b; Robertson et al. 2004). In order to investigate the consequences of supernova feedback over a range of conditions, we employ the scheme of Springel et al. (2005b), introducing a parameter $q_{\rm EOS}$ to interpolate between an isothermal equation of state ($q_{\rm EOS} = 0$) and the full multiphase equation of state ($q_{\rm EOS} = 1$) described above.

Supermassive black holes (BHs) are represented by "sink" particles that accrete gas at a rate $\dot{M}$ estimated using an Eddington-limited version of Bondi-Hoyle-Lyttleton accretion theory (Bondi 1952; Bondi & Hoyle 1944; Hoyle & Lyttleton 1939). The bolometric luminosity of the black hole is $L_{\rm bol} = \epsilon_r \dot{M} c^2$, where $\epsilon_r = 0.1$ is the radiative efficiency. We assume that a small fraction (typically $\approx$5%) of $L_{\rm bol}$ couples dynamically to the surrounding gas and that this feedback is injected into the gas as thermal energy, as described above.

We have performed several hundred simulations of colliding galaxies, varying the numerical resolution, the orbit of the encounter, the masses and structural properties of the merging galaxies, initial gas fractions, halo concentrations, and the parameters describing star formation and feedback from supernovae and black hole growth. This large set of simulations allows us to investigate merger evolution for a wide range of galaxy properties and to identify any systematic dependence of our modeling. The galaxy models are described in Springel et al. (2005b), and we briefly review their properties here.

The progenitor galaxies in our simulations have virial velocities $V_{\rm vir} = 80$, 113, 160, 226, 320, and 500 km s^{-1}. We consider cases with gas equation of state parameters $q_{\rm EOS} = 0.25$ (moderately pressurized, with a mass-weighted temperature of star-forming gas $\sim 10^{4.5}$ K) and $q_{\rm EOS} = 1.0$ (the full, "stiff" Springel-Hernquist equation of state, with a mass-weighted temperature of star-forming gas $\sim 10^{5}$ K), and initial disk gas fractions (by mass) of $f_{\rm gas} = 0.2$, 0.4, 0.8, and 1.0. Finally, we scale these models with redshift, altering the physical sizes of the galaxy components and the dark matter halo concentration in accord with cosmological evolution (Mo et al. 1998). Details are provided in Robertson et al. (2005b), and here we consider galaxy models scaled appropriately to resemble galaxies of the same $V_{\rm vir}$, $f_{\rm gas}$, and $q_{\rm EOS}$ at redshifts $z_{\rm gal} = 0$, 2, 3, and 6.

For each simulation, we generate two stable, isolated disk galaxies, each with an extended dark matter halo with a Hernquist (1990) profile, motivated by cosmological simulations (e.g., Navarro et al. 1996; Busha et al. 2005) and observations of halo properties (e.g., Rines et al. 2000, 2002, 2003, 2004), an exponential disk of gas and stars, and (optionally) a bulge. The galaxies have masses $M_{\rm vir} = V_{\rm vir}^3/(10GH_0)$ for $z_{\rm gal} = 0$, with the baryonic disk having a mass fraction $m_d = 0.041$; the bulge (when present) has a mass fraction $m_b = 0.0136$, and the rest of the mass is in dark matter typically with a concentration parameter $c = 9.0$. The disk scale length is computed based on an assumed spin parameter $\lambda = 0.033$, chosen to be near the mode in the observed λ-distribution (Vitvitska et al. 2002), and the scale

length of the bulge is set to 0.2 times the resulting value. In Hopkins et al. (2005d), we describe our analysis of simulation A3, one of our set with $V_{\rm vir} = 160$ km s^{-1}, $f_{\rm gas} = 1.0$, $q_{\rm EOS} = 1.0$, and $z_{\rm gal} = 0$, a fiducial choice with a rotation curve and mass similar to the Milky Way, and Hopkins et al. (2005a, 2005b, 2005c) used a set of simulations with the same parameters but varying $V_{\rm vir} = 80$, 113, 160, 226, and 320 km s^{-1}, which we refer to below as runs A1, A2, A3, A4, and A5, respectively.

Typically, each galaxy is initially composed of 168,000 dark matter halo particles, 8000 bulge particles (when present), 24,000 gas and 24,000 stellar disk particles, and one BH particle. We vary the numerical resolution, with many of our simulations using instead twice as many particles in each galaxy, and a subset of simulations with up to 128 times as many particles. We vary the initial seed mass of the black hole to identify any systematic dependence of our results on this choice. In most cases, we choose the seed mass either in accord with the observed $M_{\rm BH}$-σ relation or to be sufficiently small that its presence will not have an immediate dynamic effect. Given the particle numbers employed, the dark matter, gas, and star particles are all of roughly equal mass, and central cusps in the dark matter and bulge profiles are reasonably well resolved (see Fig. 2 in Springel et al. 2005b). The galaxies are then set to collide from a zero-energy orbit, and we vary the inclinations of the disks and the pericenter separation.

A representative example of the behavior of the simulations is provided in Figure 2, which shows the time sequence of a merger involving two bulgeless progenitor galaxies with virial velocities of 160 km s^{-1} and initial gas fractions of 20%. During the merger, gas is driven to the galaxy centers by gravitational tides, fueling nuclear starbursts and black hole growth. The quasar activity is short-lived and peaks twice in this merger, during both the first encounter and the final coalescence of the galaxies. To illustrate the bright, optically observable phase(s) of quasar activity that we identify below, we have added nuclear point sources in the center at the position(s) of the black hole(s) at times $T =$ 1.03, 1.39, and 1.48 Gyr, generating a surface density in correspondence to the relative luminosities of stars and quasar at these times. At other times, the accretion activity is either obscured or the black hole accretion rate is negligible. To make the appearance of the quasar visually more apparent, we have put a small part of its luminosity in "rays" around the quasar. These rays are artificial and are only a visual guide.

2.2. *Column Densities and Quasar Attenuation*

From the simulation outputs, we determine the obscuration of the black hole as a function of time during a merger by calculating the column density to a distant observer along many lines of sight. Typically, we generate $\sim$1000 radial lines of sight (rays), each with its origin at the black hole location and with directions uniformly spaced in solid angle $d\cos\theta\, d\phi$. For each ray, we begin at the origin and calculate and record the local gas properties using the SPH formalism and move a distance along the ray $\Delta r = \eta h_{\rm sml}$, where $\eta \leq 1$ and $h_{\rm sml}$ is the local SPH smoothing length. The process is repeated until a ray is sufficiently far from the origin ($\gtrsim$100 kpc) that the column has converged. We then integrate the gas properties along a particular ray to give the line-of-sight column density and mean metallicity. We have varied η and find empirically that gas properties along a ray converge rapidly and change smoothly for $\eta = 0.5$ and smaller. We similarly vary the number of rays and find that the distribution of line-of-sight properties converges for $\gtrsim$100 rays.

From the local gas properties, we use the multiphase model of the ISM described in Springel & Hernquist (2003) to determine the mass fraction in "hot" (diffuse) and "cold" (molecular and H I cloud core) phases of dense gas, and assuming pressure equilibrium, we obtain the local density of the hot and cold phases and their corresponding volume filling factors. The resulting values are in rough agreement with those of McKee & Ostriker (1977). Given a temperature for the warm, partially ionized component of the hot-phase $\sim$8000 K, determined by pressure equilibrium, we further calculate the neutral fraction of this gas, typically $\sim$0.3–0.5. We denote the neutral and total column densities as $N_{\rm H\,I}$ and $N_{\rm H}$, respectively. Using only the hot-phase density allows us to place an effective lower limit on the column density along a particular line of sight, as it assumes a given ray passes only through the diffuse ISM, with $\gtrsim$90% of the mass of the dense ISM concentrated in cold-phase "clumps." Given the small volume filling factor ($<$0.01) and cross section of cold clouds, we expect that the majority of sight lines will pass only through the hot-phase component.

Using $L_{\rm bol} = \epsilon_r \dot{M} c^2$, we model the intrinsic quasar continuum SED following Marconi et al. (2004), based on optical through hard X-ray observations (e.g., Elvis et al. 1994; George et al. 1998; Vanden Berk et al. 2001; Perola et al. 2002; Telfer et al. 2002; Ueda et al. 2003; Vignali et al. 2003), with a reflection component generated by the PEXRAV model (Magdziarz & Zdziarski 1995). This yields, for example, a B-band luminosity $\log(L_B/L_\odot) = 0.80 - 0.067\mathcal{L} + 0.017\mathcal{L}^2 - 0.0023\mathcal{L}^3$, where $\mathcal{L} = \log(L_{\rm bol}/L_\odot) - 12$, and we take $\lambda_B = 4400$ Å, but as we model the entire intrinsic SED, we can determine the bolometric correction in any frequency interval.

We then use a gas-to-dust ratio to determine the extinction along a given line of sight at optical frequencies. Observations suggest that the majority of reddened quasars have reddening curves similar to that of the Small Magellanic Cloud (SMC; Hopkins et al. 2004, Ellison et al. 2005), which has a gas-to-dust ratio lower than the Milky Way by approximately the same factor as its metallicity (Bouchet et al. 1985). Hence, we consider both a gas-to-dust ratio equal to that of the Milky Way, $(A_B/N_{\rm H\,I})_{\rm MW} = 8.47\times10^{-22}$ cm^2, and a gas-to-dust ratio scaled by metallicity, $A_B/N_{\rm H\,I} = (Z/0.02)(A_B/N_{\rm H\,I})_{\rm MW}$. In both cases we use the SMC-like reddening curve of Pei (1992). The form of the correction for hard X-ray (2–10 keV) and soft X-ray (0.5–2 keV) luminosities is similar to that of the B-band luminosity. We calculate extinction at X-ray frequencies (0.03–10 keV) using the photoelectric absorption cross sections of Morrison & McCammon (1983) and nonrelativistic Compton scattering cross sections, similarly scaled by metallicity. In determining the column density for photoelectric X-ray absorption, we ignore the inferred ionized fraction of the gas, as it is expected that the inner-shell electrons that dominate the photoelectric absorption edges will be unaffected in the temperature ranges of interest. We do not perform a full radiative transfer calculation, and therefore do not model scattering or reprocessing of radiation by dust in the infrared.

For a full comparison of quasar lifetimes and column densities obtained varying our calculation of $N_{\rm H}$, we refer to Hopkins et al. (2005a; see their Figs. 1, 5, and 6) and note their conclusion that, after accounting for clumping of most mass in the dense ISM in cold-phase structures, the column density does not depend sensitively on our assumptions of the small-scale physics of the ISM and obscuration; typically, the uncertainties in the resulting quasar lifetime as a function of luminosity are a factor of $\sim$2 at low luminosities in the B band, and smaller in, for example, the hard X-ray. Because our determination of the quasar luminosity functions is similar using the hard X-ray data alone or the hard X-ray, soft X-ray, and optical data simultaneously, the added uncertainties in our calculation of $\dot{n}(L_{\rm peak})$ in § 3.2 below due to

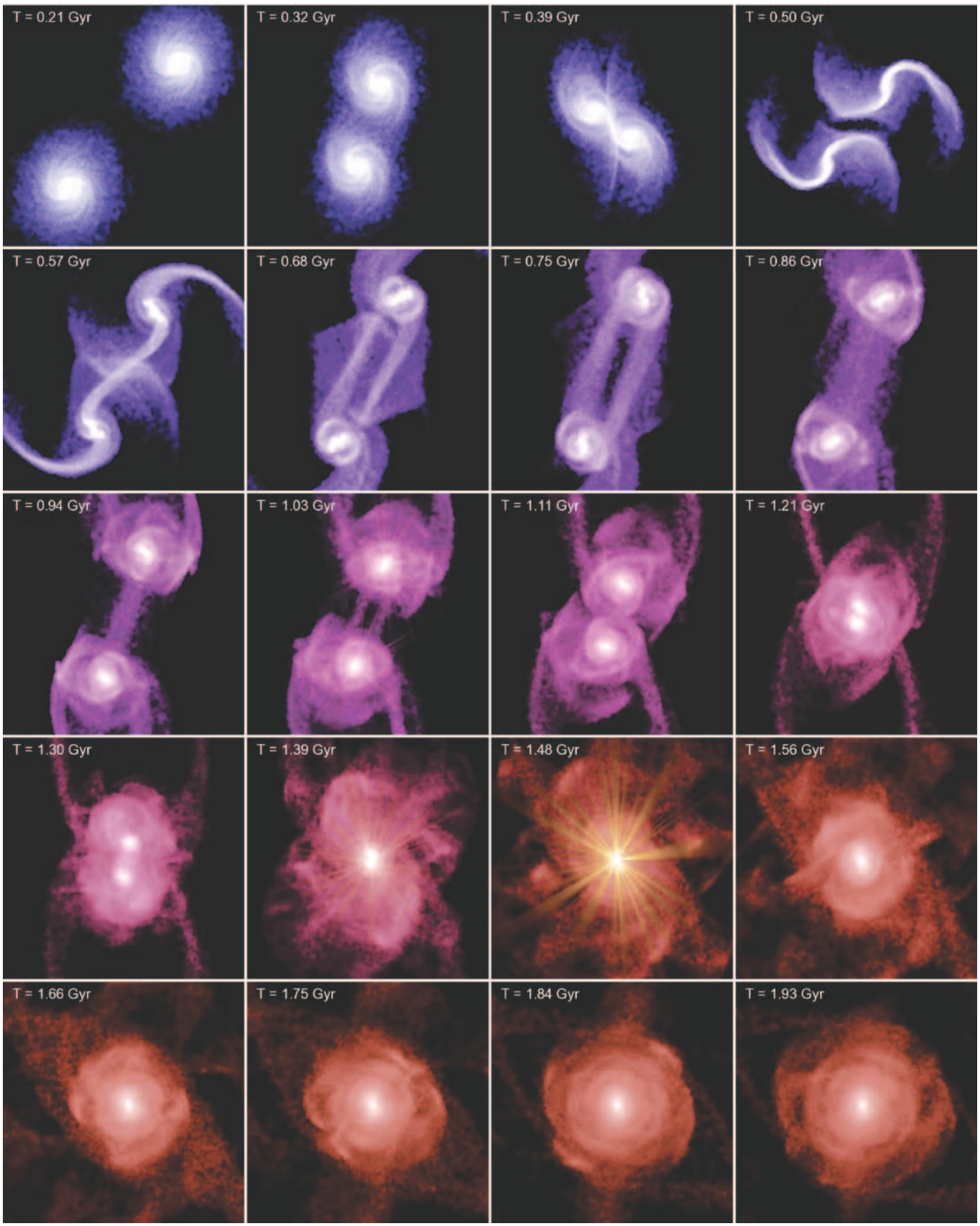

FIG. 2.—Time sequence from one of our merger simulations ($V_{\rm vir} = 160$ km s^{-1}, initial gas fraction 20%). Each panel is $80\,h^{-1}$ kpc on a side and shows the simulation time in the upper left corner. Brightness of individual pixels gives the logarithm of the projected stellar mass density, while color hue indicates the baryonic gas fraction, from 20% (*blue*) to less than 5% (*red*). At $T = 1.03$, 1.39, and 1.48 Gyr, when the black hole could be seen as an optical quasar, nuclear point sources are shown, providing a representation of the relative luminosities of stars and the quasar at these times.

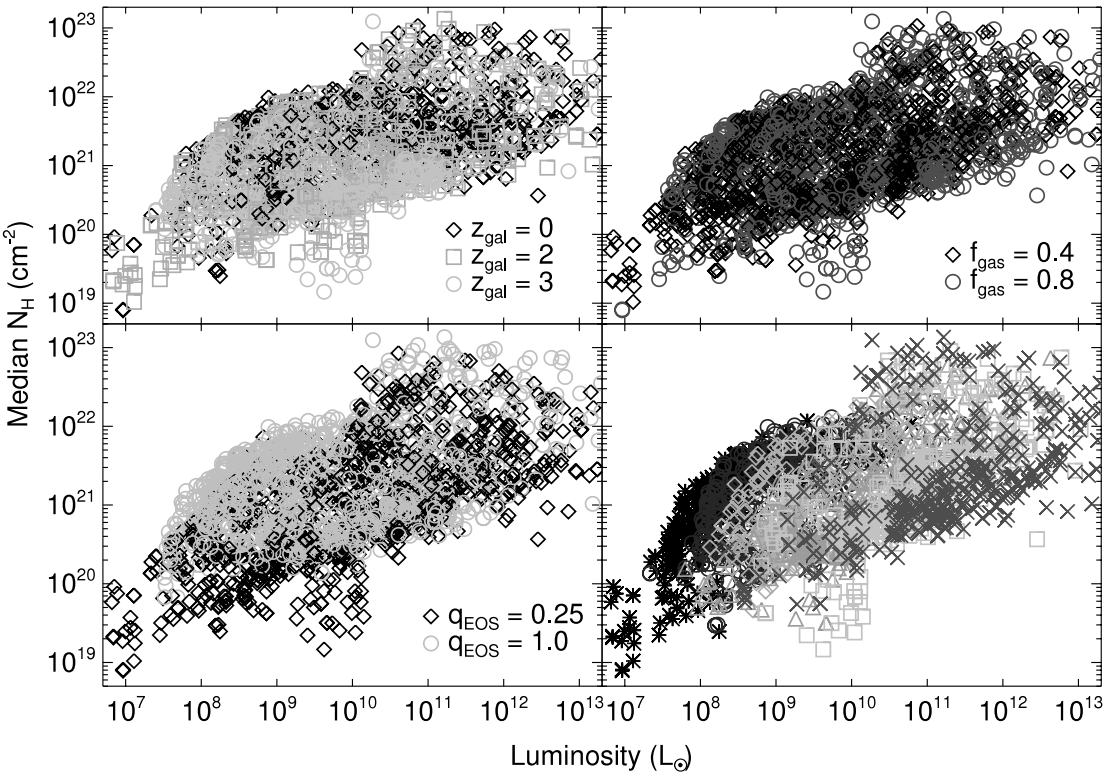

FIG. 3.—Median fitted total (neutral and ionized) column density $\bar{N}_H$ at each luminosity L in the snapshots from our series of simulations described in § 2. We compare changing concentrations and halo properties with redshift z_{gal} (*top left*), gas fractions f_{gas} (*top right*), the equation-of-state parameter q_{EOS} (*bottom left*), and virial velocity V_{vir} (*bottom right*). In the bottom right panel, simulations with $V_{vir} = 80$, 113, 160, 226, 320, and 500 km s^{-1} are shown as asterisks, circles, diamonds, triangles, squares, and crosses, respectively. Other than a possible weak sensitivity to q_{EOS}, the column density distribution as a function of luminosity shows no systematic dependence on any of the varied simulation parameters. [*See the electronic edition of the Supplement for a color version of this figure.*]

the uncertainty in our N_H calculation are small compared to the uncertainties due to degeneracies in the fitting procedure and uncertain bolometric corrections.

2.3. *The N_H Distribution as a Function of Luminosity*

Next, we consider the distribution of column densities as a function of both the instantaneous and peak quasar luminosities. For each simulation, we consider N_H values at all times with a given bolometric luminosity L (in some logarithmic interval in L), and determine the distribution of column densities at that L weighted by the total time along all sight lines with a given N_H. At each L, we approximate the simulated distribution and fit it to a lognormal form,

$$P(N_H) = \frac{1}{\sigma_{N_H}\sqrt{2\pi}} \exp\left[\frac{-\log^2(N_H/\bar{N}_H)}{2\sigma^2_{N_H}}\right]. \quad (1)$$

This provides a good fit for all but the brightest luminosities, where quasar feedback becomes important driving the blowout phase, and the quasar sweeps away surrounding gas and dust to become optically observable.

We show the resulting median column density $\bar{N}_H$ at each luminosity L in Figure 3. In the top left panel, simulations with $z_{gal} = 0$ are shown as diamonds, those with $z_{gal} = 2$ as squares, and those with $z_{gal} = 3$ as circles. In the top right, simulations with $f_{gas} = 0.4$ are shown as diamonds, those with $f_{gas} = 0.8$ as circles. In the bottom left, simulations with $q_{EOS} = 0.25$ are shown as diamonds, those with $q_{EOS} = 1.0$ as circles. And in the bottom right, simulations with $V_{vir} = 80$, 113, 160, 226, 320, and 500 km s^{-1} are shown as asterisks, circles, diamonds, triangles, squares, and crosses, respectively. Simulations with other values for these parameters (not shown for clarity, but see, e.g., Hopkins et al. 2005c) show similar trends.

While the increase in typical N_H values with luminosity appears to contradict observations suggesting that the obscured fraction decreases with luminosity, this is because the relationship shown above is dominated by quasars in growing, heavily obscured phases. In these stages, the relationship between column density and luminosity is a natural consequence of the fact that both are fueled by strong gas flows into the central regions of the galaxy; more gas inflow means higher luminosities, but also higher column densities. During these phases, the lognormal fits to column density as a function of instantaneous and peak luminosity presented in this section are reasonable approximations, but they break down in the brightest, short-lived stages of merger activity when the quasar rapidly heats the surrounding gas and drives a powerful wind, lowering the column density, resulting in a bright, optically observable quasar. Including in greater detail the effects of quasar blowout during the final stages of its growth in § 4, we find that this modeling actually predicts the observed decrease in obscured fraction with luminosity.

The relationship between N_H and L shows no strong systematic dependence on any of the simulation parameters considered. At most, there is weak sensitivity to q_{EOS}, in the sense that the simulations with $q_{EOS} = 1.0$ have slightly larger column densities at a given luminosity than those with $q_{EOS} = 0.25$. We derive an analytical model relating both the observed column density and quasar luminosity to the inflowing mass of gas in Hopkins et al. (2005c), by assuming that while it is growing, the black hole mass is proportional to the inflowing gas mass in the galaxy core (which ultimately produces the Magorrian et al. [1998] relation between

black hole and bulge mass), and assuming Bondi accretion, with obscuration along a sight line through this (spherically symmetric) gas inflow. Such a model gives the observed correlation between $N_{\rm H}$ and L, and explains the weak dependence of the column density–luminosity relation on the ISM gas equation of state. The assumptions above give a relationship of the form

$$N_{\rm H} \sim f_0 \frac{1}{m_{\rm H} R_c} \left(\frac{c_s}{c}\right) \left(\frac{cL}{G^2}\right)^{1/3}, \quad (2)$$

where $f_0 \sim 50$ is a dimensionless factor depending on the radiative efficiency, mean molecular weight, density profile, and assumed $M_{\rm BH}$-σ relation; $m_{\rm H}$ is the mass of hydrogen; R_c the radius of the galaxy core ($\sim$100 pc); and c_s is the effective sound speed in the central regions of the galaxy. A $q_{\rm EOS} = 1.0$ equation of state, with a higher effective temperature, results in a factor of $\approx$2 larger sound speed in the densest regions of the galaxy than a $q_{\rm EOS} = 0.25$ equation of state (Springel et al. 2005b), explaining the weak trend seen. In any event, the dependence is small compared to the intrinsic scatter for either equation of state in the value of $\bar{N}_{\rm H}$ at a given luminosity and further weakens at high luminosity, so it can be neglected. What may appear to be a systematic offset in $\bar{N}_{\rm H}$ with $V_{\rm vir}$ is actually just a tendency for larger $V_{\rm vir}$ systems to be at higher luminosities; there is no significant change in the dependence of $N_{\rm H}$ on L.

We use our large set of simulations to improve our fits (relative to those of Hopkins et al. 2005c) to the $N_{\rm H}$ distribution as a function of instantaneous and peak luminosities. Looking at individual simulations, there appears to be a "break" in the power-law scaling of $\bar{N}_{\rm H}$ with L at $L \sim 10^{11}\ L_\odot$. We find that the best fit to the median column density $\bar{N}_{\rm H}$ is then

$$\bar{N}_{\rm H} = \begin{cases} 10^{22.8}\ {\rm cm}^{-2} \left(\dfrac{L}{L_{\rm peak}}\right)^{0.54} & L < 10^{11}\ L_\odot, \\ 10^{21.9}\ {\rm cm}^{-2} \left(\dfrac{L}{10^{11}\ L_\odot}\right)^{0.43} & L > 10^{11}\ L_\odot. \end{cases} \quad (3)$$

Either of these two relations provides an acceptable fit to the plotted $\bar{N}_{\rm H}$ distribution if applied to the entire luminosity range ($\chi^2/\nu \approx 2.8,\ 3.2$ for the first and second relations, respectively), but their combination provides a significantly better fit ($\chi^2/\nu \approx 1.5$), although it is clear from the large scatter in $\bar{N}_{\rm H}$ values that any such fit is a rough approximation. Despite the complicated form of this equation, it is, in practice, similar to our $\bar{N}_{\rm H} \propto L^{0.35}$ fit from previous work and $\bar{N}_{\rm H} \propto L^{1/3}$ analytical scaling over the range of relevant luminosities, but is more accurate by a factor $\sim$2–3 at low ($\lesssim 10^9\ L_\odot$) luminosities. For comparison, however, we do consider this simpler form for $N_{\rm H}(L)$, as well as our more accurate fit above in our subsequent analysis, and find that it makes little difference to most observable quasar properties. At the highest luminosities, near the peak luminosities of the brightest quasars, the scatter about these fitted median $\bar{N}_{\rm H}$ values increases, and as noted above, the impact of the quasar in expelling surrounding gas becomes important and column densities vary rapidly. We consider this blowout phase in more detail in § 4.

We find that any dependence of $\sigma_{N_{\rm H}}$ (the fitted lognormal dispersion) on L or $L_{\rm peak}$ is not statistically significant, with approximately constant $\sigma_{N_{\rm H}} \approx 0.4$ for individual simulations. We similarly find no systematic dependence of $\sigma_{N_{\rm H}}$ on any of our varied simulation parameters. However, it is important to note that while the dispersion in $N_{\rm H}$ for an individual simulation is $\sigma_{N_{\rm H}} \approx 0.4$, the dispersion in $\bar{N}_{\rm H}$ across all simulations at a given luminosity is large, $\sim$1 dex. Thus, we fit the effective $\sigma_{N_{\rm H}}$ at a given luminosity for the *distribution* of quasars and find that it is $\sigma_{N_{\rm H}} \approx 1.2$. Although we have slightly revised our fits for greater accuracy at low luminosities, we note that this relation is shallower than the relation $N_{\rm H} \propto L$ roughly expected if $M_{\rm BH}$ is constant ($L \propto \rho \propto N_{\rm H}$) or $L \propto M_{\rm BH}$ always, and strongly contrasts with unification models that predict static obscuration or evolutionary models in which $N_{\rm H}$ is independent of L up to some threshold (e.g., Fabian 1999).

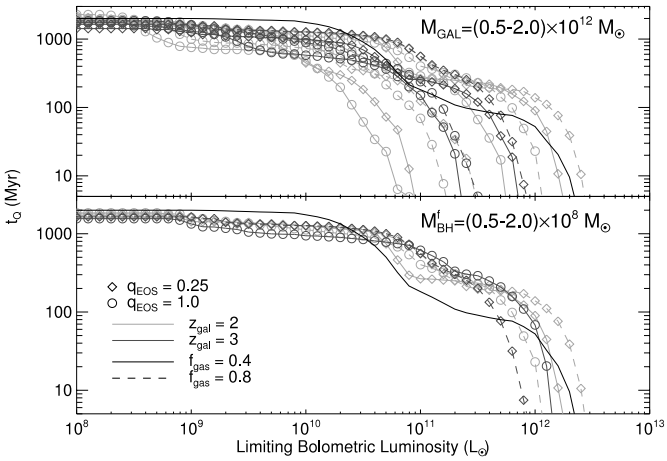

FIG. 4.—Integrated intrinsic quasar lifetime above a given reference bolometric luminosity, $t_Q(L)$, as a function of luminosity for simulations with host galaxies with total mass $M_{\rm gal} = (0.5-2.0) \times 10^{12}\ M_\odot$ (*top*), and simulations with final black hole masses $M^f_{\rm BH} = (0.5-2.0) \times 10^8\ M_\odot$ (i.e., similar peak luminosity $L_{\rm peak} \sim 10^{12}\ L_\odot$; *bottom*). The simulations cover a range in equation-of-state parameter $q_{\rm EOS}$, initial disk gas fraction $f_{\rm gas}$, galaxy redshift (for scaling of halo properties) $z_{\rm gal}$, and virial velocities $V_{\rm vir} = 113-160$ km s^{-1}. The solid line in both cases is for a merger involving Milky Way–like galaxy models, which we refer to as A3, with $f_{\rm gas} = 1,\ q_{\rm EOS} = 1,\ z_{\rm gal} = 0,$ and $V_{\rm vir} = 160$ km s^{-1}. [*See the electronic edition of the Supplement for a color version of this figure.*]

2.4. *Quasar Lifetimes and Sensitivity to Simulation Parameters*

We define the luminosity-dependent quasar lifetime $t_Q = t_Q(L_{\rm min})$ as the time a quasar has a luminosity above a certain reference luminosity $L_{\rm min}$, i.e., the total time the quasar shines at $L \geq L_{\rm min}$. For ease of comparison across frequencies, we measure the lifetime in terms of the bolometric luminosity L rather than, for example, the B-band luminosity. Knowing the distribution of column densities $N_{\rm H}$ as a function of luminosity and system properties (see § 2.3), we can then analytically or numerically calculate the distribution of observed lifetimes at any frequency if we know this intrinsic lifetime. Below $\sim$1 Myr, our estimates of t_Q become uncertain owing to the effects of quasar variability and our inability to resolve the local small-scale physics of the ISM, but this is significantly shorter than even the most rapid timescales, $\sim$10 Myr, of substantial quasar evolution.

As before, we use our diverse sample of simulations to test for systematic effects in our parameterization of the quasar lifetime. Figure 4 shows the quasar lifetime as a function of reference luminosity $L_{\rm min}$ for both a set of simulations with similar total galaxy mass, $M_{\rm gal} \approx 10^{12}\ M_\odot$, and one with similar final black hole mass (i.e., similar peak quasar luminosity), $M^f_{\rm BH} \approx 10^8\ M_\odot$. In each case, the simulations cover a range in $q_{\rm EOS}, f_{\rm gas}, z_{\rm gal}$, and $V_{\rm vir}$.

As Figure 4 demonstrates, at a given $M_{\rm gal}$ there is a wide range of lifetimes, with a systematic dependence on several quantities. For example, for fixed $M_{\rm gal}$, a lower $q_{\rm EOS}$ means that the gas is

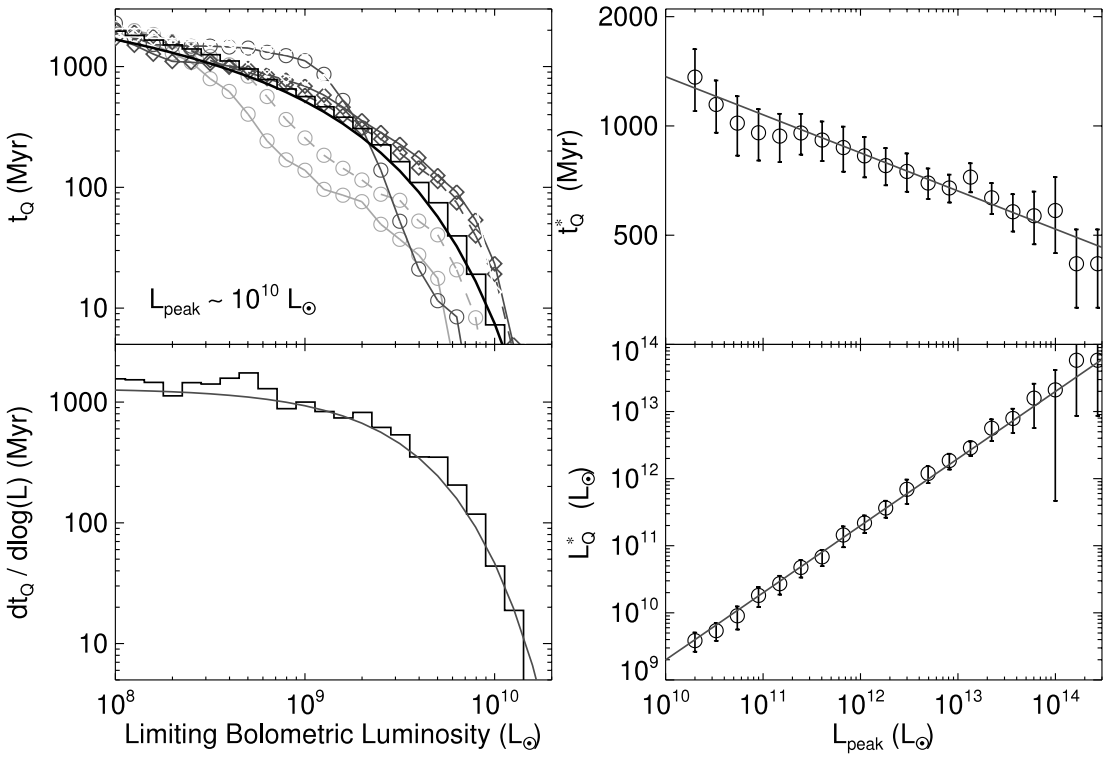

FIG. 5.—Fits to the quasar lifetime as a function of luminosity from our simulations. The top left panel shows the intrinsic, bolometric quasar lifetime t_Q of a set of simulations with L_{peak} within a factor of 2 of $10^{10}\ L_\odot$, in the manner of Fig. 4. The histogram shows the geometric mean of these lifetimes, and the histogram in the bottom left panel shows the differential lifetime $dt/d\log L$ from this geometric mean. The thick solid line in the top left panel and solid line in the bottom left panel show the best-fit to our analytical form, $dt/d\log L = t_Q^* \exp(-L/L_Q^*)$. The top right panel shows the fitted t_Q^* and resulting errors in each peak luminosity (final black hole mass) interval, and the best-fit power law to $t_Q^*(L_{peak})$ (*solid line*). The bottom right panel shows the fitted L_Q^* and resulting errors in each peak luminosity (final black hole mass) interval, and the best-fit proportionality $L_Q^* \propto L_{peak}$ (*solid line*). [*See the electronic edition of the Supplement for a color version of this figure.*]

less pressurized and more easily collapses to high density, resulting in larger M_{BH}^f and longer lifetimes at higher luminosities. Similarly, higher f_{gas} provides more fuel for black hole growth at fixed M_{gal}. However, for a given M_{BH}^f, the lifetime t_Q as a function of L_{min} is similar across simulations and shows no systematic dependence on any of the varied parameters. We find this for all final black hole masses in our simulations, in the range $M_{BH}^f \sim 10^6$–$10^{10}\ M_\odot$. We have further tested this as a function of resolution, comparing with alternate realizations of our fiducial A3 simulation with up to 128 times as many particles, and find similar results as a function of M_{BH}^f.

From Figure 4 it is clear that the final black hole mass or peak luminosity is a better variable to use in describing the lifetime than the host galaxy mass. The lack of any systematic dependence of either the quasar lifetime or $N_H(L, L_{peak})$ on host galaxy properties implies that our earlier results (Hopkins et al. 2005a, 2005b, 2005c, 2005d) are reliable and can be applied to a wide range of host galaxy properties, redshifts, and luminosities, although we refine and expand the various fits of these works and their applications herein. Furthermore, the large scatter in t_Q at a given galaxy mass has important implications for the quasar correlation function as a function of luminosity, as one cannot associate a single quasar luminosity with hosts of a given mass (see Lidz et al. 2006).

Although the truncated power laws we have previously fitted to t_Q using only the A-series simulations (Hopkins et al. 2005a) provide acceptable fits to all our runs, we use our new, larger set of simulations to improve the accuracy of the fits and average over peculiarities of individual simulations, giving a more robust prediction of the lifetime as a function of instantaneous and peak luminosity. For a given peak luminosity L_{peak}, we consider simulations with an L_{peak} within a factor of 2 and take the geometric mean of their lifetimes $t_Q(L)$ (we ignore any points where $t_Q < 1$ Myr, as our calculated lifetimes are uncertain below this limit). We can then differentiate this numerically to obtain $dt/d\log L$ (the time spent in a given logarithmic luminosity interval), and fit some functions to both curves simultaneously. Figure 5 illustrates this and shows the results of our fitting. We find that both the integrated lifetime $t_Q(L)$ and the differential lifetime $dt/d\log L$ are well fitted by an exponential,

$$dt/d\log L = t_Q^* \exp\left(-L/L_Q^*\right), \tag{4}$$

where both t_Q^* and L_Q^* are functions of M_{BH}^f or L_{peak}. The best-fit such $dt/d\log L$ is shown in the figure as a solid line for simulations with $L_{peak} \sim 10^{10}\ L_\odot$ and agrees well with both the numerical derivative $dt/d\log L$ (*bottom left*, *histogram*) and the geometric mean $t_Q(L)$ (*top left*, *histogram*). This of course implies

$$t_Q(L) = t_Q^* \int_L^{L_{peak}} e^{-L/L_Q^*}\, d\log L, \tag{5}$$

but we are primarily interested in $dt/d\log L$ in our subsequent analysis.

Although our fitted lifetime involves an exponential, it is in no way similar to the exponential light curve of constant-Eddington-ratio black hole growth or the model in, e.g., Haiman & Loeb (1998), which gives $dt/d\log L = \text{constant} \sim t_S \ll t_Q^*$.

Our functional form also has the advantage that although it should formally be truncated with $dt/d\log L = 0$ for $L > L_{peak}$, the values in this regime fall off so quickly that we can safely use

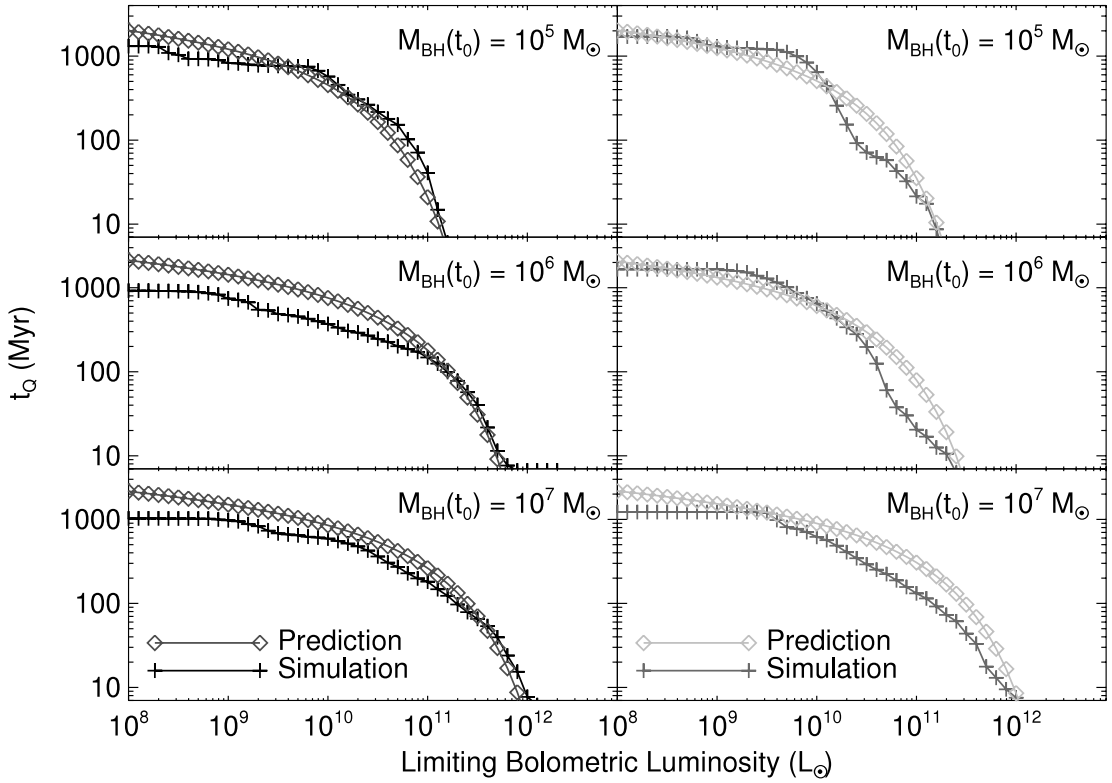

FIG. 6.—Predicted quasar lifetime as a function of luminosity compared to that obtained in simulations with and without bulges and with different initial seed black hole masses. All simulations shown in this plot are initially identical to our fiducial A3 (Milky Way like) case, but with or without an initial stellar bulge and with an initial seed black hole mass as labeled. Diamonds show the predicted quasar lifetime t_Q, a function of the peak luminosity of each simulated quasar, determined from the fits shown in Fig. 5. Crosses show the lifetime determined directly in the simulations. [*See the electronic edition of the Supplement for a color version of this figure.*]

the above fit for all large L. Similarly, at $L \lesssim 10^{-4} L_{\rm peak}$, $dt/d\log L$ falls below the constant t_Q^* to which this equation asymptotes. Furthermore, in this regime, the fits above begin to differ significantly from those obtained by fitting, for example, truncated power laws or Schechter functions. However, these luminosities are well below those we generally consider and well below the luminosities where the contribution of a quasar with some $L_{\rm peak}$ is significant to the observed quantities we predict. Moreover, this turndown (i.e., the lower value predicted by an exponential as opposed to a power-law or Schechter function at low luminosities) is at least in part an artifact of the finite simulation duration. The values here are also significantly more uncertain, as by these low relative accretion rates, the system is likely to be accreting in some low-efficiency, ADAF (advection-dominated accretion flow) state (e.g., Narayan & Yi 1995), which we do not implement directly in our simulations. Rather than introduce additional uncertainties into our modeling when they do not affect our predictions, we adopt these exponential fits that are accurate at $L \gtrsim (10^{-4}$ to $10^{-3}) L_{\rm peak}$. However, for purposes where the faint-end behavior of the quasar lifetime is important, such as predicting the value and evolution of the faint-end quasar luminosity function slope with redshift, a more detailed examination of the lifetime at low luminosities and relaxation of quasars after the blowout phase is necessary, and we consider these issues separately in Hopkins et al. (2006a).

We also note that in Hopkins et al. (2005b) we considered several extreme limits to our modeling, neglecting all times before the final merger and applying an ADAF correction at low accretion rates (taken into account a posteriori by rescaling the radiative efficiency ϵ_r with accretion rate, given the assumption that such low accretion rates do not have a large dynamical effect on the system regardless of radiative efficiency), and found that this does not change our results: the lifetime at low luminosities may be slightly altered but the key qualitative point, that the quasar lifetime increases with decreasing luminosity, is robust against a wide range of limits designed to decrease the lifetime at low luminosities.

Figure 5 further shows the fitted t_Q^* (*top right*) and L_Q^* (*bottom right*) as a function of peak quasar luminosity for each $L_{\rm peak}$. We find that L_Q^*, the luminosity above which the lifetime rapidly decreases, is proportional to $L_{\rm peak}$,

$$L_Q^* = \alpha_L L_{\rm peak}, \tag{6}$$

with a best-fit coefficient $\alpha_L = 0.20$ (*solid line*). The weak dependence of t_Q^* on $L_{\rm peak}$ is well described by a power law,

$$t_Q^* = t_*^{(10)} \left(\frac{L_{\rm peak}}{10^{10}\ L_\odot} \right)^{\alpha_T}, \tag{7}$$

with $t_*^{(10)} = 1.37 \times 10^9$ yr and $\alpha_T = -0.11$.

The presence or absence of a stellar bulge in the progenitors can have a significant impact on the quasar light curve (Springel et al. 2005b), primarily affecting the strength of the strong accretion phase associated with initial passage of the merging galaxies (e.g., Mihos & Hernquist 1994). Likewise, the seed mass of the simulation black holes could have an effect, as black holes with smaller initial masses will spend more time growing to large sizes, and more massive black holes may be able to shut down early phases of accretion in mergers in minor blowout events. In Figure 6 we show various tests to examine the robustness of our fitted quasar lifetimes to these variations. We have rerun our

fiducial Milky Way–like A3 simulation both with (*right panels*) and without (*left panels*) initial stellar bulges in the merging galaxies and varying the initial black hole seed masses from 10^4 to $10^7\ M_\odot$. In each case we compare the lifetime t_Q determined directly from the simulations (*crosses*) to that predicted from our fits above (*diamonds*), based only on the peak luminosity (final black hole mass) of the simulated quasar. Again, we find that varying these simulation parameters can have a significant effect on the final black hole mass, but that the quasar lifetime as a function of peak luminosity is a robust quantity, independent of initial black hole mass or the presence or absence of a bulge in the quasar host.

We can integrate the total radiative output of our model quasars,

$$E_{\rm rad} = \int_{L_{\rm min}}^{L_{\rm peak}} L \frac{dt}{d\log L}\, d\log L, \tag{8}$$

and using our fitted formulae and $L_{\rm min} \ll L_Q^*$, we find

$$E_{\rm rad} = L_Q^* t_Q^* \log e \left(1 - e^{-L_{\rm peak}/L_Q^*}\right). \tag{9}$$

Knowing $E_{\rm rad} = \epsilon_r M_{\rm BH}^f c^2$, we can compare the final black hole mass as a function of peak luminosity to what we would expect if the peak luminosity were the Eddington luminosity of a black hole with mass $M_{\rm Edd}$, $L_{\rm Edd} = \epsilon_r M_{\rm Edd} c^2/t_{\rm S}$, where $t_{\rm S}$ is the Salpeter time for $\epsilon_r = 0.1$. Equating $E_{\rm rad} = \epsilon_r M_{\rm BH}^f c^2$ with the value calculated in equation (9), and using the definition of the Eddington mass at $L = L_{\rm peak}$ and our fitted $L_Q^* = \alpha_L L_{\rm peak}$, we obtain

$$\frac{M_{\rm BH}^f(L_{\rm peak})}{M_{\rm Edd}(L_{\rm peak})} = \alpha_L \left(\frac{t_Q^*}{t_{\rm S}}\right) \log e \approx 1.24 f_T, \tag{10}$$

where $f_T = (L_{\rm peak}/10^{13}\ L_\odot)^{-0.11}$ for the power-law fit to t_Q^*. For our calculations explicitly involving black hole mass, we adopt this conversion unless otherwise noted, as we have performed our primary calculation [i.e., calculated $\dot{n}(L_{\rm peak})$ in terms of peak luminosity]. Moreover, although this agrees well with the black hole masses in our simulations as a function of peak luminosity (as it must if the fitted quasar lifetimes are accurate), this allows us to smoothly interpolate to the highest black hole masses ($\sim$ a few $\times 10^9$–$10^{10}\ M_\odot$), which are of particular interest in examining the black hole population but for which the number of simulations we have with a given final black hole mass drops rapidly.

This gives explicitly the modifications to the black hole mass compared to that inferred from the light-bulb and constant-Eddington-ratio models that we outline below in § 2.5, in which quasars shine at constant luminosity or follow exponential light curves and for which $M_{\rm BH}^f = M_{\rm Edd}(L_{\rm peak})/l$, where l, the (constant) Eddington ratio, is generally adopted. The corrections are small, and therefore most of the black hole mass is accumulated in the bright, near-peak quasar phase, in good agreement with observational estimates (e.g., Soltan 1982; Yu & Tremaine 2002); we discuss this in greater detail in §§ 4 and 6. Furthermore, the increase of f_T with decreasing $L_{\rm peak}$ implies that lower mass quasars accumulate a larger fraction of their mass in slower, subpeak accretion after the final merger, while high-mass objects acquire essentially all their mass in the peak quasar phase. This is seen directly in our simulations and is qualitatively in good agreement with expectations from simulations and semianalytical models in which the $M_{\rm BH}$-σ relation is set by black hole feedback in a strong quasar phase. Compared to the assumption that $M_{\rm BH}^f = M_{\rm Edd}(L_{\rm peak})$, this formula introduces a small but nontrivial correction in the relic supermassive black hole mass function implied by the quasar luminosity function and $\dot{n}(L_{\rm peak})$ (see § 6).

The predictions of our model for the quasar lifetime and evolution can be applied to observations that attempt to constrain the quasar lifetime from individual quasars, for example using the proximity effect in the Lyα forest (Bajtlik et al. 1988; Haiman & Cen 2002; Jakobsen et al. 2003; Yu & Lu 2005) and multiepoch observations (Martini & Schneider 2003). However, many observations designed to constrain the quasar lifetime do so not for individual quasars, but using demographic or integral arguments based on the population of quasars in some luminosity interval (e.g., Soltan 1982; Haehnelt et al. 1998; Yu & Tremaine 2002; Yu & Lu 2004; Porciani et al. 2004; Grazian et al. 2004). Our prediction for these observations is similar but slightly more complex, as an observed luminosity function at a given luminosity will consist of sources with different peak luminosities $L_{\rm peak}$, but the same instantaneous luminosity L. Furthermore, the lifetime being probed may be either the integrated quasar lifetime above some luminosity threshold or the differential lifetime at a particular luminosity.

For a given determination of the quasar luminosity function using our model for quasar lifetimes and some distribution of peak luminosities, we can predict the distribution of quasar lifetimes as a function of the observed luminosity interval. Figure 7 shows an example of such a result, using the determination of the luminosity function below in § 3.2, at redshift $z = 0.5$. We consider several bolometric luminosities spanning the luminosity function from 10^9 to $10^{14}\ L_\odot$, and for each, the distribution of sources (peak luminosities) and the corresponding distribution of quasar lifetimes. We show both the distribution of integrated quasar lifetimes t_Q (*left panel*) and the distribution of differential quasar lifetimes $dt/d\log L$ (*right panel*). The evolution with redshift is weak, with the lifetime increasing by $\sim$1.5–2 at a given luminosity at $z = 2$. There is furthermore an ambiguity of a factor $\sim$2, as some of the quasars observed at a given luminosity will only be entering a peak quasar phase, whereas the lifetimes shown are integrated over the whole quasar evolution. This prediction is quite different from that of the optical quasar phase that we describe below in § 4 and in Hopkins et al. (2005d), as it considers only the intrinsic bolometric luminosity, but our modeling and the fits provided above for the bolometric lifetime and column density distributions should enable the prediction of these quantities, considering attenuation, in any wave band. In either case, it is clear that the lifetime distribution for lower luminosity quasars is increasingly more strongly peaked and centered around longer lifetimes, in good agreement with the limited observational evidence from, e.g., Adelberger & Steidel (2005). This is a consequence of the fact that in our model, quasar lifetimes decrease with increasing luminosity. The range spanned in the figure corresponds well to the range of quasar lifetimes implied by the observations above and others (e.g., Martini 2004 and references therein).

2.5. *Alternative Models of Quasar Evolution*

Our modeling reproduces at least the observed hard X-ray quasar luminosity function by construction, since we use the observed quasar luminosity functions to determine the birthrate of quasars of a given $L_{\rm peak}$, $\dot{n}(L_{\rm peak})$, in § 3.2. It is therefore useful to consider in detail the differences in our subsequent predictions between various models for the quasar lifetime and obscuration, in order to determine to what extent these predictions are implied by any model that successfully reproduces the observed quasar luminosity function, and to what extent they are independent of the observed luminosity functions and instead depend on the

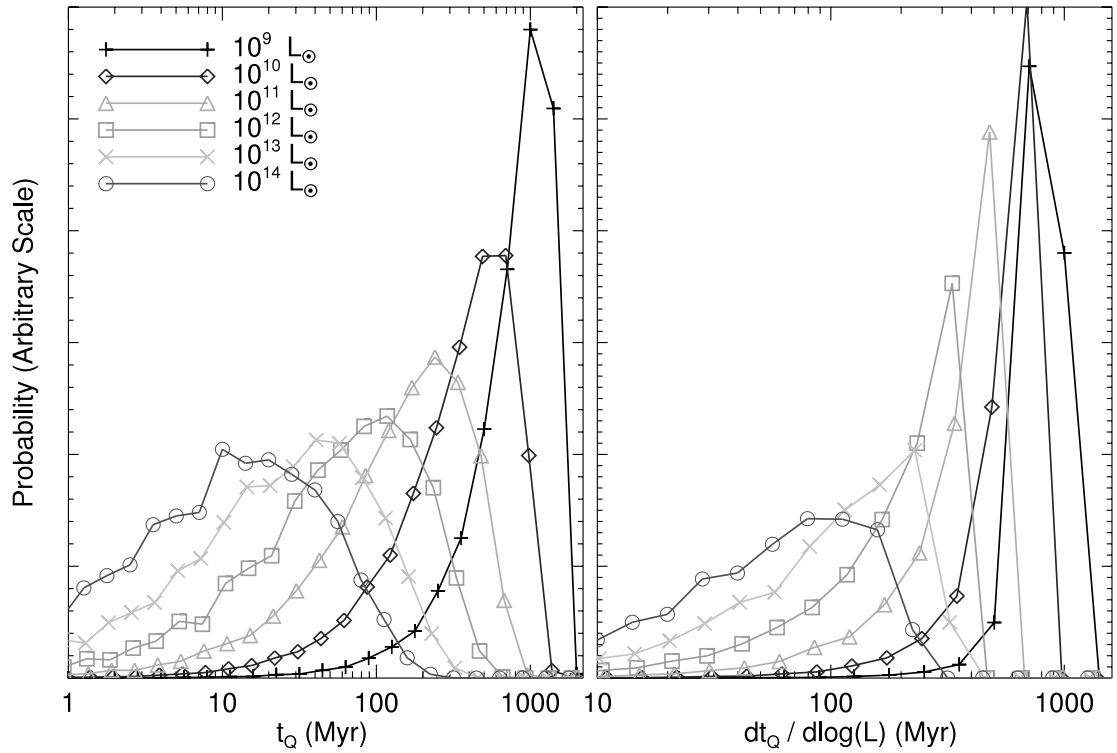

FIG. 7.—Predicted distribution (fractional number density per logarithmic interval in lifetime) of quasar lifetimes at different bolometric luminosities, for the luminosity function determined in § 3 at $z = 0.5$. The left panel plots the distribution of integrated lifetimes t_Q (time spent over the course of each quasar lifetime above the given luminosity). The right panel plots the distribution of differential lifetimes $dt/d\log L$ (time spent by each quasar in a logarithmic interval about the given luminosity). [*See the electronic edition of the Supplement for a color version of this figure.*]

model of quasar evolution adopted. To this end, we define two models for the quasar lifetime and two models for the distribution of quasar column densities, combinations of which have been commonly used in most previous analyses of quasars.

For the quasar lifetime, we consider the following two cases:

Light-bulb model (e.g., Small & Blandford 1992; Kauffmann & Haehnelt 2000; Wyithe & Loeb 2003; Haiman et al. 2004).—The simplest possible model for the quasar light curve, the "feast or famine," or "light bulb," model assumes that quasars have only two states, "on" and "off." Quasars turn "on," shine at a fixed bolometric luminosity $L = L_{\rm peak}$, defined by a "constant" Eddington ratio (i.e., $L_{\rm peak} = lM_{\rm BH}^{f}$) and constant quasar lifetime $t_{Q,\,\rm LB}$. Models where quasars live arbitrarily long with slowly evolving mean volume emissivity or mean light curve (e.g., Small & Blandford 1992; Haiman & Menou 2000; Kauffmann & Haehnelt 2000) are equivalent to the light-bulb scenario, as they still assume that quasars observed at a luminosity L radiate at that approximately constant luminosity over some universal lifetime $t_{Q,\rm LB}$ at a particular redshift. We adopt $l = 0.3$ and $t_{Q,\,\rm LB} = 10^7$ yr, as is commonly assumed in theoretical work and suggested by observations (given this prior) (e.g., Yu & Tremaine 2002; Martini 2004; Soltan 1982; Yu & Lu 2004; Porciani et al. 2004; Grazian et al. 2004), and similar to the *e*-folding time of a black hole with canonical radiative efficiency $\epsilon_r = 0.1$ (Salpeter 1964) or the dynamical time in a typical galactic disk or central regions of the merger. These choices control only the normalization of $\dot{n}(L_{\rm peak})$ and therefore do not affect most of our predictions. Where the normalization (i.e., value of the constant t_Q or l) is important, we allow it to vary in order to produce the best possible fit to the observations.

Exponential (fixed Eddington ratio) model.—A somewhat more physical model of the quasar light curve is obtained by assuming growth at a constant Eddington ratio, as is commonly adopted in, for example, semianalytical models that attempt to reproduce quasar luminosity functions (e.g., Kauffmann & Haehnelt 2000; Wyithe & Loeb 2003; Volonteri et al. 2003). In this model, a black hole accretes at a fixed Eddington ratio l from an initial mass M_i to a final mass M_f [or equivalently, a final luminosity $L_f = lL_{\rm Edd}(M_f)$], and then shuts off. This gives exponential mass and luminosity growth, and the time spent in any logarithmic luminosity bin is constant,

$$dt/d\log(L) = t_{\rm S}[\ln(10)/l] \tag{11}$$

for $L_i < L < L_f$. This is true for any exponential light curve (i.e., this model includes cases with an exponential *decline* in quasar luminosity) $f(t) \propto e^{\pm t/t_*}$, such as that of Haiman & Loeb (1998), with only the normalization $dt/d\log(L) = t_* \ln(10)$ changed, and thus any such model will give identical results with correspondingly different normalizations. As with the light-bulb model, we are free to choose the characteristic Eddington ratio and corresponding timescale for this light curve, and we adopt $l = 0.3$ (i.e., $t_* \sim 10^8$ yr) in general. Again, however, we allow the normalization to vary freely where it is important, such that these models have the best chance to reproduce the observations. For our purposes, models in which this timescale is determined by, for example, the galaxy dynamical time and thus are somewhat dependent on host galaxy mass or redshift are nearly identical to this scenario. Further, insofar as the dynamical time increases weakly with increasing host galaxy mass (as, e.g., for a spheroid with $M_{\rm BH} \propto M_{\rm vir} \sim a\sigma^2/G$, where a is the spheroid scale length and $M_{\rm BH} \propto \sigma^4$, such that $t_{\rm dyn} \sim a/\sigma \propto \sigma \propto M_{\rm vir}^{1/4}$), this produces behavior qualitatively opposite to our predictions (of increasing lifetime with decreasing instantaneous luminosity) and yields

results that are even more discrepant from our predictions and the observations than the constant (host galaxy independent) case.

A wide variety of light-bulb or exponential (constant Eddington ratio) models are possible, allowing for different distributions of typical Eddington ratios and/or quasar lifetimes (see, e.g., Steed & Weinberg 2003 for an extensive comparison of several classes of such models), but for our purposes they are essentially identical insofar as they do not capture the essential qualitative features of our quasar lifetimes, namely, that the quasar lifetime depends on both instantaneous and peak luminosities and increases with decreasing instantaneous luminosity.

We fit both of the simple models above to the observed quasar luminosity functions in the same manner described in § 3 (i.e., in the same manner as we fit our more complicated models of quasar evolution) to determine $\dot{n}(L_{\rm peak})_{\rm LB}$ for the light-bulb model and $\dot{n}(L_{\rm peak})_{\rm Edd}$ for the fixed-Eddington-ratio model (see eqs. [15] and [16], respectively). Thus all three models of the quasar light curve, the light-bulb and fixed-Eddington-ratio models and our luminosity-dependent lifetimes model, produce an essentially identical bolometric luminosity function.

We also consider two commonly adopted alternative models for the column density distribution and quasar obscuration:

Standard (luminosity independent) torus (e.g., Antonucci 1993).—This is the canonical obscuration model, based on observations of local, low-luminosity Seyfert galaxies (e.g., Risaliti et al. 1999). The column density distribution is derived from the torus geometry, where we assume the torus inner radius lies at a distance R_T from the black hole, with a height H_T and a density distribution $\rho(\theta) \propto \exp(-\gamma|\cos\theta|)$, where θ is the polar angle and the torus lies in the $\theta = 0$ plane. This results in a column density as a function of viewing angle of

$$N_{\rm H}(\theta) = N_{\rm H,\,0}\exp(-\gamma|\cos\theta|)\cos(90-\theta) \times\sqrt{\left(\frac{R_T}{H_T}\right)^2 - \sec^2(90-\theta)\left[\left(\frac{R_T}{H_T}\right)^2 - 1\right]} \quad (12)$$

(Treister et al. 2004). Here $N_{\rm H,\,0}$ is the column density along a line of sight through the torus in the equatorial plane and γ parameterizes the exponential decay of density with viewing angle. This is a phenomenological model, and as a result the parameters are essentially all free. We adopt typical values, an equatorial column density $N_{\rm H,\,0} = 10^{24}$ cm^{-2}, radius-to-height ratio $R_T/H_T = 1.1$, and density profile $\gamma = 4$. This combination of parameters follows Treister et al. (2004) and is designed to fit the observed X-ray column density distribution and give a ratio of obscured to unobscured quasars of $\sim$3, similar to the mean locally observed value (e.g., Risaliti et al. 1999).

Receding (luminosity dependent) torus (e.g., Lawrence 1991).—Many observations suggest that the fraction of obscured objects depends on luminosity (Steffen et al. 2003; Ueda et al. 2003; Hasinger 2004; Grimes et al. 2004; Sazonov & Revnivtsev 2004; Barger et al. 2005; Simpson 2005). Therefore, some theoretical works have adopted a "receding torus" model, in which the torus radius R_T (i.e., distance from the quasar) is allowed to vary with luminosity, but the height and other parameters remain constant. The torus radius is assumed to increase with luminosity, enlarging the opening angle and thus the fraction of unobscured quasars. In this case, the column densities are identical to those shown above, but now $R_T/H_T = (L/L_0)^{0.5}$, where $L_0 \approx 10^{11}\ L_\odot$ is the luminosity at which the ratio of obscured to unobscured quasars is $\approx 3:1$ and the power-law slope is chosen to fit the dependence of obscured fraction on luminosity.

Both of these column density distributions represent phenomenological models with several free parameters, explicitly chosen to reproduce the observed differences in quasar luminosity functions and column density distributions. Despite this, it is not clear that these functional forms represent the best possible fit to the observations they are designed to reproduce. Furthermore, comparison of our results in which column density distributions depend on luminosity and peak luminosity elucidates the importance of proper modeling of the dependence of column density on quasar evolution.

3. THE QUASAR LUMINOSITY FUNCTION

3.1. *The Effect of Luminosity-dependent Quasar Lifetimes*

Given quasar lifetimes as functions of both instantaneous and peak luminosities, the observed quasar luminosity function (in the absence of selection effects) is a convolution of the lifetime with the intrinsic distribution of sources with a given $L_{\rm peak}$. If sources of a given L are created at a rate $\dot{n}(L, t)$ (per unit comoving volume) at cosmological time $t_{\rm H} \sim 1/H(z)$ and live for some lifetime $\Delta t_Q(L)$, the total comoving number density observed will be

$$\Delta n = \int_{t_{\rm H}}^{t_{\rm H}+\Delta t_Q(L)} \dot{n}(L, t)\, dt, \quad (13)$$

which, for a cosmologically evolving $\dot{n}(L, t)$, can be expanded about $\dot{n}(L, t_{\rm H})$, yielding $\Delta n = \dot{n}(L, t_{\rm H})\Delta t_Q(L)$ to first order in $\Delta t_Q(L)/t_{\rm H}$. Considering a complete distribution of sources with some $L_{\rm peak}$, we similarly obtain the luminosity function

$$\phi(L) \equiv \frac{d\Phi}{d\log L}(L) = \int \frac{dt(L, L_{\rm peak})}{d\log L}\dot{n}(L_{\rm peak})\, d\log(L_{\rm peak}). \quad (14)$$

Throughout, we will denote the differential luminosity function, i.e., the comoving number density of quasars in some logarithmic luminosity interval, as $\phi \equiv d\Phi/d\log L$. Here $\dot{n}(L_{\rm peak})$ is the comoving number density of sources created per unit cosmological time per logarithmic interval in $L_{\rm peak}$, at some redshift, and $dt/d\log L$ is the differential quasar lifetime, i.e., the total time that a quasar with a given $L_{\rm peak}$ spends in a logarithmic interval in bolometric luminosity L. This formulation implicitly accounts for the "duty cycle" (the fraction of active quasars at a given time), which is proportional to the lifetime at a given luminosity. Corrections to this formula owing to finite lifetimes are of order $(dt/d\log L)/t_{\rm H}$, which for the luminosities and redshifts considered here (except for Fig. 11) are never larger than $\sim\frac{1}{5}$ and are generally $\ll 1$, which is significantly smaller than the uncertainty in the luminosity function itself.

We next consider the implications of our luminosity-dependent quasar lifetimes for the relation between the observed luminosity function and the distribution of peak luminosities (i.e., intrinsic properties of quasar systems). In traditional models of quasar lifetimes and light curves, this relation is trivial. For example, models in which quasars "turn on" at fixed luminosity for some fixed lifetime (i.e., the light-bulb model defined in § 2.5) imply

$$\dot{n}(L_{\rm peak})_{\rm LB} \propto \phi(L = L_{\rm peak}), \quad (15)$$

and models in which quasar light curves are a pure exponential growth or decay with some cutoff(s) (e.g., exponential or fixed-Eddington-ratio models) imply

$$\dot{n}(L_{\rm peak})_{\rm Edd} \propto \left.\frac{d\phi}{d\log L}\right|_{L=L_{\rm peak}}. \quad (16)$$

These both have essentially *identical* shape to the observed luminosity function, qualitatively different from our model prediction that $\dot{n}(L_{\rm peak})$ should turn over at luminosities approximately below the break in the observed luminosity function (see, e.g., Fig. 1 of Hopkins et al. 2006b). The luminosity-dependent quasar lifetimes determined from our simulations imply a new interpretation of the luminosity function, with $\dot{n}(L_{\rm peak})$ tracing the bright end of the luminosity function similar to traditional models, but then peaking and turning over below $L_{\rm peak} \sim L_{\rm break}$, the break luminosity in standard double-power-law luminosity functions. In our deconvolution of the luminosity function, the faint end corresponds primarily to sources in sub-Eddington phases transitioning into or out of the phase(s) of peak quasar activity. There is also some contribution to the faint-end lifetime from quasars accreting efficiently (i.e., growing exponentially at high Eddington ratio) early in their activity and on their way to becoming brighter sources, but this becomes an increasingly small fraction of the lifetime at lower luminosities. For example, in Figure 7 of Hopkins et al. (2005a), direct calculation of the quasar lifetime shows that sub-Eddington phases begin to dominate the lifetime for $L \lesssim 0.1 L_{\rm peak}$, with $\gtrsim$90% of the lifetime at $L \sim 10^{-3} L_{\rm peak}$ corresponding to sub-Eddington growth. By definition, a fixed-Eddington-ratio or light-bulb model is dominated at all luminosities by a fixed, usually large, Eddington ratio. Even models that assume an exponential decline in the quasar luminosity from some peak, although they clearly must spend a significant amount of time at low Eddington ratios, have an identical $\dot{n}(L_{\rm peak}) = \dot{n}(L_{\rm peak})_{\rm Edd}$ (modulo an arbitrary normalization), and predict far less time at most observable ($\gtrsim 10^{-4} L_{\rm peak}$) low luminosities and accretion rates (because the accretion rates fall off so rapidly); i.e., the population at any observed luminosity is still dominated by objects near their peak.

From our new, large set of simulations, we test this model of the relationship between the distribution of peak quasar luminosities and observed luminosity functions, namely, our assertion that $\dot{n}(L_{\rm peak})$ should peak around the observed break in the luminosity function and turn over below this peak, with the observed luminosity function faint-end slope dominated by sources with peak luminosities near the break in sub-Eddington (subpeak luminosity) states. In particular, we wish to ensure that this behavior for $\dot{n}(L_{\rm peak})$ is real, and not some artifact of our fitting functions for the quasar lifetime.

Figure 8 shows the best-fit $\dot{n}(L_{\rm peak})$ distribution (*thick histogram*) fitted to the Ueda et al. (2003) hard X-ray quasar luminosity function (*thin curve*) at redshift $z = 0.5$, as well as the resulting best-fit luminosity function (*thin histogram*). For ease of comparison with other quasar luminosities, we rescale the luminosity function to the bolometric luminosity using the corrections of Marconi et al. (2004). We determine $\dot{n}(L_{\rm peak})$ by logarithmically binning the range of $L_{\rm peak}$ and considering for each bin all simulations with $L_{\rm peak}$ in the given range. For each bin, then, we take the average binned time the simulations spend in each luminosity interval, and take that to be the quasar lifetime $dt/d\log L$. We then fit to the observed luminosity function of Ueda et al. (2003), fitting

$$\phi(L) \approx \sum_i \dot{n}_i(L_{{\rm peak},\,i}) \left\langle \frac{\Delta t(L,\ L_{{\rm peak},\,i})}{\Delta \log L} \right\rangle \qquad (17)$$

and allowing $\dot{n}_i(L_{{\rm peak},\,i})$ to be a free coefficient for each binned $L_{\rm peak} = L_{{\rm peak},\,i}$. Despite our large number of simulations, the numerical binning process makes this result noisy, especially at the extreme ends of the luminosity function. However, the relevant result is clear: the qualitative behavior of $\dot{n}(L_{\rm peak})$ described above is unchanged. For further discussion of the qualitative differences between the $\dot{n}(L_{\rm peak})$ distribution from different quasar models and the robust nature of our interpretation even under restrictive assumptions (e.g., ignoring the early phases of merger activity or applying various models for radiative efficiency as a function of accretion rate), we refer to Hopkins et al. (2005b).

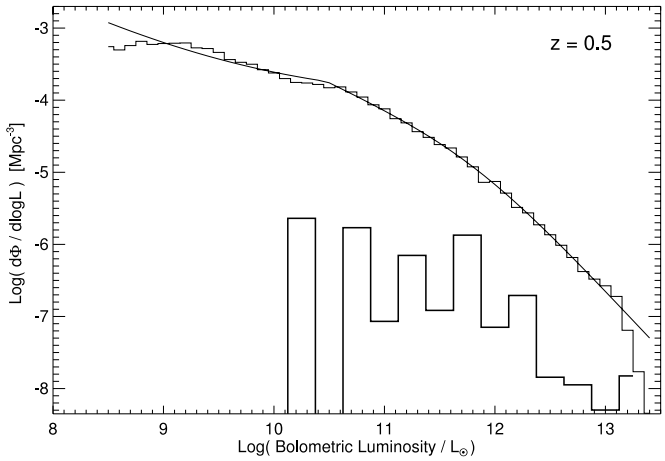

FIG. 8.—We reproduce (*thin histogram*) the luminosity function of Ueda et al. (2003) at redshift $z = 0.5$ (*thin curve*) using the binned differential quasar lifetime $dt/d\log L$ directly from our simulations and a fitted distribution of peak luminosities $\dot{n}(L_{\rm peak})$ (*thick histogram*). For each bin in $\log(L_{\rm peak})$, we average the binned differential lifetime of a set of simulations with peak luminosity in the bin. This clearly demonstrates that our key qualitative result, that the faint end of the luminosity function is reproduced by quasars with *peak* luminosity around the break luminosity but observed primarily in sub-Eddington states (luminosities $L \ll L_{\rm peak}$), is not an artifact of our fitting formulae or extrapolation to extreme luminosities.

3.2. *The Luminosity Function at Different Frequencies and Redshifts*

Given a distribution of peak luminosities $\dot{n}(L_{\rm peak})$, we can use our model of quasar lifetimes and the column density distribution as a function of instantaneous and peak luminosities to predict the luminosity function at any frequency. From a distribution of $N_{\rm H}$ values and some a priori known minimum observed luminosity $L_\nu^{\rm min}$, the fraction $f_{\rm obs}$ of quasars with a peak luminosity $L_{\rm peak}$ and instantaneous bolometric luminosity L that lie above the luminosity threshold is given by the fraction of $N_{\rm H}$ values below a critical $N_{\rm H}^{\rm max}$, where $L_\nu^{\rm min} = f_\nu L \exp(-\sigma_\nu N_{\rm H}^{\rm max})$. Here $f_\nu(L) \equiv L_\nu/L$ is a bolometric correction and σ_ν is the cross section at frequency ν. Thus,

$$N_{\rm H}^{\rm max}(\nu,\, L,\, L_\nu^{\rm min}) = \frac{1}{\sigma_\nu} \ln\left[\frac{f_\nu(L) L}{L_\nu^{\rm min}}\right], \qquad (18)$$

and for the lognormal distribution above,

$$f_{\rm obs}(\nu,\, L,\, L_{\rm peak},\, L_\nu^{\rm min}) = \frac{1}{2}\left\{1 + {\rm erf}\left[\frac{\log\left(N_{\rm H}^{\rm max}/\bar{N}_{\rm H}\right)}{\sqrt{2}\sigma_{N_{\rm H}}}\right]\right\}. \qquad (19)$$

This results in a luminosity function (in terms of the bolometric luminosity)

$$\phi(\nu,\, L,\, L_\nu^{\rm min}) = \int f_{\rm obs}(\nu,\, L,\, L_{\rm peak},\, L_\nu^{\rm min}) \times \frac{dt(L,\, L_{\rm peak})}{d\log L}\, \dot{n}(L_{\rm peak})\, d\log(L_{\rm peak}), \qquad (20)$$

where $\phi(\nu, L, L_\nu^{\rm min})$ is the number density of sources with bolometric luminosity L per logarithmic interval in L, with an observed luminosity at frequency ν above $L_\nu^{\rm min}$.

Based on the direct fit for $\dot{n}(L_{\rm peak})$ in Figure 8, we wish to consider a functional form for $\dot{n}(L_{\rm peak})$ with a well-defined peak and falloff in either direction in $\log(L_{\rm peak})$. Therefore, we take $\dot{n}(L_{\rm peak})$ to be a lognormal distribution, with

$$\dot{n}(L_{\rm peak}) = \dot{n}_* \frac{1}{\sigma_* \sqrt{2\pi}} \exp\left\{ -\frac{1}{2} \left[\frac{\log(L_{\rm peak}/L_*)}{\sigma_*} \right]^2 \right\}. \quad (21)$$

Here $\dot{n}_*$ is the total number of quasars being created or activated per unit comoving volume per unit time; L_* is the center of the lognormal, the characteristic peak luminosity of quasars being born [i.e., the peak luminosity at which $\dot{n}(L_{\rm peak})$ itself peaks], which is directly related to the break luminosity in the observed luminosity function; and σ_* is the width of the lognormal in $\dot{n}(L_{\rm peak})$ and determines the slope of the bright end of the luminosity function. Since our model predicts that the bright end of the luminosity function is made up primarily of sources at high Eddington ratio near their peak luminosity, i.e., essentially identical to light-bulb or fixed-Eddington-ratio models, the bright-end slope is a fitted quantity, determined by whatever physical processes regulate the bright-end slope of the active black hole mass function (possibly feedback from outflows or threshold cooling processes; e.g., Wyithe & Loeb 2003; Scannapieco & Oh 2004; Dekel & Birnboim 2006), unlike the faint-end slope, which is a consequence of the quasar lifetime itself and is only weakly dependent on the underlying faint-end active black hole mass or $\dot{n}(L_{\rm peak})$ distribution.

We note that although this choice of fitting function has appropriate general qualities, it is ultimately somewhat arbitrary, and we choose it primarily for its simplicity and its capacity to match the data with a minimum of free parameters. We could instead, for example, have chosen a double-power-law form with $\dot{n}(L_{\rm peak}) = \dot{n}_* / [(L_{\rm peak}/L_*)^{\gamma_1} + (L_{\rm peak}/L_*)^{\gamma_2}]$ and $\gamma_1 < \gamma_2$, but given that the entire faint end of the luminosity function is dominated by objects with $L_{\rm peak} \sim L_*$, the observed luminosity function has essentially no power to constrain the faint-end slope γ_1, other than setting an upper limit $\gamma_1 \lesssim 0$. The "true" $\dot{n}(L_{\rm peak})$ will, of course, be a complicated function of both halo merger rates at a given redshift and the distribution of host galaxy properties including, but not necessarily limited to, masses, concentrations, and gas fractions.

Having chosen a form for $\dot{n}(L_{\rm peak})$, we can then fit to an observed luminosity function to determine ($\dot{n}_*$, L_*, σ_*). We take advantage of the capability of our model to predict the luminosity function at multiple frequencies and consider both fits to just the Ueda et al. (2003) hard X-ray (2–10 keV) luminosity function, $\phi_{\rm HX}$, and fits to the Ueda et al. (2003) and Miyaji et al. (2001) soft X-ray (0.5–2 keV; $\phi_{\rm SX}$) and Croom et al. (2004) optical B-band (4400 Å; ϕ_B) luminosity functions *simultaneously*. These observations agree with other, more recent determinations of $\phi_{\rm HX}$, $\phi_{\rm SX}$, and ϕ_B (e.g., Barger et al. 2005; Hasinger et al. 2005; Richards et al. 2005, respectively) at most luminosities, and therefore we do not expect revisions to the observed luminosity functions to dramatically change our results. In order to avoid numerical artifacts from fitting to extrapolated, low-luminosity slopes in the analytical forms of these luminosity functions, we directly fit to the binned luminosity function data. Thus, we fit each luminosity function in all redshift intervals for which we have binned data.

We find good fits ($\chi^2/\nu = 68.8/104 \approx 0.66$) to all luminosity functions at all redshifts with a pure peak-luminosity evolution (PPLE) model, for which

$$L_* = L_*^0 \exp(k_L \tau), \quad \dot{n}_* = \text{constant}, \quad \sigma_* = \text{constant}, \quad (22)$$

where τ is the fractional look-back time ($\tau \equiv H_0 \int_0^z dt$) and k_L is a dimensionless constant fitted with L_*, $\dot{n}_*$, σ_*. It is important to distinguish this pure *peak*-luminosity evolution from "standard" pure luminosity evolution (PLE) models (e.g., Boyle et al. 1988), as the shape (in addition to just the horizontal normalization) of the quasar luminosity function can change as a function of redshift if the quasar lifetime is not a completely self-similar function of redshift or luminosity.

We do not find significant improvement in the fits if we additionally allow $\dot{n}_*$ or σ_* to evolve with redshift ($\Delta\chi^2 \sim 1$–2, depending on the adopted form for the evolution), and therefore consider only the simplest parameterization above (eq. [22]). We also find acceptable fits for a pure density evolution (PDE) model, with L_* = constant and $\dot{n}_* = \dot{n}_*^0 \exp(k_N \tau)$ (both keeping σ_* fixed and allowing it to evolve as well). However, the fits are somewhat poorer ($\chi^2/\nu \approx 1$), and the resulting parameters overproduce the present-day density of low-mass supermassive black holes and the intensity of the X-ray background by an order of magnitude, so we do not consider them further. In either case, there is a considerable degeneracy between the parameters σ_* and L_*, where a decrease in L_* can be compensated by a corresponding increase in σ_*. This degeneracy is present because, as indicated above, the observed luminosity function only weakly constrains the faint-end slope of $\dot{n}(L_{\rm peak})$.

The observations shown are insufficient at high redshift to strongly resolve the "turnover" in the total comoving quasar density at $z \sim 2$–3, and thus we acknowledge that there must be corrections to this fitted evolution at higher redshift, which we address below. However, as we primarily consider low redshifts, $z \lesssim 3$, and show that the supermassive black hole population and X-ray background are dominated by quasars at redshifts for which our $\dot{n}(L_{\rm peak})$ distribution is well determined, this is not a significant source of error in most of our calculations even if we extrapolate our evolution to $z \gg 3$.

Figure 9 shows the resulting best-fit PPLE luminosity functions from the best-fit $\dot{n}(L_{\rm peak})$ distribution, for redshifts $z = 0$–3. This has the best-fit ($\chi^2/\nu = 0.67$) values ($\log L_*, k_L, \log \dot{n}_*, \sigma_*$) = (9.94, 5.61, −6.29, 0.91) with corresponding errors (0.29, 0.28, 0.13, 0.09). Here L_* is in solar luminosities and $\dot{n}_*$ in comoving Mpc^{-3} Myr^{-1}. Fitting to the hard X-ray data alone gives a similar fit, with the slightly different values ($\log L_*, k_L, \log \dot{n}_*, \sigma_*$) = (9.54, 4.90, −5.86, 1.03) ± (0.66, 0.43, 0.37, 0.13), $\chi^2/\nu = 0.7$ (note the degeneracy between L_* and σ_* in the two fits). Our best-fit value of $k_L = 5.6$ compares favorably to the value of ∼6 found by, e.g., Boyle et al. (2000) and Croom et al. (2004) for the evolution of the break luminosity in the observed luminosity function, demonstrating that the break luminosity traces the *peak* in the $\dot{n}(L_{\rm peak})$ distribution at all redshifts. These fits and the errors were obtained by least-squares minimization over all data points (comparing each to the predicted curve at its redshift and luminosity), assuming the functional form we have adopted for $\dot{n}(L_{\rm peak})$.

The agreement we obtain at all redshifts, in the hard X-ray (*solid line*), soft X-ray (*dashed line*), and B band (*dotted line*), is good. This is not at all guaranteed by our procedure, as the fit is highly overconstrained because we fit three luminosity functions each at five redshifts to only four free parameters. Of course, the choice of the functional form for $\dot{n}(L_{\rm peak})$ ensures that we should

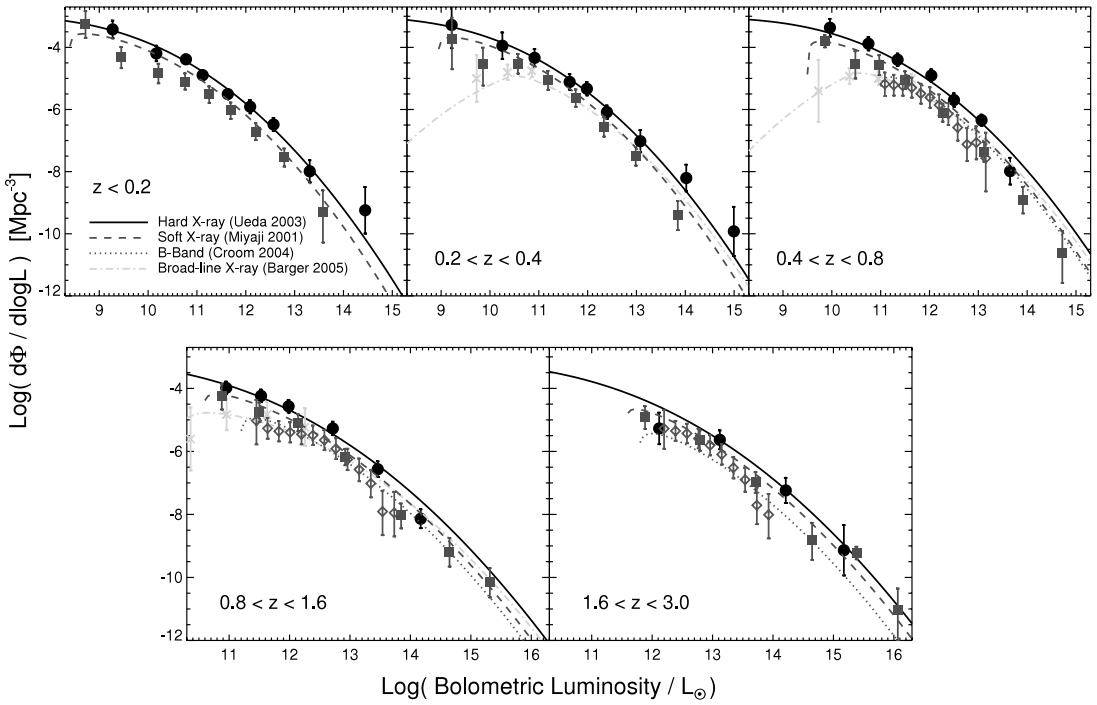

FIG. 9.—Best-fit luminosity function from the PPLE $\dot{n}(L_{\rm peak})$ distribution, for redshifts $z = 0-3$. From our fitted lognormal $\dot{n}(L_{\rm peak})$ distribution, we simultaneously reproduce the luminosity function in the hard X-ray (2–10 keV; *solid line*), soft X-ray (0.5–2 keV; *dashed line*), and optical *B* band (4400 Å; *dotted line*) at all redshifts. Moreover, we reproduce the distribution of broad-line quasars in hard X-ray selected samples (*dot-dashed line*), as described in § 4. All quantities have been rescaled to bolometric luminosities for ease of comparison, using the corrections of Marconi et al. (2004), with the plotted error bars representing both quoted measurement errors and the estimated errors in the bolometric corrections. The observations are from Miyaji et al. (2001) (soft X-ray; *squares*), Ueda et al. (2003) (hard X-ray; *circles*), Croom et al. (2004) (*B* band; *diamonds*), and Barger et al. (2005) (X-ray-selected broad-line quasars; *crosses*). [*See the electronic edition of the Supplement for a color version of this figure.*]

be able to reproduce at least one luminosity function and its evolution (e.g., the hard X-ray luminosity function, which is least affected by attenuation), but our modeling of the column density distributions in mergers allows us to simultaneously reproduce the luminosity functions in different wave bands without imposing assumptions about obscured fractions or sources of attenuation. Expressed as bolometric luminosity functions, ϕ_B, $\phi_{\rm SX}$, and $\phi_{\rm HX}$ would be identical in the absence of obscuration, similar to the predicted $\phi_{\rm HX}$, as obscuration is minimal in the hard X-ray.

For redshifts $z \leq 1$, in our Figure 10 we reproduce Figure 2 of Hopkins et al. (2005c), which shows in detail the agreement between hard X-ray (Ueda et al. 2003), soft X-ray (Miyaji et al. 2000), and optical (Boyle et al. 2000) luminosity functions resulting from the time- and luminosity-dependent column density distributions derived from the simulations. The differential extinction predicted for different frequencies (and magnitude limits) of observed samples based on the column density distributions in our simulations accounts for the different shape of the luminosity function in each band, and the evolution of the luminosity function with redshift is driven by a changing L_*, the peak of the $\dot{n}(L_{\rm peak})$ distribution (eq. [22]). We emphasize that in our analysis, the key quantity constrained by observations is the fitted $\dot{n}(L_{\rm peak})$ distribution with redshift. All other quantities and distributions are derived from the basic input physics of our simulations, with no further assumptions or adjustable factors in our modeling beyond the prescription for Bondi (Eddington limited) accretion and ~5% energy deposition in the ISM, which are themselves constrained by observations and theory as discussed in § 2 and in Di Matteo et al. (2005).

We can, of course, fit the previously defined simpler model of quasar lifetimes, either a light-bulb or exponential light curve/fixed Eddington ratio model, and obtain an identical hard X-ray luminosity function. We determine these fits (see also eqs. [15] and [16]) and use them throughout when we compare the predictions of such models (described in § 2.5) to those of our simulated quasar lifetimes in our subsequent analysis. Applying a standard torus model to any model of the luminosity function reproduces, by design, the mean offset between the *B*-band and hard X-ray luminosity functions, as the parameters of this model are *tuned* to reproduce this offset. As many observations show, the fraction of broad-line quasars increases with luminosity (Steffen et al. 2003; Ueda et al. 2003; Hasinger 2004; Sazonov & Revnivtsev 2004; Barger et al. 2005; Simpson 2005), and so reproducing the relationship between *B*-band and hard X-ray luminosity functions requires adding parameters to the standard torus model that allow luminosity-dependent scalings, i.e., the class of "receding torus" models. These, again by construction, reproduce the distinction between hard X-ray and *B*-band quasar luminosity functions, including the dependence of this difference on luminosity. These are, however, phenomenological models designed to fit these observations. Our simulations, on the other hand, provide a self-consistent description of the column density, which predicts the *differences* between hard X-ray, soft X-ray, and optical luminosity functions without the addition of tunable parameters or model features designed to reproduce these observations.

Our fits are accurate down to low luminosities, as is clear from our prediction for the X-ray luminosity function at bolometric luminosities $L \sim 10^9\ L_\odot$. Furthermore, we have calculated the predicted $z \lesssim 0.1$ luminosity function in the *B* band as well as in Hα emission, using the conversion between the two from Hao et al. (2005) and comparing directly to their luminosity functions for Seyfert galaxies and low-luminosity active galactic nuclei (AGNs) (both type I and II), and find that our distribution $\dot{n}(L_{\rm peak})$

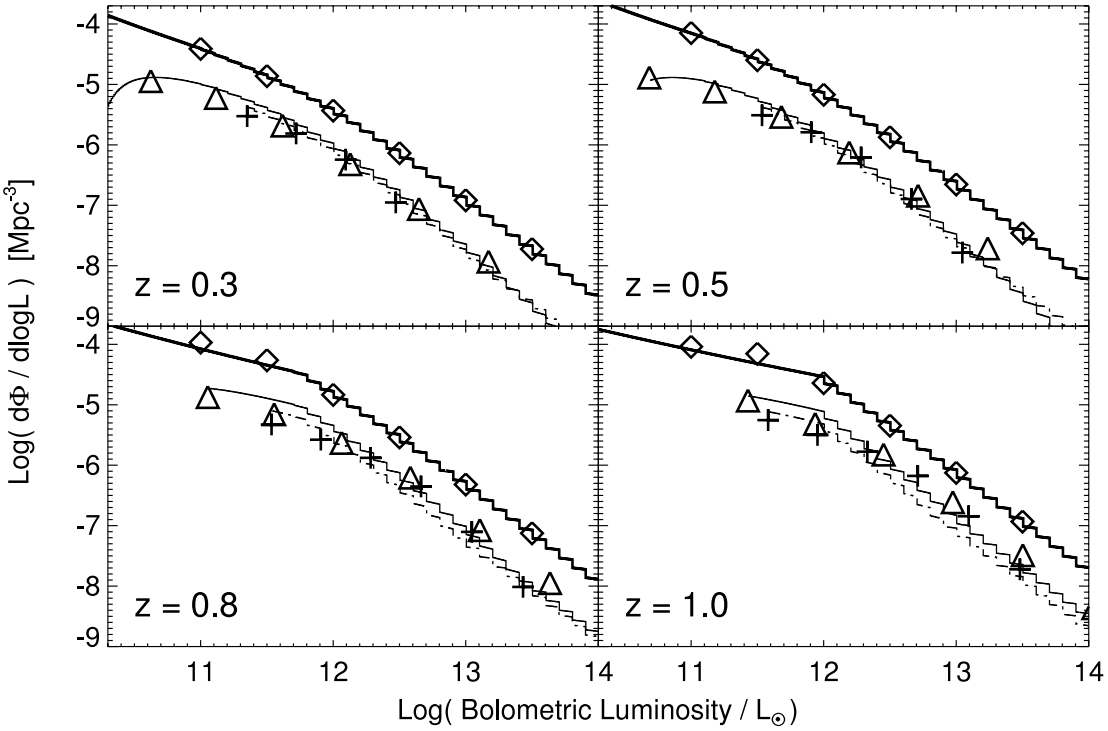

FIG. 10.—Hard X-ray (*thick line*), soft X-ray (*thin line*), and *B*-band (*dot-dashed line*) luminosity functions determined from our model of quasar lifetimes and column densities, based on a distribution of intrinsic source properties fitted to the observed hard X-ray luminosity function and the limiting magnitudes of observed samples, at the different redshifts shown. All quantities are rescaled to bolometric luminosities with the bolometric corrections of Marconi et al. (2004). Symbols show the observed luminosity functions for hard X-rays (Ueda et al. 2003; *diamonds*), soft X-rays (Miyaji et al. 2000; *triangles*), and *B* band (Boyle et al. 2000; *crosses*). Reproduced from Hopkins et al. (2005c).

and model for quasar lifetimes and obscuration reproduces the complete observed luminosity function down to a *B*-band luminosity $M_B \sim -16$. Although our prediction falls below the observed Seyfert luminosity function at fainter magnitudes, there is no reason to believe that mergers should be responsible for all nuclear activity at these luminosities (and indeed alternative fueling mechanisms for such faint objects likely exist); it is surprising, in fact, that this picture reproduces the observed AGN activity to such faint luminosities.

Using the bolometric corrections of Elvis et al. (1994) instead of Marconi et al. (2004) results in a significantly steeper cutoff in the luminosity function at high bolometric luminosities, as the bolometric luminosity inferred for the brightest observed X-ray quasars is almost an order of magnitude smaller using the Elvis et al. (1994) corrections. However, this is because the Elvis et al. (1994) bolometric corrections do not account for any dependence on luminosity, and further, the quasars in the sample of Elvis et al. (1994) are X-ray bright (Elvis et al. 2002), whereas it has been well-established that the ratio of bolometric luminosity to hard or soft X-ray luminosity increases with increasing luminosity (e.g., Wilkes et al. 1994; Green et al. 1995; Vignali et al. 2003; Strateva et al. 2005). Recent comparisons between large samples of quasars selected by both optical and X-ray surveys (Risaliti & Elvis 2005) further suggest that this is an intrinsic correlation, not driven by, for example, the dependence of obscuration on luminosity. For a direct comparison of the bolometric luminosity functions resulting from the two corrections, we refer to Hopkins et al. (2005c). Our analysis uses the form for the UV–to–X-ray flux ratio, α_{ox}, from Vignali et al. (2003), but our results are relatively insensitive to the different values found in the literature. It is important to account for this dependence, as it creates a significant difference in the high-luminosity end of the bolometric quasar luminosity function and implies that a nonnegligible fraction of the brightest quasars are not seen in optical surveys (see the discussion in Marconi et al. 2004; Richards et al. 2005).

Finally, our fitted form for the evolution of the break luminosity, with $L_* \propto \exp(k_L \tau)$, cannot continue to arbitrarily high redshift. At redshifts $z \gtrsim 2-3$, this asymptotes because $\tau \rightarrow 1$, whereas the observed quasar population declines above $z \sim 2$. This difference is not important for most of our calculated observables, as they are either independent of high-redshift evolution or evolve with cosmic time in some fashion as $\propto \int \dot{n}(L_{\rm peak})\, dt$, with little time and thus negligible contributions to integrated totals at high redshifts. However, some quantities, in particular the high-mass end of the black hole mass function (see § 6), which is dominated by the small number of the brightest quasars at high redshifts, can receive large relative contributions from these terms. Therefore, it is important in estimating these quantities to be aware of the turnover in the quasar density at high redshifts.

We quantify this in Figure 11, where we show the predicted broad-line luminosity function (where the broad-line phase is determined below in § 6) in six luminosity intervals from $z \sim 1.2$ to 4.8. The intervals are those of the COMBO-17 (Classifying Objects by Medium-Band Observations in 17 Filters) luminosity function from Wolf et al. (2003), but we further compare to the observed luminosity functions of Warren et al. (1994), Schmidt et al. (1995), Kennefick et al. (1995), Fan et al. (2001), and Richards et al. (2005) at the appropriate (labeled) redshifts. At each redshift $z > 2$, we take the fitted $\dot{n}(L_{\rm peak})$ distribution above (eqs. [21] and [22]) and rescale it according to an exponential cutoff: either PDE, $\dot{n}(L_{\rm peak}) \rightarrow \dot{n}(L_{\rm peak}) \times 10^{-\alpha_{\rm PDE}(z-2)}$, or PPLE, $L_* \rightarrow L_* \times 10^{-\alpha_{\rm PPLE}(z-2)}$. Fitting to the data gives $\alpha_{\rm PDE} \sim 0.65$ and $\alpha_{\rm PPLE} \sim 0.55$ ($\chi^2/\nu \approx 1.3$ for both), in reasonable agreement

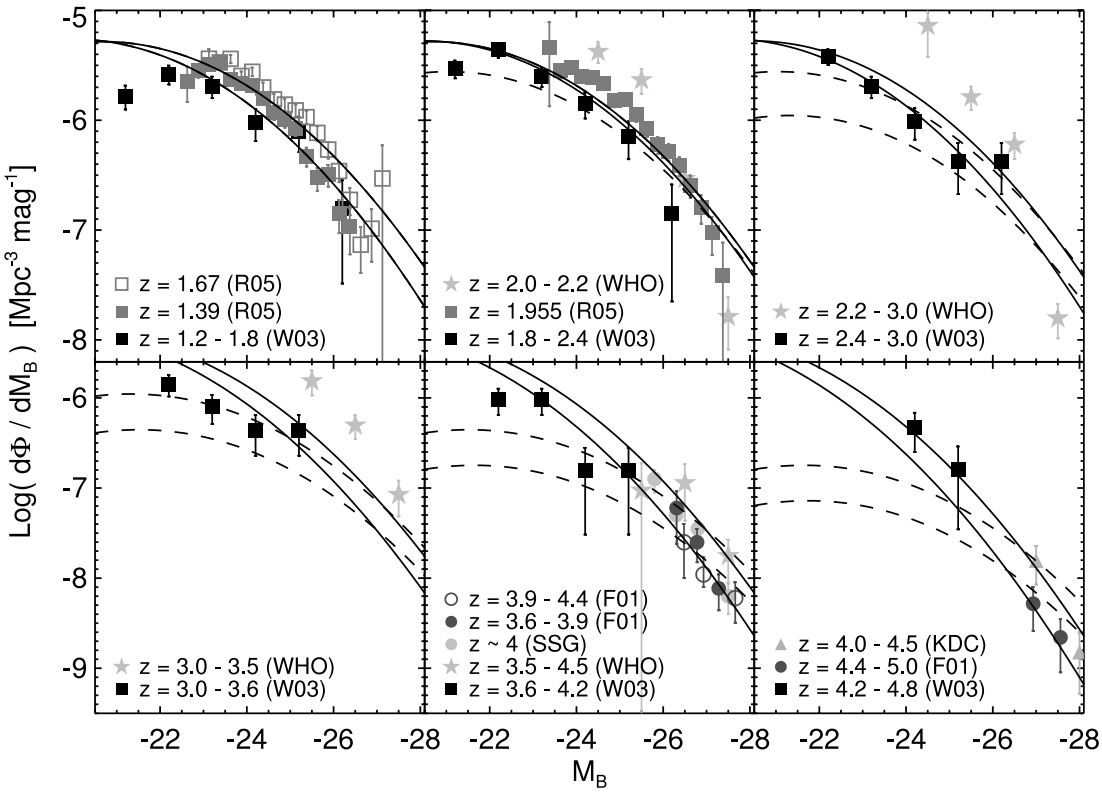

FIG. 11.—Running our predicted broad-line luminosity function (determined in §§ 3 and 4) to high redshifts, with either total density (*dashed lines*) or break luminosity (L_*; *solid lines*) decreasing exponentially with redshift above $z = 2$. In each panel, our prediction is shown for the minimum and maximum redshift of the corresponding interval from the COMBO-17 luminosity function of Wolf et al. (2003) (W03; *squares*). Other references for the observations shown are Richards et al. (2005; R05), Warren et al. (1994; WHO), Fan et al. (2001; F01), Schmidt et al. (1995; SSG), and Kennefick et al. (1995; KDC). [*See the electronic edition of the Supplement for a color version of this figure.*]

with the density evolution of, e.g., Fan et al. (2001). We note that this evolution, extrapolated as far as $z \sim 6$, is consistent also with the constraints on $z \sim 6$ quasars from Fan et al. (2004), especially in the PPLE case.

In each panel, we plot the resulting broad-line luminosity function (see § 4), for both the minimum and maximum redshift of the redshift bin, and both the PPLE (*solid lines*) and PDE (*dashed lines*) cases. The degeneracy between these possibilities is well known, as current observations do not resolve the break in the luminosity function. Furthermore, the predicted luminosity function should be considered uncertain, especially at low luminosities, as the quasar lifetime at these luminosities and redshifts can become comparable to the age of the universe, at which point our formalism for the luminosity function as a function of $\dot{n}(L_{\rm peak})$ becomes inaccurate. However, we are able to make testable predictions, based on differences between the two models in integrated *galaxy* properties (for example, color-magnitude diagrams of red sequence galaxies at low masses or the fraction of recently formed spheroids as a function of mass and redshift), which distinguish the PPLE and PDE models for the evolution of the quasar luminosity function at $z \gtrsim 2-3$ (Hopkins et al. 2006b). Owing to these degeneracies and the poor constraints on the observed high-redshift luminosity functions, we have not considered them (those at $z > 3$) in our fits to $\dot{n}(L_{\rm peak})$, but use them here to roughly constrain the turnover in the quasar density above $z \sim 2$ (i.e., fitting to $\alpha_{\rm PDE}$ and $\alpha_{\rm PPLE}$). Which form of the turnover we use makes little difference in our subsequent analysis, but, as discussed above, including *some* turnover is important in calculating select quantities such as the extreme high-mass end of the black hole mass function.

3.3. *The Observed $N_{\rm H}$ Distribution*

Given the column density distributions and quasar lifetimes calculated from our simulations in § 2, and the quantity $\dot{n}(L_{\rm peak})$ determined above (§ 3.2), we can predict the distribution of column densities observed in a given sample. This will depend not only on the range of observed luminosities and the redshift of the sample, but also on the minimum observed magnitude and frequency (i.e., the selection function) of the sample. For a nearly complete sample or estimate of the luminosity function, for example the hard X-ray luminosity function, at least to $N_{\rm H} \sim 10^{25}$ cm^{-2}, we can integrate the $N_{\rm H}(L, L_{\rm peak})$ distribution over the $\dot{n}(L_{\rm peak})$ distribution (weighted by the lifetime at L).

Figure 12 plots the resulting distribution of column densities for this analysis. The left panel reproduces and expands on a portion of Figure 3 of Hopkins et al. (2005a), showing the distribution of column densities (scaled linearly) expected from the characteristic quasars $L_{\rm peak} \sim L_*$ of the luminosity function observed in optical samples, based on the simulated column density distributions as a function of luminosity and peak luminosity (*solid line*). Specifically, we plot the distribution of neutral $N_{\rm H\,I}$ values requiring that the observed B-band luminosity be above some reference value $L_{B,\,\rm min}$. The smooth curve shown is the best fit to the E_{B-V} distribution of bright SDSS (Sloan Digital Sky Survey) quasars with $z < 2.2$ from Hopkins et al. (2004). The curve has been rescaled in terms of the column density (inverting our gas-to-dust prescription) and plotted about a peak (mode) $N_{\rm H\,I}$ (undetermined in Hopkins et al. 2004) of $N_{\rm H\,I} \approx 0.5 \times 10^{21}$ cm^{-2}. The observationally implied E_{B-V} distribution is determined from fitting to the distribution of photometric reddening in all SDSS

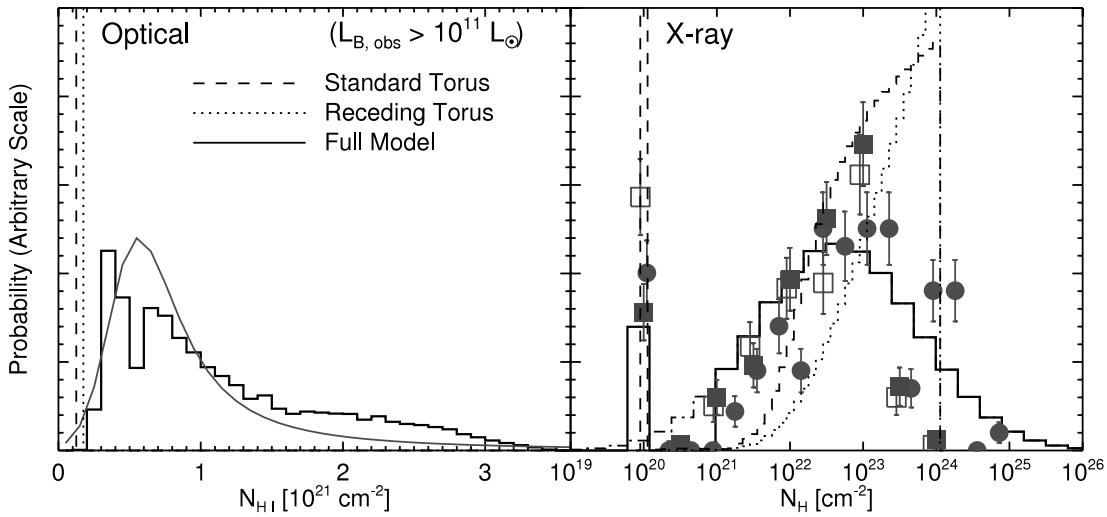

FIG. 12.—*Left*: Distribution of column densities expected from the characteristic quasars $L_{peak} \sim L_*$ of the luminosity function observed in optical samples, for a standard torus model of quasar obscuration (*dashed line*), a receding torus model (*dotted line*), and the distributions of column densities as a function of instantaneous and peak luminosity in our simulations (*solid line*). The distribution of neutral $N_{H\,I}$ values is obtained requiring an observed *B*-band luminosity $>10^{11}\ L_\odot$. The curve is the best fit to the E_{B-V} distribution of bright SDSS quasars with $z < 2.2$, from Hopkins et al. (2004), rescaled to column densities and plotted about a peak (mode) $N_{H\,I}$ (undetermined in Hopkins et al. 2004) of $N_{H\,I} \approx 0.5 \times 10^{21}$ cm^{-2}. The *i*-band absolute magnitude limit imposed in the observed sample, $M_i < -22$, corresponds approximately to our plotted *B*-band limit $L_{B,\,obs} > 10^{11}\ L_\odot$. Reproduced from Hopkins et al. (2005a). *Right*: Integrated distribution of total (neutral and ionized) column densities expected for a complete hard X-ray sample, from the column densities of our simulations and the $i(L_{peak})$ distribution. The distribution below 10^{21} cm^{-2} is shown (*dot-dashed line*) and replotted as a single bin at $N_H = 10^{20}$ cm^{-2} for our modeled columns. Data shown are the results of Treister et al. (2004; *squares*) and Mainieri et al. (2005; *circles*), with assumed Poisson errors. Filled squares assume an intrinsic photon index $\Gamma = 1.9$ for the soft X-ray quasar spectrum; open squares assume $\Gamma = 1.7$. [*See the electronic edition of the Supplement for a color version of this figure.*]

bands (i.e., using the five-band photometry as a proxy for spectral fitting) in Sloan quasars, relative to the modal quasar colors at each redshift, for quasars with an absolute magnitude limit $M_i < -22$. The *i*-band absolute magnitude limit imposed in the observed sample, $M_i < -22$, corresponds approximately to our plotted *B*-band limit $L_{B,obs} > 10^{11}\ L_\odot$. This estimate does not account for bright but strongly reddened quasars having their colors altered to the point where color selection criteria of quasar surveys will not include them. However, this effect would only serve to bring our distribution into better agreement with observations, as it would slightly lower the high-$N_{H\,I}$ tail. We also consider the predictions of a standard torus model and receding (luminosity dependent) torus model in the figure (*dashed and dotted lines, respectively*). These should not be taken literally in this case; they reflect that these phenomenological models do not predict the distribution of low/moderate column densities, but rather assume that all lines of sight not intersecting the torus are "unobscured" and encounter some constant, small column density (usually chosen to be $N_H \sim 10^{20}$ cm^{-2}).

The right panel of Figure 12 shows the integrated distribution (in log N_H) for a complete hard X-ray sample, both as predicted from our simulations based on the joint distribution of column density, luminosity, and peak luminosity (*solid line*), and for both the standard torus model (*dashed line*) and receding torus model (*dotted line*) described in § 2.5. The data shown are the results of Treister et al. (2004; *squares*) and Mainieri et al. (2005; *circles*), with assumed Poisson errors, from multiband *Chandra* and *Hubble Space Telescope* observations of GOODS (Great Observatories Origins Deep Survey) fields. The filled squares are obtained by assuming an intrinsic photon index for the soft X-ray quasar spectrum of $\Gamma = 1.9$, the open squares assuming $\Gamma = 1.7$. For the sake of direct comparison with observed distributions, objects with $N_H < 10^{21}$ cm^{-2}, for which only an upper limit to the column density would be determined in X-ray observations, are grouped together and plotted as a single bin at $N_H = 10^{20}$ cm^{-2}. The actual distribution below 10^{21} cm^{-2} is shown as a dot-dashed line. We note that our model of the quasar spectrum assumes a photon index $\Gamma = 1.9$ in the soft X-ray, but this has no effect on the column densities calculated from the surrounding gas in our simulations.

The agreement between the observed column density distribution and the result of our simulations once the same selection effect is applied supports our model of quasar evolution, and the good agreement extends to both optical and X-ray samples. Probing to fainter luminosities or frequencies less affected by attenuation broadens the column density distribution, as is seen from the inferred column density distributions in the X-ray. This broadening occurs because, at lower luminosities, observers will see both intrinsically bright periods extinguished by larger column densities (broadening the distribution to larger N_H values) and intrinsically faint periods with small column densities (broadening the distribution to smaller N_H values). The distribution as a function of reference luminosity is a natural consequence of the dynamics of the quasar activity. Throughout much of the duration of bright quasar activity, column densities rise to high levels as a result of the same process that feeds accretion, producing the well-known reddened population of quasars (e.g., Webster et al. 1995; Brotherton et al. 2001; Francis et al. 2001; Richards et al. 2001, 2003; Gregg et al. 2002; White et al. 2003), extending to bright quasars strongly reddened by large $N_{H\,I}$. Furthermore, a significant number of quasars are extinguished from optical samples or attenuated to lower luminosities, giving rise to the distinction between luminosity functions in the hard X-ray, soft X-ray, and optical.

The standard torus model described in § 2.5, although unable to predict the distribution of column densities seen in optically, relatively unobscured quasars, does a fair job of reproducing the observed distribution of X-ray column densities. The parameters of the model are, of course, chosen to reproduce the data shown (the model parameters are taken from Treister et al. 2004).

Nevertheless, our prediction is still a better fit to the observed distribution, with $\chi^2/\nu \approx 2$ as opposed to $\chi^2/\nu \approx 7$ (although the absolute values depend on the estimated systematic errors in the column density estimations). The receding torus model fares even more poorly in reproducing the observed column density distributions and is ruled out at high significance ($\chi^2/\nu \approx 10$), although this can be alleviated if the observed samples are assumed to be incomplete above $N_{\rm H} \sim 10^{23}$ cm^{-2}. This disagreement results because, in order to match the observed scaling of broad-line fraction with luminosity (see § 4 below), this model assumes a larger covering fraction for the torus at lower luminosities, normalized to an obscured fraction similar to that of the standard torus model near the break in the observed quasar luminosity function. However, since quasars with luminosities below the break dominate the total number counts, this predicts that the cumulative column density distribution must be significantly more dominated by objects with large covering angles, giving a larger Compton-thick population, inconsistent with the actual observed column density distribution.

Although we do not see a significant fraction of extremely Compton-thick column densities $N_{\rm H} \gtrsim 10^{26}$ cm^{-2} in the distributions from our simulations, our model does not rule out such values. It is possible that bright quasars in unusually massive galaxies or quasars in higher redshift, compact galaxies that we have not simulated may, during peak accretion periods, reach such values in their typical column densities. Moreover, as our model assumes that ∼90% of the mass of the densest gas is clumped into cold-phase molecular clouds, a small fraction of sight lines will pass through such clouds and measure column densities similar to those shown for the "cold-phase gas" in, for example, Figure 2 of Hopkins et al. (2005d), $N_{\rm H} \gtrsim 10^{25}-10^{26}$ cm^{-2}.

Furthermore, we have not determined the "shape" at any instant of the obscuration (e.g., the dependence of obscuration on radial direction), as in practice, for most of the most strongly obscured phases in peak merger activity, the central regions of the merging galaxies are highly chaotic. Generally, the scale of the obscuration in the peak merger phases is ∼100 pc, quite different from that implied by most traditional molecular torus models, but we note that our resolution limits, ∼20 pc in the dense central regions of the merger, prevent our ruling out collapse of gas in the central regions into a smaller but more dense torus. However, several efforts to model traditional tori through radiative transfer simulations (e.g., Granato & Danese 1994; Schartmann et al. 2005) suggest significant column densities produced on scales of ∼100–200 pc, comparable to our predictions, and we note that only the solid angle covered by a torus, not the absolute torus scale, is constrained in the typical phenomenological torus model (e.g., Antonucci 1993).

Whether the obscuration of bright quasars originates on larger scales than is generally assumed is observationally testable, either through direct probes of polarized scattered light tracing the obscuring/reflecting structure (e.g., Zakamska et al. 2005) or through correlations between obscuration and, for example, host galaxy morphologies and inclinations (e.g., Donley et al. 2005). These larger scales typical of the central regions of a galaxy are widely accepted as the scales of obscuration in starbursting systems (e.g., Soifer et al. 1984a, 1984b; Sanders et al. 1986, 1988a, 1988b; for a review, see, e.g., Soifer et al. 1987), which in our modeling is associated with rapid obscured quasar growth and precedes the quasar phase. Thus, it is natural to associate obscuration with these large scales in any picture that associates starbursts and rapid black hole growth or quasar activity, as opposed to the smaller, approximately parsec scales implied by torus models primarily developed to reproduce observations of quiescent, low-luminosity type II AGNs, which are usually not directly associated with merger activity. These low-luminosity AGNs are in a relaxed state, suggesting the possibility that the remaining cold gas in the central regions of our merger remnants will collapse once the violent effects of the merger and bright quasar phase have passed, producing a more traditional small torus in a quiescent nucleus. The central point is that regardless of the form of obscuration, the typical magnitude of the obscuration is a strongly evolving function of time, luminosity, and host system properties, and the observed column density distributions reflect this evolution.

4. BROAD-LINE QUASARS

4.1. *Determining the Broad-Line Phase*

Optical samples typically identify quasars through their colors, relying on the characteristic nonstellar power-law continua of such objects. However, observations of X-ray–selected AGNs show a large population of so-called type 2 AGNs, most of which have Seyfert-like luminosities and typical spectra in X-rays and wavelengths longward of 1 μm (e.g., Elvis et al. 1994) but are optically obscured to the point where no broad lines are visible. Their optical continua, in other words, resemble those of typical galaxies, and thus they are not identified by conventional color selection techniques in optical quasar surveys. Traditional unification models (Antonucci 1993) have postulated a static torus as the explanation for the existence of the type 2 population, with such objects viewed through the dusty torus and thus optically obscured. Moreover, both synthesis models of the X-ray background (Setti & Woltjer 1989; Madau et al. 1994; Comastri et al. 1995; Gilli et al. 1999, 2001) and recent direct observations in large surveys (e.g., Zakamska et al. 2004, 2005) indicate the existence of a population of type 2 quasars, with similar obscuration but intrinsic (unobscured) quasar-like luminosities.

Observations of both radio-loud (Hill et al. 1996; Simpson et al. 1999; Willott et al. 2000; Simpson & Rawlings 2000; Grimes et al. 2004) and radio-quiet (Steffen et al. 2003; Ueda et al. 2003; Hasinger 2004; Sazonov & Revnivtsev 2004; Barger et al. 2005; Simpson 2005) quasars, however, have shown that the broad-line fraction increases with luminosity, with broad-line objects representing a large fraction of all AGNs at luminosities above the "break" in the luminosity function and rapidly falling off at luminosities below the break. Modifications to the standard torus unification model explain this via a luminosity-dependent inner torus radius (Lawrence 1991), but this represents a tunable modification to a purely phenomenological model. Furthermore, as the observations have improved, it has become clear that even these luminosity-dependent torus models cannot produce acceptable fits to the broad-line fraction as a function of luminosity (e.g., Simpson 2005). However, we have shown above that the obscuring column, even at a given luminosity, is an evolutionary effect, dominated by different stages of gas inflow in different merging systems giving rise to varying typical column densities, rather than a single static structure. It is of interest, then, to calculate when quasars will be observed as broad-line objects, and to compare this with observations of broad-line quasars and their population as a function of luminosity.

Figure 13 shows the *B*-band luminosity as a function of time for both the quasars and host galaxies in three representative simulations: the A2, A3, and A5 cases described in detail in § 2.1. These simulations each have $f_{\rm gas} = 1.0$, $q_{\rm EOS} = 1.0$, and $z_{\rm gal} = 0$, with virial velocities $V_{\rm vir} = 113, 160$, and 320 km s^{-1} and with resulting final black hole masses $M_{\rm BH}^{f} = 3 \times 10^7$, 3×10^8, and 2×10^9 $M_\odot$, respectively. The thick dark line in each

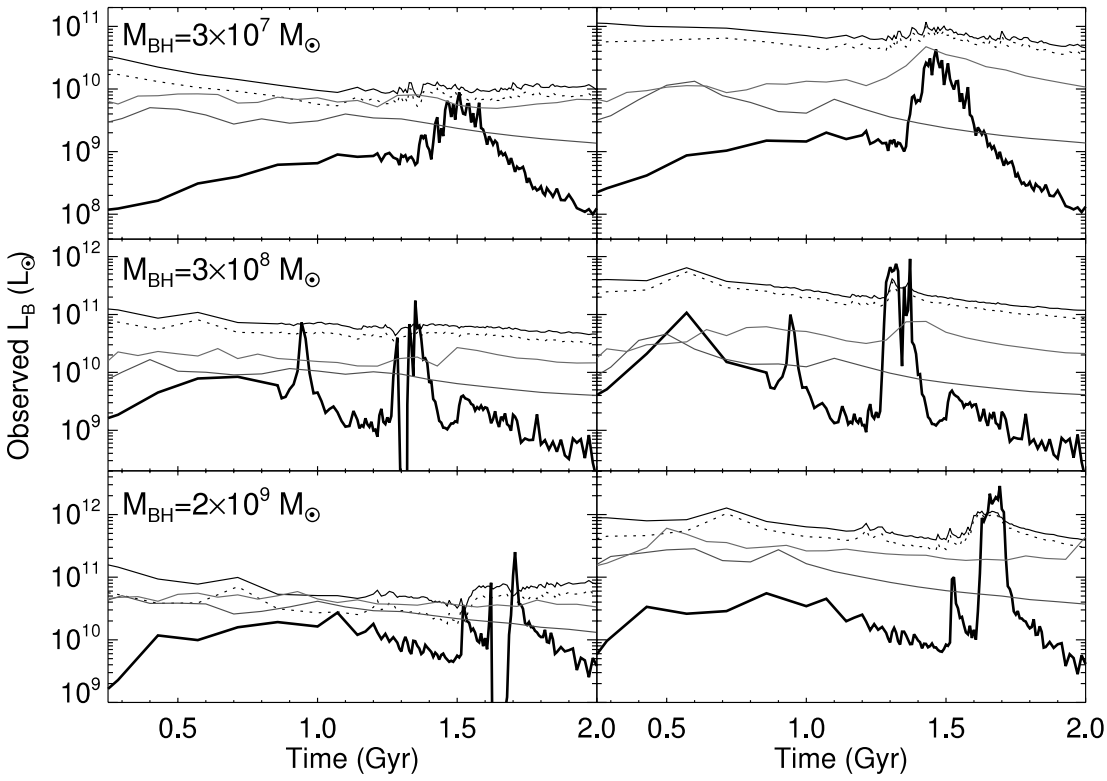

FIG. 13.—Intrinsic (*right*) and median attenuated (*left*) *B*-band luminosity of the quasar (*thick dark line*) and host galaxy (*thin dark line*, integrated over all stars; *dotted dark line*, ignoring bulge stars) as a function of time. Results are shown from three representative simulations: A2, A3, and A5 (see § 2.1) with $q_{\rm EOS} = 1.0$, $z_{\rm gal} = 0$, and virial velocities $V_{\rm vir} = 113,\ 160,$ and 320 km s^{-1}. Each quasar should be observable as a broad-line AGN when $L_{B,\,\rm QSO} \gtrsim L_{B,\,\rm host}$. Darker lines show the stellar light curve with gas fraction $f_{\rm gas} = 1.0$, the upper light solid line is for $f_{\rm gas} = 0.4$, and the lower light solid line is for $f_{\rm gas} = 0.2$; quasar light curves are similar for each gas fraction. [*See the electronic edition of the Supplement for a color version of this figure.*]

case shows the quasar *B*-band luminosity, and the thin dark line shows the integrated *B*-band luminosity of all stars in the galaxy. New stars are formed self-consistently in the simulations according to the ISM gas properties, equation of state, and star formation model described in Springel & Hernquist (2003), with the age and metallicity taken from the local star-forming ISM gas, which is enriched by supernova feedback from previous star formation. We then use the stellar population synthesis model of Bruzual & Charlot (2003) to determine the *B*-band luminosity (the *B*-band mass-to-light ratio) of new stars based on the stellar age and metallicity. The dotted line shows the result neglecting bulge particles, which must be initialized at the beginning of the simulation with random or uniform ages and metallicities instead of those quantities being determined self-consistently from the simulation physics. The right panels plot the intrinsic values of these quantities, and the left panels plot the median observed values of these quantities, where we have used our method for determining column densities and dust attenuation (§ 2.2) to every star and bulge particle for each line of sight.

Unfortunately, the host galaxy luminosity does *not* scale with instantaneous and peak quasar luminosity as do, for example, the quasar lifetime and obscuration. Rather, there are important systematic dependencies, the largest of which is the dependence on host galaxy gas fraction. If the host galaxies are more massive, more concentrated, or have a weaker ISM equation of state pressurization, then they will more effectively drive gas into the central regions and maintain high gas densities for longer periods of time, as the deeper potential well or lack of gas pressure requires more heat input from the quasar before the gas can be expelled. These conditions will generally produce a quasar with a larger peak luminosity (final black hole mass), but also form more new stars, meaning that the *B*-band relation between host and quasar luminosity is roughly preserved.

However, the black hole consumes only a small fraction of the available gas (comparison of, e.g., the stellar mass and black hole mass suggests the black hole consumes $\sim$0.1% of the gas mass), and so, at least above some threshold $f_{\rm gas} \lesssim 0.1$, the quasar peak luminosity does not significantly depend on the galaxy gas fraction (see, e.g., Fig. 2 of Robertson et al. 2005b). But, the mass of new stars formed *during* the merger does strongly depend on the available gas. For example, simulations that are otherwise identical but have initial $f_{\rm gas} = 0.2, 0.4, 0.8,$ and 1.0 (i.e., an increasing fraction of the initial disk mass in gas instead of stars) produce similar peak quasar luminosity and final total stellar mass (within $\sim$30% of one another), reflecting the conversion of most gas into stars and the fact that the peak quasar luminosity is determined more by the depth of the potential well than the total available gas supply. But, the mass of *new* stars formed in a merger scales roughly as $M_{*,\,\rm new} \propto f_{\rm gas}$ (as it must if the initial gas fraction does not change the final total stellar mass), and since young stellar populations dominate the observed *B*-band luminosity (especially during the peak merger and starburst phases associated with the bright quasar phase of interest), this implies roughly that $L_B \propto f_{\rm gas}$.

We demonstrate this explicitly in Figure 13, where we show in each panel the host galaxy and stellar *B*-band light curves for otherwise identical simulations with different gas fractions, $f_{\rm gas} = 0.2$ (*lower light solid line*), 0.4 (*upper light solid line*), and 1.0 (*dark lines*). In each of these cases, the quasar light curve is nearly identical (we show only the $f_{\rm gas} = 1.0$ quasar light curve, for clarity, but the others are within $\sim$30% of the curve shown at most times, with no systematic offset).

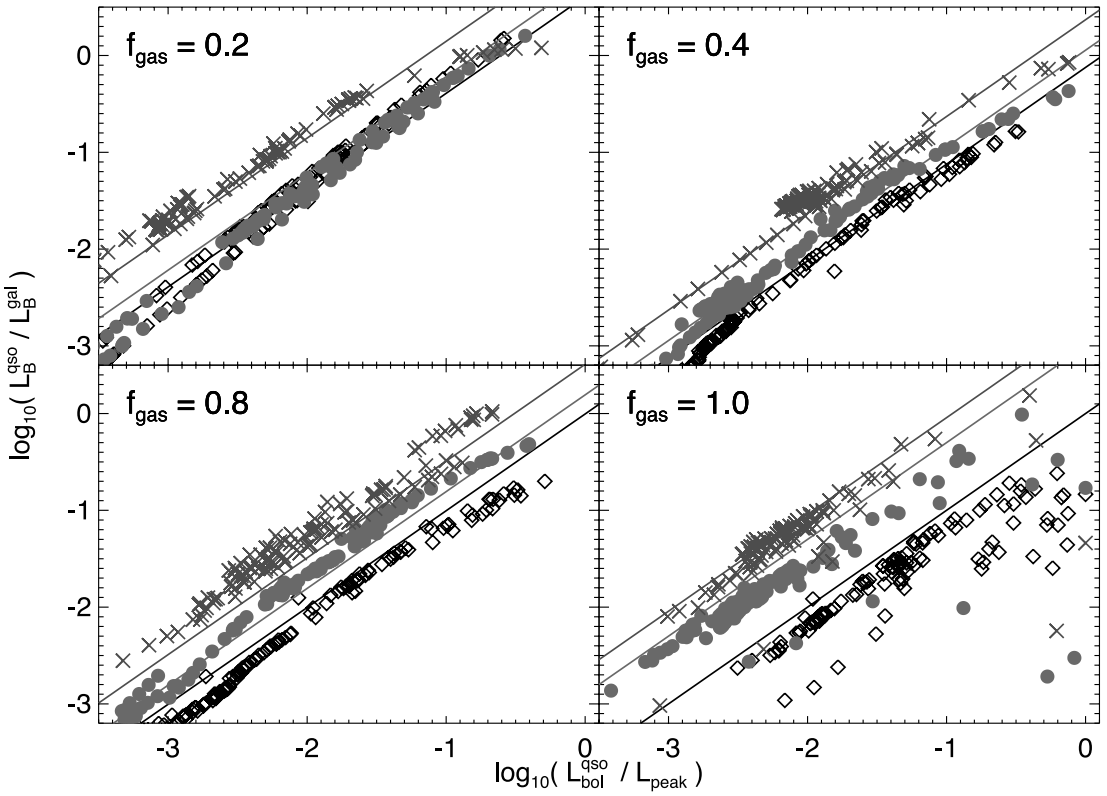

FIG. 14.—Ratio of observed (attenuated) *B*-band quasar luminosity to host galaxy luminosity as a function of the ratio of instantaneous to peak quasar bolometric luminosity. Results are from simulations A2 (*diamonds*), A3 (*circles*), and A5 (*crosses*) (the same simulations shown in Fig. 13) with $q_{\rm EOS} = 1.0$, $z_{\rm gal} = 0$, and virial velocities $V_{\rm vir} = 113,\ 160,$ and 320 km s^{-1}. Each panel shows the same simulations, except for a different initial gas fraction $f_{\rm gas} = 0.2,\ 0.4,\ 0.8,$ and 1.0, as labeled. Solid lines show the predictions of eq. (26). [*See the electronic edition of the Supplement for a color version of this figure.*]

In order for a quasar to be classified as a "broad-line" object, the optical spectrum must be visible and identified as such in the observed sample. This is clearly related to the ratio of quasar to host galaxy luminosity, but the threshold for classification is not obvious. In an X-ray- or IR-selected sample, optical follow-up should be able to disentangle host galaxy light and identify quasar broad-line spectra with fluxes a factor of several fainter than the host. However, automated optical selection based on color or morphological criteria might well exclude objects unless the quasar luminosity is a factor of several greater than that of the host galaxy. Therefore, there is significant systematic uncertainty in the theoretical definition of a broad-line quasar. To first order, based on the above arguments, we can classify "broad-line quasars" as objects in which the quasar optical luminosity is larger than some multiple $f_{\rm BL}$ of the host galaxy optical luminosity. Because the relevant ratio is different depending on the survey and selection techniques, we consider the range $f_{\rm BL} = 0.3-3$, with a rough median $f_{\rm BL} = 1$. Furthermore, because our simulations do not allow us to model the broad-line regions of the quasar or spectral line structures as influenced by, for example, reddening and dust absorption, we adopt the *B*-band luminosity of the quasar and host galaxy as a proxy for optical luminosity and more complex (but often quite sample-specific) color and morphological selection criteria.

In Figure 13 the *B*-band host galaxy luminosity is quite flat as a function of time, relative to the quasar *B*-band luminosity, and is roughly given by $L_B^{\rm gal}/L_\odot \sim M_{*,\,\rm new}/M_\odot$, where $M_{*,\,\rm new}$ is the mass of new stars formed in the merger. As noted above, this scales approximately linearly with initial gas fraction at fixed final total stellar mass M_*, giving $L_B^{\rm gal}/L_\odot \approx c_{\rm gal}(M_*/M_\odot) f_{\rm gas}$, where $c_{\rm gal}$ is a correction of order unity that we can fit from the simulations (essentially a mean mass-to-light ratio for the newly formed stars). The bolometric correction of the quasar is usually defined by $L_{\rm bol}^{\rm qso} = c_B L_B^{\rm qso}$, and the quasar peak luminosity is $L_{\rm peak} = c_L L_{\rm Edd}(M_{\rm BH}^f)$, where again c_L is a correction factor of order unity that we can calculate from our form for the quasar lifetime (see eq. [10]) or measure in the simulations.

If we require that the quasar *B*-band luminosity be larger than a factor $f_{\rm BL}$ of the host galaxy *B*-band luminosity, we obtain

$$L_{\rm bol}^{\rm qso}/L_\odot > f_{\rm BL} c_B c_{\rm gal}(M_*/M_\odot) f_{\rm gas}. \tag{23}$$

Dividing this through by $L_{\rm peak}$, we have

$$\frac{L_{\rm bol}^{\rm qso}}{L_{\rm peak}} \gtrsim 0.4 f_{\rm gas} f_{\rm BL} \left(\frac{c_{\rm gal}}{1.0}\right) \left(\frac{c_B}{12.0}\right) \left(\frac{M_{\rm BH}^f/M_*}{0.001}\right)^{-1} \left(\frac{c_L}{1.24}\right)^{-1}. \tag{24}$$

We can test this scaling relation against the results of our simulations, and do so in Figure 14. Rearranging the equations above gives

$$\frac{L_B^{\rm qso}}{L_B^{\rm gal}} \approx 3.4 f_{\rm gas}^{-1} \frac{L_{\rm bol}^{\rm qso}}{L_{\rm peak}} \times \left(\frac{c_{\rm gal}}{1.0}\right)^{-1} \left(\frac{c_B}{12.0}\right)^{-1} \left(\frac{M_{\rm BH}^f/M_*}{0.001}\right) \left(\frac{c_L}{1.24}\right), \tag{25}$$

which we can compare to our direct calculation of $L_B^{\rm qso}/L_B^{\rm gal}$ and $L_{\rm bol}^{\rm qso}/L_{\rm peak}$ for each simulation snapshot.

Ultimately, we are not interested so much in the intrinsic B-band luminosity of the quasar and host galaxy, but rather the observed luminosities; i.e., we are interested in the ratio $L^{\rm qso}_{B,\,\rm obs}/L^{\rm gal}_{B,\,\rm obs} = (L^{\rm qso}_{B}/L^{\rm gal}_{B})\{\exp\,[-(\tau_Q - \tau_G)]\}$, where τ_Q and τ_G are "effective" optical depths that we use to denote the mean attenuation of quasar and host galaxy B-band luminosities, respectively. We have considered the distribution of column densities attenuating the quasar as a function of instantaneous and peak quasar luminosity in detail in § 2.3 above; the attenuation of the host galaxy as a function of luminosity, observed band, halo mass, and star formation rate is discussed in detail in Jonsson et al. (2006). Combining these fits gives, roughly, $\exp\,[-(\tau_Q - \tau_G)] \sim (M^{f}_{\rm BH}/10^8\ M_\odot)^{0.16}$, but a better approximation can be determined directly from the simulations.

This scaling can be understood roughly using toy models of uniformly mixed luminous sources within the galaxy described by Jonsson et al. (2006), after accounting for the fact that the luminosity (star formation rate) dependent portion of the attenuation scales with luminosity in a manner similar to our quasar attenuation (compare our $\tau_Q \propto N_{\rm H} \propto L^{0.43-0.54}_{\rm qso}$ to their $\tau_G \propto L^{0.55}_{B,\,\rm gal}$). The key consequence of this is that more massive systems (higher bulge and black hole masses) have their host galaxy light proportionally more attenuated in mergers, meaning that (as suggested by the comparison of light curves in Fig. 13) the quasar is more likely to be observed with an optical luminosity larger than that of its host.

Figure 14 plots the ratio of the observed (attenuated) B-band quasar luminosity to the observed host galaxy B-band luminosity as a function of the ratio of instantaneous to peak quasar bolometric luminosity. We show the results for four different gas fractions $f_{\rm gas} = 0.2,\ 0.4,\ 0.8,$ and 1.0, as labeled. For each gas fraction, we consider our simulations A2 (*diamonds*), A3 (*circles*), and A5 (*crosses*) (the same simulations shown in Fig. 13) with $q_{\rm EOS} = 1.0,\ z_{\rm gal} = 0$, and virial velocities $V_{\rm vir} = 113, 160,$ and 320 km s^{-1}, using the labeled initial gas fraction. The solid lines in each panel show the predictions of combining the scalings expected for the intrinsic luminosities (eq. [25]) and attenuations as above, giving

$$\frac{L^{\rm qso}_{B,\,\rm obs}}{L^{\rm gal}_{B,\,\rm obs}} = 7.9\frac{1}{f_{\rm gas}}\left(\frac{M^{f}_{\rm BH}}{10^8\ M_\odot}\right)^{0.2}\frac{L}{L_{\rm peak}}, \qquad (26)$$

where the solid lines each use the $M^{f}_{\rm BH}$ and $f_{\rm gas}$ of the simulation of the corresponding symbol and panel. This scaling provides a good estimate of the observed optical quasar-to-galaxy luminosity ratio, including the complicated effects of attenuation, evolving mass-to-light ratios, metallicities, and host galaxy properties, as a function of gas fraction, final black hole mass, and the ratio of the current to peak quasar luminosity. Although, for clarity, we have not shown a range of simulations varying other parameters, we find that this scaling is robust to the large number of quantities we have considered in our simulations; there are systematic offsets in, for example, $L_{\rm peak}$ and $M^{f}_{\rm BH}$ with changes such as, for example, different ISM equations of state, but the scaling in terms of $L_{\rm peak}$ and $M^{f}_{\rm BH}$ is unchanged.

Because the ratio of observed quasar and host galaxy B-band luminosities in our simulations obeys the scaling of equation (26), we can use it to predict the properties of broad-line quasars, defined by $L^{\rm qso}_{B,\,\rm obs} > f_{\rm BL}\,L^{\rm gal}_{B,\,\rm obs}$. To do so, however, we must assume a typical host galaxy gas fraction. Unfortunately, because our empirical modeling in terms of the quasar lifetime as a function of L and $L_{\rm peak}$ does not have a systematic dependence on host galaxy gas fraction (see § 2.4), we have no constraint on this parameter. It is, however, convenient for several reasons to consider $f_{\rm gas} = 0.3$ as a typical value for bright quasars.

First, such a gas fraction is capable of yielding the brightest observed quasars; second, scaling a Milky Way–like disk with the observed $z = 0$ gas fraction of $\sim$0.1 to the redshifts of peak quasar activity gives a similar gas fraction (e.g., Springel et al. 2005a); third, gas fractions of $\gtrsim$30% in major mergers are needed to explain the observed fundamental plane (Robertson et al. 2006), kinematic properties (T. J. Cox et al. 2006, in preparation), and central phase space densities (Hernquist et al. 1993) of elliptical galaxies; fourth, this choice implies that the brightest quasars with $M^{f}_{\rm BH} \sim 10^{10}\ M_\odot$ attain observed B-band luminosities $\sim$1000 times that of their hosts at their peaks, as is observed (e.g., McLure & Dunlop 2004). Finally, and most importantly, the assumed $f_{\rm gas}$ and $f_{\rm BL}$ are degenerate in our predictions for the broad-line population, as they both enter linearly in the ratio of host galaxy to quasar B-band luminosity. Therefore, the range of $f_{\rm BL} = 0.3-3$ that we consider (for a fixed median $f_{\rm gas} = 0.3$) can be equivalently considered, for a fixed median $f_{\rm BL} = 1$, to represent a theoretical uncertainty in the host galaxy gas fraction, $f_{\rm gas} = 0.1-0.9$, i.e., spanning the range from present, relatively gas-poor Milky Way–like disks to almost completely gaseous disks. This, then, gives for our broad-line criterion

$$\frac{L}{L_{\rm peak}} \gtrsim 0.2\left(\frac{f_{\rm BL}}{1.0}\right)\left(\frac{f_{\rm gas}}{0.3}\right)\left(\frac{M^{f}_{\rm BH}}{10^7\ M_\odot}\right)^{-0.2}. \qquad (27)$$

The broad-line phase is thus, as is clear from Figure 13 and implicit in our definition of the broad-line phase, closely associated with the final blowout stages of quasar evolution, when the mass of the quasar reaches that corresponding to its location on the $M_{\rm BH}$-σ relation and gas is expelled from the central regions of the galaxy, shutting down accretion (Di Matteo et al. 2005). We note that combining the equation above with our fitted quasar lifetimes gives an integrated time when the quasar would be observable as a broad-line object of $t_{\rm BL} \sim 10-20$ Myr, in good agreement with the optically observable bright quasar lifetimes that we calculate directly from our quasar light curves, including the effects of attenuation, and with empirical estimates of the quasar lifetime that are based directly on optically selected, broad-line quasar samples.

The $(M^{f}_{\rm BH}/10^7\ M_\odot)^{0.2}$ term in the above equation reflects the fact that below a certain peak luminosity, quasars are less likely to reach luminosities above that of the host galaxy, as can be seen in the uppermost panels of Figure 13 for a final black hole mass of $M^{f}_{\rm BH} = 3\times10^7$; i.e., the smallest AGNs are proportionally less optically luminous than their hosts. This does not imply that such systems are not inherently broad-line objects, but only that the host galaxy light will increasingly dominate at lower luminosities. We also caution against extrapolating this to large or small $M^{f}_{\rm BH}$, as the attenuation becomes more difficult to predict at these peak luminosities and the linear formula above is not always accurate (see Fig. 14).

We can use this estimate of the broad-line phase and our model of the quasar lifetime to calculate the total energy radiated in this bright, optically observable stage following the calculation of § 2.4, but with a minimum luminosity determined by equation (27). This gives an integrated fraction of $\sim$0.3–0.4 ($\sim\exp\,[-0.2f_{\rm BL}(\,f_{\rm gas}/0.3)/\alpha_L]$) of the total radiant energy emitted during the broad-line phase. Thus, despite the short duration of this optical quasar stage, a large fraction of the total radiated energy is emitted (as it represents the final e-folding in the growth of the black hole) when most of the final black hole mass (§ 2.4)

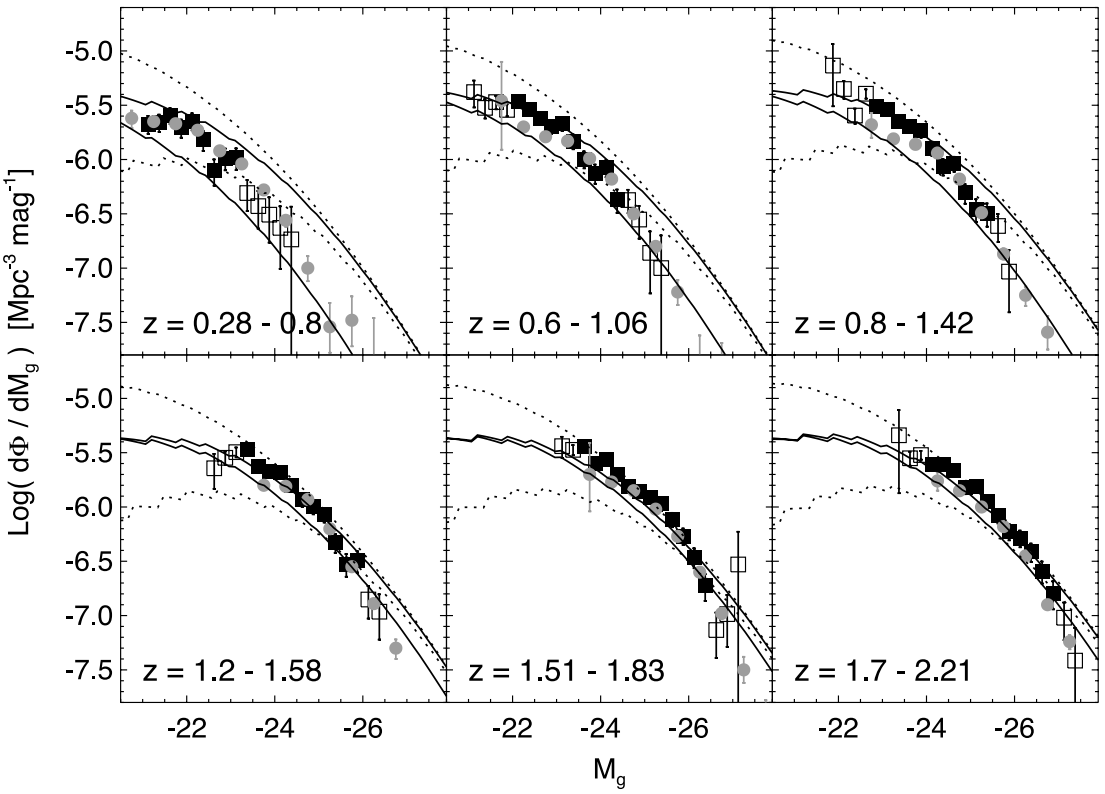

FIG. 15.—Broad-line quasar luminosity function of Richards et al. (2005) from the 2dF-SDSS (2SLAQ) survey (*squares*) and that of Croom et al. (2004; *circles*) from the 2dF survey, compared to our predicted broad-line luminosity function from our determination of the relative quasar and host galaxy luminosities in our simulations (*solid line*), where we estimate that quasars are observable as broad-line objects when their observed *B*-band luminosity is greater than a factor $f_{\rm BL}$ of that of the host galaxy. Solid lines are shown for the minimum and maximum observed redshift in each bin (as labeled), assuming $f_{\rm BL} = 1$. Dotted lines show the result for $f_{\rm BL} = 0.3$ and $f_{\rm BL} = 3$, at the mean redshift of the bin, i.e., corresponding to broad-line luminosity functions in surveys that are complete out to quasars with observed optical luminosities $\sim$0.3 and 3 times that of the host galaxy, respectively, or alternatively, reflecting nearly complete theoretical uncertainty regarding merging galaxy gas fractions ($f_{\rm gas} = 0.1-0.9$). Open squares show observations with uncertain incompleteness corrections in Richards et al. (2005). [*See the electronic edition of the Supplement for a color version of this figure.*]

is accumulated. Accounting for the luminosity dependence of our bolometric corrections (with the optical fraction of the quasar energy increasing with bolometric luminosity), as well as for the small fraction of objects observable at lower luminosities (with larger typical obscuring column densities), increases this fraction to as much as $\sim$0.6–0.7 for bright quasars. Therefore, despite the fact that the *duration* of the optically observable broad-line quasar phase may be $\sim$1/10 that of the obscured quasar growth phase, the changing quasar luminosity over this period and nontrivial quasar lifetime as a function of luminosity implies only small corrections to counting arguments such as that of Soltan (1982), which rely on the total observed optical quasar flux density to estimate the relic supermassive black hole density.

4.2. *The Broad-Line Fraction as a Function of Luminosity*

By estimating the time that a quasar with some $L_{\rm peak}$ will be observable as a broad-line quasar at a given luminosity, we can then calculate the broad-line quasar luminosity function in the same fashion as the complete quasar luminosity function in § 3.2. Instead of the full quasar lifetime $dt/d\log L$, we consider only the time during which broad-lines would be observed (i.e., during which the quasar spectrum would be recognized, as opposed to the host galaxy spectrum), as identified in our simulations (§ 4.1).

For a sample selected in hard X-rays (i.e., the selection function only being relevant at column densities $\gtrsim 10^{24}$ cm^{-2}), we show the resulting broad-line luminosity function in Figure 9 (*dot-dashed lines*) and compare it to the broad-line quasar luminosity function identified in the hard X-ray luminosity function of Barger et al. (2005). The agreement is good at all luminosities, and our model explains both the fact that broad-line quasars dominate the luminosity function at luminosities well above the break in the luminosity function, and the downturn in the broad-line quasar population at luminosities below the peak. Essentially, the broad-line quasar population more closely traces the shape of the $\dot{n}(L_{\rm peak})$ distribution, giving rise to the observed behavior as a dual consequence of luminosity-dependent quasar lifetimes and the evolutionary nature of quasar obscuration in our simulations.

Figure 15 compares our theoretical predictions to the 2dF-SDSS (2SLAQ) *g*-band luminosity function of broad-line quasars from Richards et al. (2005; *squares*), as well as the *B*-band luminosity function from Croom et al. (2004; *circles*), at several redshifts from $z \sim 0.3$ to 2, over which range the surveys are expected to be relatively complete (for broad-line quasars). The 2dF-SDSS result is the most recent determination of the broad-line luminosity function, but compares well with previous determinations by, e.g., Boyle et al. (1988, 1990, 1991, 2000), Koo & Kron (1988), Marano et al. (1988), Zitelli et al. (1992), and Croom et al. (2004). Open squares correspond to bins in luminosity that have been corrected for incompleteness following Page & Carrera (2000), but this correction is uncertain as the bins are not uniformly sampled. We compare this at each redshift to the prediction of our determination of the quasar broad-line phase, where we estimate that the quasar is observable as a broad-line object when its observed *B*-band luminosity is greater than a factor

$f_{\rm BL} = 1$ of that of the host galaxy. We calculate this for both the minimum and maximum observed redshift of each bin to show the range owing to evolution of the luminosity function over each interval in redshift. The systematic uncertainty in our prediction can be estimated from the dotted lines, which show the prediction (at the mean redshift of the bin) if we instead require the observed quasar *B*-band luminosity to be above a factor of 0.3 (*upper lines*) or 3 (*lower lines*) of the observed host galaxy *B*-band luminosity, which as discussed in § 4.1 can alternatively be considered an uncertainty in host galaxy gas fraction, with $f_{\rm gas} = 0.1$ and $f_{\rm gas} = 0.9$, respectively.

The agreement at all luminosities and redshifts shown is encouraging, given the simplicity of our determination of the broad-line phase from the simulations, but the systematic uncertainties are large, emphasizing the importance of calculating detailed selection effects in contrasting, for example, broad-line samples from optical and X-ray surveys, as opposed to assuming a constant obscured fraction at a given luminosity based on the ratio of luminosity functions, as has been adopted in previous phenomenological models. The difference between different choices of $f_{\rm gas}$ is suppressed at the high-luminosity (and correspondingly high-redshift) end of the luminosity function, because the quasar-to-galaxy *B*-band luminosity ratio scales as $\propto (M_{\rm BH}^{f})^{0.2}$; i.e., regardless of the choice of $f_{\rm BL}$, quasars increasingly overwhelm their host galaxy in large systems near their peak luminosity. However, at low luminosity, the predictions rapidly diverge, implying that a measurement of the faint end of the broad-line quasar luminosity function, with a reliable calibration of $f_{\rm BL}$, can constrain the typical gas fractions of quasar host galaxies and the evolution of these gas fractions with redshift.

By dividing out the predicted luminosity function $\phi_{\rm HX}$, we can estimate the fraction of broad-line objects observed in reasonably complete X-ray samples as a function of luminosity. This is shown in Figure 16, where for ease of comparison we have shown the broad-line fraction as a function of hard X-ray (2–10 keV) luminosity. Our prediction, based on determining the time a quasar with a given luminosity L and peak luminosity $L_{\rm peak}$ in our simulations will be observable with a *B*-band luminosity greater than a fraction $f_{\rm BL} = 1.0$ of the host galaxy observed *B*-band luminosity, is shown as the thick dark line. This is compared to the observations of Ueda et al. (2003; *squares*), Hasinger (2004; *circles*), Grimes et al. (2004; *triangles*), and Simpson (2005; *diamonds*). The data from Hasinger (2004) have been scaled from soft X-ray (0.5–2 keV) using our bolometric corrections, and the data from Grimes et al. (2004) and Simpson (2005) have been converted from [O III] luminosity as in Simpson (2005) using the mean correction for Seyfert galaxies (Mulchaey et al. 1994), $L_{\rm [O\ III]} = 0.015 \times L_{\rm 2-10\ keV}$.

We also plot as upper and lower dashed lines the results of changing $f_{\rm BL}$, the fraction of the host galaxy *B*-band luminosity above which the quasar *B*-band luminosity must be observed for identification as a broad-line object, considering $f_{\rm BL} = 0.3$ and 3, respectively. We determine this for the low-redshift $z \lesssim 0.3$ quasar distribution, from which most of the data are drawn. The dot-dashed line shows the difference at high redshift, if just $z \gtrsim 1$ quasars are considered (for $f_{\rm BL} = 1$). The broad-line fraction is systematically lower, primarily because the break luminosity in the luminosity function moves to higher luminosity with redshift, meaning that at a fixed luminosity below the break, a smaller fraction of observed objects are at $L \sim L_{\rm peak}$ in the blowout phase of peak optical quasar luminosity. Finally, the dotted line shows the results assuming a light-bulb model for the broad-line phase [but still using our $\dot{n}(L_{\rm peak})$ distribution; otherwise, this translates to a constant obscured fraction with luminosity] lifetimes, with a fixed broad-line lifetime of $t_Q = 20$ Myr.

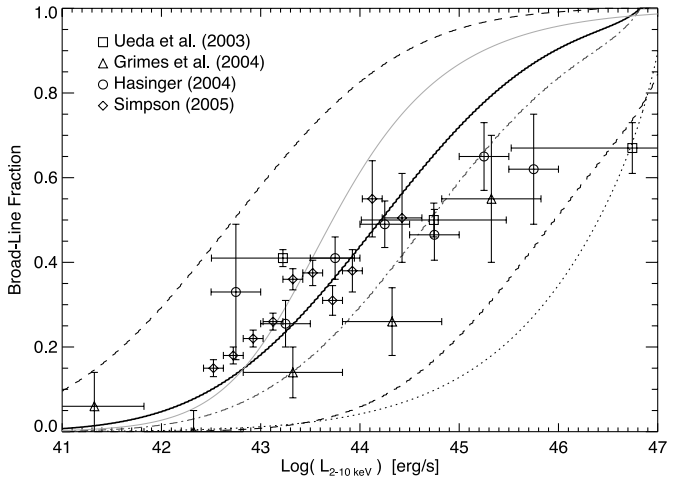

FIG. 16.—Predicted broad-line fraction of a complete X-ray sample at low $z \lesssim 0.3$ redshift, from our simulations (where the object is observable as a broad-line quasar when it has an observed *B*-band luminosity greater than a factor $f_{\rm BL} = 1.0$ of that of its host galaxy), is shown (*thick dark line*). The results, changing our $f_{\rm BL}$ to 0.3 and 3.0, are shown, or equivalently, of assuming a host galaxy gas fraction $f_{\rm gas} = 0.1$ or 0.9 instead of $\sim$0.3 (*dashed lines*), as are the results assuming a light-bulb model where quasars spend a fixed time $t_Q = 20$ Myr as broad-line objects with a luminosity of $L_{\rm peak}$ (*dotted line*). For comparison, the (scaled to 2–10 keV luminosity) observations of Ueda et al. (2003; *squares*), Hasinger (2004; *circles*), Grimes et al. (2004; *triangles*), and Simpson (2005; *diamonds*) are shown. The predicted result at higher redshift ($z \gtrsim 1$) is shown (*dot-dashed line*), offset owing to the shift in break luminosity of the luminosity function with redshift. The best-fit luminosity-dependent torus model, *fitted* to the data, is shown as the light solid line. The best-fit static torus model is a constant broad-line fraction $\sim$0.3 (not shown for clarity). [*See the electronic edition of the Supplement for a color version of this figure.*]

The prediction of the most basic torus model, with constant broad-line fraction of $\sim$0.36, is ruled out to high significance ($\chi^2/\nu = 18.5$ or 17.2 if we consider all data points or only the most well-constrained data, respectively, from Simpson [2005]). Furthermore, the light solid line shows the best-fit luminosity-dependent torus model, in which the broad-line fraction is given by (e.g., Simpson 1998; Grimes et al. 2004)

$$f = 1 - 1/\sqrt{1 + 3L/L_0}, \tag{28}$$

where L_0 is the luminosity where the number of broad-line objects is equal to the number of non-broad-line objects. This fit is at best marginally acceptable over a narrow range in luminosities ($\chi^2/\nu = 14.0,\ 7.3$). Modified luminosity-dependent, receding torus models have been proposed that give a better fit to the data by, for example, allowing the torus height to vary with luminosity (e.g., Simpson 2005), but there is no physical motivation for these changes, and they introduce such variation through additional free parameters that allow a curve of essentially arbitrary slope to be fitted to the data.

However, the prediction of our model agrees reasonably well ($\chi^2/\nu = 4.0,\ 1.2$) with the observations over the entire range covered, a span of 6 orders of magnitude in luminosity. We emphasize that our prediction, which matches the data better than standard torus models that are actually *fitted* to the data, is not a fit to the observations. Instead, it is derived from the physics of our simulations, including the black hole accretion and feedback that are critical in driving the blowout phase, which constitutes most of the time a quasar is visible as a broad-line object by our estimation, and from the $\dot{n}(L_{\rm peak})$ distribution implied by our model of quasar lifetimes and the *bolometric* quasar luminosity

function. The agreement suggests that our choice of the parameter combination $f_{\rm BL} f_{\rm gas} = 0.3$ is a good approximation. As noted above, this implies that calibrating $f_{\rm BL}$ for an observed sample, combined with the mean broad-line fraction and our modeling, can provide a constraint (albeit model-dependent) on the host galaxy gas fraction of quasars at a given redshift, which cannot necessarily be directly measured even with difficult, detailed host galaxy probes, as gas is rapidly converted into stars throughout the merger. The uncertainty plotted, while large, actually represents a larger theoretical uncertainty: as discussed above, if an observational sample were well-defined such that it were complete to broad-line objects with observed optical luminosity above a fraction $f_{\rm BL}$ of the host galaxy luminosity, the range that we consider would correspond to a range $f_{\rm gas} = 0.1-0.9$ in the quasar host galaxy gas fraction, which the observations could then constrain.

In our modeling, the broad-line fraction as a function of luminosity does not depend sensitively on the observed luminosity function, as evidenced by the relatively similar prediction at high redshift. The evolution that we do predict with redshift, in fact, agrees well with that found by Barger et al. (2005) over the redshift range $z = 0.1-1.2$ (see also La Franca et al. 2005), an aspect of the observations that is not reproduced in any static or luminosity-dependent torus model but follows from the evolution of the quasar luminosity function in our picture for quasar growth. However, we do caution that gas fractions may systematically evolve with redshift, and as discussed above, a higher gas fraction will give generally shorter broad-line lifetimes using our criteria of quasar optical luminosity being higher than some fraction of the host galaxy luminosity, which will also contribute to the evolution in the mean broad-line fraction with redshift. Finally, neglecting the role of luminosity-dependent quasar lifetimes gives unacceptable fits to the data ($\chi^2/\nu = 66.0, 77.5$), as the broad-line fraction as a function of luminosity is a consequence of both the evolution of obscuration and the dependence of lifetime on luminosity.

Our model for quasar evolution provides a direct physical motivation for the change in broad-line fraction with luminosity and suggests that it is not a complicated selection effect. As an observational sample considers higher luminosities (i.e., approaches and passes the break in the observed luminosity function), a comparison of the luminosity function and the underlying $\dot{n}(L_{\rm peak})$ shows that it is increasingly dominated by sources near their peak luminosity in the final stages of Eddington limited growth. The final stages of this growth expel the large gas densities obscuring the quasar, rendering it a bright, optically observable broad-line object for a short time. Therefore, we expect that the fraction of broad-line objects should increase with luminosity in quasar samples, as indicated by the observations.

Many observational measures do not consider a direct optical analysis of the quasar spectrum in estimating the fraction of broad-line objects as a function of luminosity. For example, Ueda et al. (2003) adopt a proxy, classifying as "obscured" any quasars with an X-ray-identified column density $N_{\rm H} > 10^{22}$ cm^{-2}, and as "unobscured" quasars below this column density. We can compare to their observations using the column density distributions as a function of luminosity from our simulations, which cover the entire range in luminosity of the observed sample. Specifically, we use a Monte Carlo realization of these distributions, employing our fitted $\dot{n}(L_{\rm peak})$ distribution at each redshift to produce a list of quasar peak luminosities and then generating all other properties based on the probability distribution of a given property in simulations with a similar peak luminosity. We describe this methodology in detail in § 8 and provide several such mock quasar distributions at different redshifts.

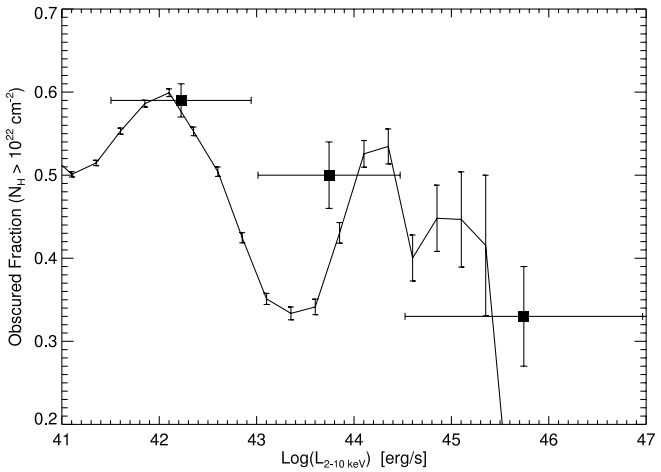

FIG. 17.—Predicted "obscured" fraction (*solid line*) in an X-ray sample with identical redshift and luminosity range to that of Ueda et al. (2003), as a function of hard X-ray (2–10 keV) luminosity. Vertical error bars show Poisson errors estimated from the total time at a given luminosity across all our simulations (absolute values of the error bars should not be taken literally). The "obscured" fraction is defined as the fraction of quasars with X-ray column densities $N_{\rm H} > 10^{22}$ cm^{-2} in bins of $\Delta \log L_{2-10\,\rm keV}$. The observations from Ueda et al. (2003) are shown as filled squares.

In Figure 17 we compare our estimated obscured and unobscured fractions as a function of hard X-ray luminosity, using the same definitions, as well as redshift and luminosity limits, as the observed sample. The solid line shows our prediction, with vertical error bars representing Poisson errors, where the number of "counts" is proportional to the total time spent by simulations at the plotted luminosity (the absolute value of these errors should not be taken seriously). The obscured fraction is determined in bins of luminosity $\Delta \log L_{2-10\,\rm keV} = 0.5$. Despite our large number of simulations, there is still some artificial "noise" owing to incomplete coverage of the merger parameter space, namely, the apparent oscillations in the obscured fraction. However, the mean trend agrees well with that observed (*squares*), suggesting that the success of our modeling in reproducing the fraction of broad-line objects as a function of luminosity is not a consequence of the definitions chosen above. We do not show the predictions of the standard and luminosity-dependent torus models, as (because essentially any line of sight through the torus encounters a column density $N_{\rm H} > 10^{22}$ cm^{-2}) the predictions of these models are identical to those shown and compared to the same observations in Figure 16.

Our prediction that the fraction of broad-line objects should rise with increasing luminosity is counterintuitive, given our fitted column density distributions in which typical (median) column densities increase with increasing luminosity. This is primarily due to the simplicity of our $N_{\rm H}$ fits; we assume that the distribution is lognormal at all times, but a detailed inspection of the cumulative (time integrated) column density distribution shows that at bright (near peak) luminosities, the distribution is in fact bimodal (see, e.g., Fig. 3 of Hopkins et al. 2005a and Fig. 2 of Hopkins et al. 2005c), representing both the heavily obscured growth phase and the blowout phase that we have identified here as the broad-line phase. Over most of a simulation, we find the general trend shown in Figure 3 and discussed above, namely, that typical column densities increase with intrinsic (unobscured) luminosity. This is because the total time at moderate to large luminosities is dominated by black holes growing in the obscured/starburst stages; here, the same gas inflows fueling black hole growth also give rise to large column densities and starbursts that obscure the

black hole activity. However, when the quasar nears its final, peak luminosity, there is a rapid blowout phase as feedback from the growing accretion heats the surrounding gas, driving a strong wind and eventually terminating rapid accretion, leaving a remnant with a black hole satisfying the $M_{\rm BH}$-σ relation. This can be identified with the traditional bright optical quasar phase, as the final stage of black hole growth with a rapidly declining density (allowing the quasar to be observed in optical samples), giving typical luminosities, column densities, and lifetimes of optical quasars. In these stages, larger luminosities imply more violent blowout events; i.e., a brighter peak luminosity quasar more effectively expels the nearby gas and dust, rendering a dramatic decrease in column density at these bright stages (see Hopkins et al. 2006a).

We are essentially modeling this bimodality in more detail by directly determining the broad-line phase from our simulations. However, the broad-line fraction–luminosity relation that we predict is also a consequence of the more complicated relationship between column density, peak luminosity, and bolometric and observed luminosity, as opposed to the predictions from a model with correlation between $N_{\rm H}$ and only observed luminosity. The key point is that we find, near the *peak* luminosity of the quasar, as feedback drives away gas and slows down accretion, the typical column densities fall rapidly with luminosity in a manner similar to that observed. In our model of the luminosity function, quasars below the observed break are either accreting efficiently in early stages of growth or are in sub-Eddington phases coming into or out of their peak quasar activity. Around and above the break, the luminosity function becomes dominated by sources at high Eddington ratio at or near their peak luminosities. Based on the above calculation, we then *expect* what is observed, that in this range of luminosities, the fraction of objects observed with large column densities will rapidly decrease with luminosity as the observed sample is increasingly dominated by sources at their peak luminosities in this blowout phase. This also further emphasizes that the evolution of quasars dominates over static geometrical effects in determining the observed column density distribution at any given luminosity.

Finally, if host galaxy contamination were not a factor, we would expect from our column density model that, at low luminosities ($L \lesssim 10^{10}\ L_{\odot}$, well below the range of most observations shown), the broad-line fraction would again increase (i.e., the obscured fraction would decrease), as the lack of gas to power significant accretion would also imply a lack of gas to produce obscuring columns. However, at these luminosities, typical of faint Seyfert galaxies or LINERs, our modeling becomes uncertain; it is quite possible, as discussed previously, that cold gas remaining in relaxed systems could collapse to form a traditional dense molecular torus on approximately parsec scales, well below our resolution limits. Furthermore, host galaxy light is likely to overwhelm any AGN broad-line contribution, and selection effects will also become significant at these luminosities.

4.3. *The Distribution of Active Broad-Line Quasar Masses*

Our determination of the broad-line or optical phase in quasar evolution allows us to make a further prediction, namely, the mass distribution of currently active broad-line quasars. At some redshift, the total number density of observed, currently active broad-line quasars with a given $L_{\rm peak}$ will be (in the absence of selection effects)

$$n_{\rm BL}(L_{\rm peak}) \approx \dot{n}(L_{\rm peak}) t_{\rm BL}(L_{\rm peak}), \quad (29)$$

where $t_{\rm BL}(L_{\rm peak})$ is the total integrated time that a quasar with peak luminosity $L_{\rm peak}$ spends as a broad-line object (using our criterion for the ratio of the observed quasar B-band luminosity to that of the host galaxy), given by integrating our formulae in § 4.1 or directly calculated from the simulations. Since we have determined roughly that a quasar should be observable as a broad-line object at times with $L \gtrsim 0.2 L_{\rm peak}$ primarily just after it reaches its peak luminosity, in the blowout phase of its evolution, we expect the instantaneous black hole mass at the time of observation as a broad-line quasar to be, on average, $M_{\rm BH}^{\rm BL} \approx M_{\rm BH}^{f}(L_{\rm peak})$, where $M_{\rm BH}^{f} \sim M_{\rm Edd}(L_{\rm peak})$, modulo the order unity corrections described in § 2.4. Using our fitted $\dot{n}(L_{\rm peak})$ distribution from the luminosity function, extrapolated to low redshift ($z \sim 0$), and combining it with the integrated broad-line lifetimes from our simulations as above, we obtain the differential number density of sources in a logarithmic interval in $L_{\rm peak}$. Finally, we use our equation (10) for $M_{\rm BH}^{f}(L_{\rm peak})$ determined from our fitted quasar lifetimes (demanding that $E_{\rm rad} = \epsilon_r M_{\rm BH}^{f} c^2$) to convert this to a distribution in black hole mass.

Our predicted $n(M_{\rm BH})$, i.e., the number of observed *active* quasars at low redshift in a logarithmic interval of black hole mass, is shown in Figure 18. We consider the complete distribution of active quasar masses, for both broad-line and non-broad-line objects, in the left panel of the figure, and the distribution of broad-line objects only, $n(M_{\rm BH}^{\rm BL})$, in the right panel. On the left, we show the complete distribution that would be observed without any observational limits (*dashed line*). We calculate this from the distributions of Eddington ratios in our simulations, as a function of current and peak luminosity, and our fit to $\dot{n}(L_{\rm peak})$ (as, e.g., for our Monte Carlo realizations). We also consider the observed distribution if we apply the luminosity limit for completeness from the SDSS sample of Heckman et al. (2004; *dotted line*), $L_{[\rm O\ III]} > 10^6\ L_{\odot}$, which using their bolometric corrections yields $L > 3.5 \times 10^9\ L_{\odot}$, and then additionally applying their minimum velocity dispersion $\sigma > 70$ km s^{-1} (*dot-dashed line*). Finally, we can weight this distribution by luminosity (*solid line*) to compare directly to that determined in their Figure 1. The circles are taken from the luminosity-weighted black hole mass function of Heckman et al. (2004), which serves as a rough estimate of the *active* black hole mass distribution given their selection effects. Vertical error bars represent the range in parameterizations of the mass function from Heckman et al. (2004), including whether or not star formation is corrected for and limiting the sample to luminosities $L \gtrsim 10^{10}\ L_{\odot}$ or Eddington ratios > 0.01. Horizontal errors represent an uncertainty of 0.2 dex in the black hole mass estimation (representative of uncertainties in the $M_{\rm BH}$-σ relation used). The agreement is good, especially given the significant effects of the selection criteria and luminosity weighting.

We also consider the predictions of a light-bulb or exponential/fixed-Eddington-ratio model of the quasar lifetime for the active black hole mass distribution (*lighter lines*). For purposes of the active black hole mass function, the two predictions are identical and independent of the assumed quasar lifetime (modulo the arbitrary normalization), as both assume that all observed quasars are accreting at a fixed Eddington ratio, giving the distribution of active black hole masses. The dashed line shows the prediction for the complete active black hole mass function, which rises sharply to lower luminosities, as it must given a luminosity function that increases monotonically to lower luminosities. The solid line shows the prediction of such a model with the complete set of selection effects from Heckman et al. (2004) described above applied, as with the solid black line showing the prediction of our modeling. Here we chose the characteristic Eddington ratio ≈ 1.0 by fitting the predicted curve to the Heckman et al. (2004) observations. Note that both the characteristic Eddington ratio and lifetime (normalization) of the curve are fitted, so the relative

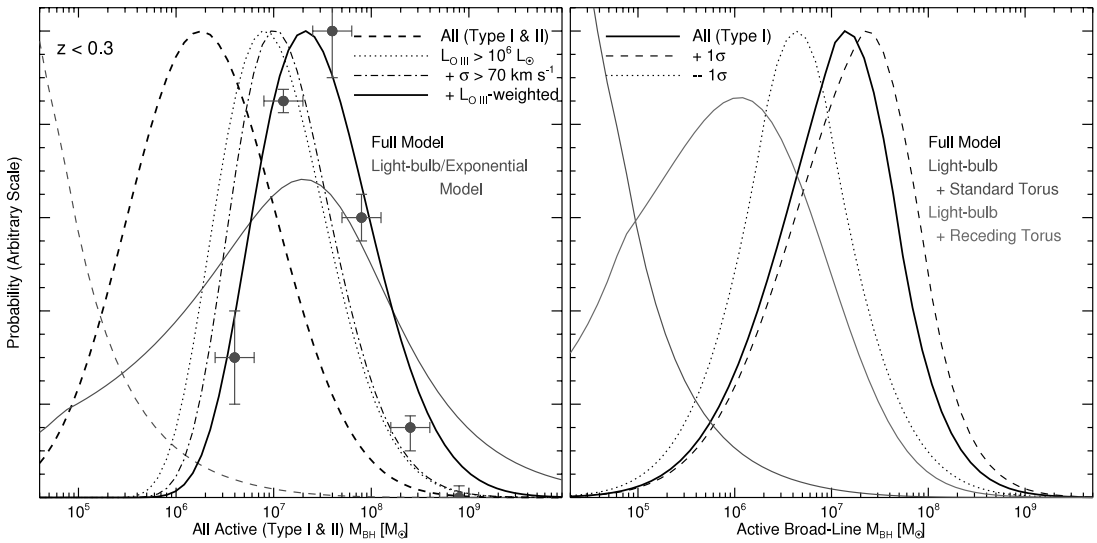

FIG. 18.—Predicted distribution of currently active black hole masses, considering both all types (types I and II; *left*) and only those visible as broad-line quasars (type I; *right*), at low $z \lesssim 0.3$ redshift, from our $\dot{n}(L_{\rm peak})$ distribution and the estimation of the broad-line phase directly from the simulations. In the left panel (all quasar types), we consider the result with arbitrarily faint luminosity limits (*dashed line*), and with the luminosity completeness limit (*dotted line*) and both the luminosity limit and velocity dispersion limit (*dash-dotted line*) of the SDSS sample of Heckman et al. (2004). We then consider the mass distribution with these limits, weighted by O III luminosity, for direct comparison to the mass function of Heckman et al. (2004), shown as circles (vertical error bars represent the range in different parameterizations of the luminosity-weighted mass function from Heckman et al. [2004, their Fig. 1], horizontal error bars a $\sim$0.2 dex uncertainty in the black hole mass). Darker lines show this for our full model; lighter lines show the full distribution (*dashed line*) and distribution with the same weighting and selection effects as Heckman et al. (2004; *solid line*) for a light-bulb or exponential light curve model of quasar evolution. *Right*: Distribution of active broad-line quasar masses (*solid lines*, where an object is a broad-line quasar if the observed quasar *B*-band luminosity is above a factor $f_{\rm BL} = 1$ of that of the host galaxy; dotted and dashed lines show the result if $f_{\rm BL} = 0.3$ or 3, respectively). The three rightmost curves show the prediction of the full model; the two left curves show the predictions of a light-bulb/exponential light curve model with a standard torus model (*descending curve*) and receding torus model (*peaked curve*) used to determine the broad-line fraction. [*See the electronic edition of the Supplement for a color version of this figure.*]

normalization of this curve and our full model prediction are not the same; for example, the predicted total absolute number of active $M_{\rm BH} > 10^9\ M_\odot$ quasars is higher in the full model than in the light-bulb or exponential models. Still, it is clear that these models produce too broad a distribution of active black hole masses, in disagreement with the observations. We could, of course, obtain an arbitrarily close agreement with the observations if we fit to the *distribution* of accretion rates, but such a model would recover a quasar lifetime and accretion rate distribution quite similar to ours, as is evident from the agreement between the predictions of our simulations and the observations. A purely empirical model of this type is considered by, for example, Merloni (2004), who finds that qualitative evolution in the quasar lifetime and antihierarchical black hole assembly similar to that predicted by our modeling is implied by the combination of quasar luminosity functions and the black hole mass function.

In the right panel of Figure 18, we show our predicted mass distribution for low-redshift, active broad-line quasars (*solid black lines*), where we estimate that an object is a broad-line quasar if the observed quasar *B*-band luminosity is above a factor $f_{\rm BL} = 1$ of that of the host galaxy; dotted and dashed lines show the result if $f_{\rm BL} = 0.3$ or 3, respectively, parameterizing the range of different observed samples. As discussed above, the range of $f_{\rm BL}$ shown can be, alternatively, thought of as a parameterization of uncertainty in the host galaxy gas fraction, if (in an observed sample) the sensitivity to seeing quasar broad lines against host galaxy contamination is known. Therefore, the location of the peak in the active broad-line black hole mass function can be used, just as the mean broad-line fraction versus luminosity, as a test of the typical gas fractions of bright quasar host galaxies and can constrain potential evolution in these gas fractions with redshift.

The prediction shown is testable but appears to be in good agreement with preliminary results for the distribution of active broad-line black hole masses from the SDSS (e.g., McLure & Dunlop 2004). The observations may show fewer low-mass black holes than we predict, but this is expected, as observed samples are likely incomplete at the low luminosities of these objects (even at the Eddington limit, a $10^5\ M_\odot$ black hole has magnitude $M_g \sim -16$). If, in our model, we were to consider instead a standard torus scenario for the definition of the broad-line phase, we would predict the same curve as that shown in the left half of the figure (*black dashed curve*; our prediction for the cumulative active black hole mass function). This is because the standard torus model predicts that a constant fraction of objects are broad-line quasars, regardless of mass or luminosity, thus giving identical distributions of type I and type II quasar masses. If we consider a luminosity-dependent or receding torus model, the prediction is nearly identical to the black line shown. This is because, as shown in Figure 16, our prediction for the broad-line fraction as a function of luminosity is similar to that of the receding torus model. The differences in the model predictions for the broad-line fraction as a function of luminosity do manifest in the prediction for the active broad-line black hole mass function, but the difference in these models is smaller than the $\sim$1 σ range from different values of $f_{\rm BL}$ shown. However, if we consider different models for the quasar light curve or lifetime, the predicted active broad-line mass function is quite different (as is the cumulative active black hole mass function).

We show the predictions of a light-bulb or exponential light curve model for quasar evolution in the the right panel of Figure 18, adopting either a standard torus model (*left descending curve*) or receding torus model (*left peaked curve*) to determine the broad-line fraction as a function of luminosity. For the standard torus model, this predicts that the broad-line mass function should trace the observed luminosity function, rising monotonically to lower black hole masses in power-law fashion (just as

seen in the light dashed line in the left half of the figure for the cumulative black hole mass function). For the receding torus model, the active black hole mass function shows a peak (because, at lower luminosities, there are more observed quasars, but a larger fraction of them are obscured). However, the location of this peak is at roughly an order of magnitude smaller black hole mass than for our prediction. This assumes a typical Eddington ratio of $\sim$1, which we have fitted to the cumulative black hole mass function; the peak in the broad-line active black hole mass function in these models could be shifted to larger black hole masses by assuming a smaller typical Eddington ratio, but this would only worsen the agreement with the cumulative black hole mass function of Heckman et al. (2004). Furthermore, a robust difference between the models is that the light-bulb or Eddington-limited/exponential models predict, for the standard torus case, no turnover in the active broad-line black hole mass function, and for the receding torus case, a broader distribution in active broad-line quasar black hole masses than is predicted in our modeling. Roughly, the lognormal width of this distribution in our model is $\sim$0.6 dex, whereas the light-bulb or exponential light curve models have a distribution with width $\sim$1.0 dex. As noted above, we obtain a similar prediction if we adopt our full obscuration model instead of the receding torus model here. A determination of the range of active, broad-line quasar masses can, therefore, constrain quasar lifetimes and light curves.

Our model makes an accurate prediction for the distribution of *active* black hole masses, even at $z \sim 0$, where our extrapolation of the luminosity function is uncertain. It is important to distinguish this from the predicted relic black hole mass distribution derived in § 6, which must account for all quasars, i.e., $\dot{n}(L_{\rm peak})$ integrated over redshift. We additionally find for broad-line quasars, as we expect from our prediction of the broad-line phase, that these objects are primarily radiating at large Eddington ratios, $l \sim 0.2-1$, but we address this in more detail in § 5. The success of this prediction serves not only to support our model but also implies that we can extrapolate to fairly low luminosities, even bright Seyfert systems at $z \sim 0$. This suggests that many of these systems, at least at the bright end, may be related to our assumed quasar evolution model, fueled by similar mechanisms and either exhibiting weak interactions among galaxies or relaxing from an earlier, brighter stage in their evolution. As we speculate in § 8, our description of self-regulated black hole growth may also be relevant to fainter Seyfert galaxies, even those that reside in apparently undisturbed galaxies.

5. THE DISTRIBUTION OF EDDINGTON RATIOS

In traditional models of quasar lifetimes and light curves, the Eddington ratio $l \equiv L/L_{\rm Edd}$ is generally assumed to be constant. Even complex models of the quasar population that allow for a wide range of Eddington ratios according to some probability distribution $P(l)$ implicitly associate a fixed Eddington ratio with each individual quasar and do not allow for $P(l)$ to depend on instantaneous luminosity or host system properties. However, this is a misleading assumption in the context of our model, as the Eddington ratio varies in a complicated manner over most of the quasar light curve. Furthermore, the integrated time at a given Eddington ratio is different in different systems, with more massive, higher peak luminosity systems spending more time at large ($l \sim 1$) Eddington ratios.

The probability of being at a given Eddington ratio should properly be thought of as a conditional joint distribution $P(l|L, L_{\rm peak})$ in both instantaneous and peak luminosity, just as the quasar "lifetime" is more properly a conditional distribution $t_Q(L|L_{\rm peak})$. Rather than adopting a uniform Eddington ratio or Eddington ratio distribution, empirical estimates must consider more detailed formulations such as the framework presented in Steed & Weinberg (2003), which allows for a conditional bivariate Eddington ratio distribution and can therefore incorporate these physically motivated dependencies and complications in deconvolving observations of the quasar luminosity function to determine, for example, Eddington ratio distributions, active black hole mass functions, and other physical quantities.

Figure 19 shows the distribution of Eddington ratios as a function of luminosity for the fiducial, Milky Way–like A3 simulation ($V_{\rm vir} = 160$ km s^{-1}). Over the course of the simulation, the system spends a roughly comparable amount of time at a wide range of Eddington ratios from $l \sim 0.001$ to 1. At high luminosities, $L > 10^{12}\ L_\odot$ for a system with $L_{\rm peak} \approx 10^{13}\ L_\odot$, the range of Eddington ratios is concentrated at high values $l \sim 0.5-1$, with some time spent at ratios as low as $l \sim 0.1$. Note, however, that the y-axis of the plot is scaled logarithmically, so the time spent at $l \sim 0.1$ in this luminosity interval is a factor of $\sim$5 smaller than the time spent at $l \gtrsim 0.5$. Considering lower luminosities $10^{11}\ L_\odot < L < 10^{12}\ L_\odot$, the distribution of Eddington ratios broadens down to $l \sim 0.01$. Going to lower luminosities still, $L < 10^{11}\ L_\odot$, the distribution broadens further, with comparable time spent at ratios as low as $l \sim 0.001$, and becomes somewhat bimodal. At large luminosities near $L_{\rm peak}$, the system is primarily in Eddington-limited or near-Eddington growth. However, as we consider lower luminosities, we include both early times when the black hole is growing efficiently (high l) and late or intermediate times when the black hole is more massive but the accretion rate falls (low l). As we go to lower luminosities, the *total* time spent in sub-Eddington states increasingly dominates the time spent at $l \sim 1$, although the time spent at any given value of l is fairly flat with log (l).

Roughly, at some luminosity L there is a constant probability of being in some logarithmic interval in l,

$$P(l|L, L_{\rm peak}) \sim \left[\log\left(\frac{L_{\rm peak}}{L}\right)\right]^{-1}, \quad \frac{L}{L_{\rm peak}} < l < 1, \qquad (30)$$

and $P(l|L, L_{\rm peak}) = 0$ otherwise. This is especially clear if we compare the distribution of Eddington ratios in each luminosity range obtained if we consider only times after the final merger of the black holes (*dotted histograms*). At the highest luminosities, the distribution is identical to that obtained previously, since all the time at these luminosities is during the final merger. However, as we move to lower luminosities, the characteristic l moves systematically lower, as we are seeing only the relaxation after the final blowout near $L_{\rm peak}$, with characteristic Eddington ratio $l = L/L_{\rm peak}$ at any given luminosity L. These trends are also clear if we consider the distribution of *instantaneous* black hole masses in each luminosity interval shown in the figure, which is trivially related to the Eddington ratio distribution at a given luminosity L as

$$M_{\rm BH} = M_0 \frac{L}{l L_{\rm Edd}(M_0)} = \frac{L t_{\rm S}}{l \epsilon_r c^2}. \qquad (31)$$

Of course, it is clear here that $M_{\rm BH} \approx M_{\rm BH}^f = 3 \times 10^8\ M_\odot$ if we consider only times after the final merger.

It has also been argued from observations of stellar black hole binaries that a transition between accretion states occurs at a critical Eddington ratio $\dot{m} \equiv \dot{M}/\dot{M}_{\rm Edd}$, from radiatively inefficient accretion flows at low accretion rates (e.g., Esin et al. 1997) to radiatively efficient accretion through a standard Shakura & Sunyaev (1973) disk. Although the critical Eddington ratio for

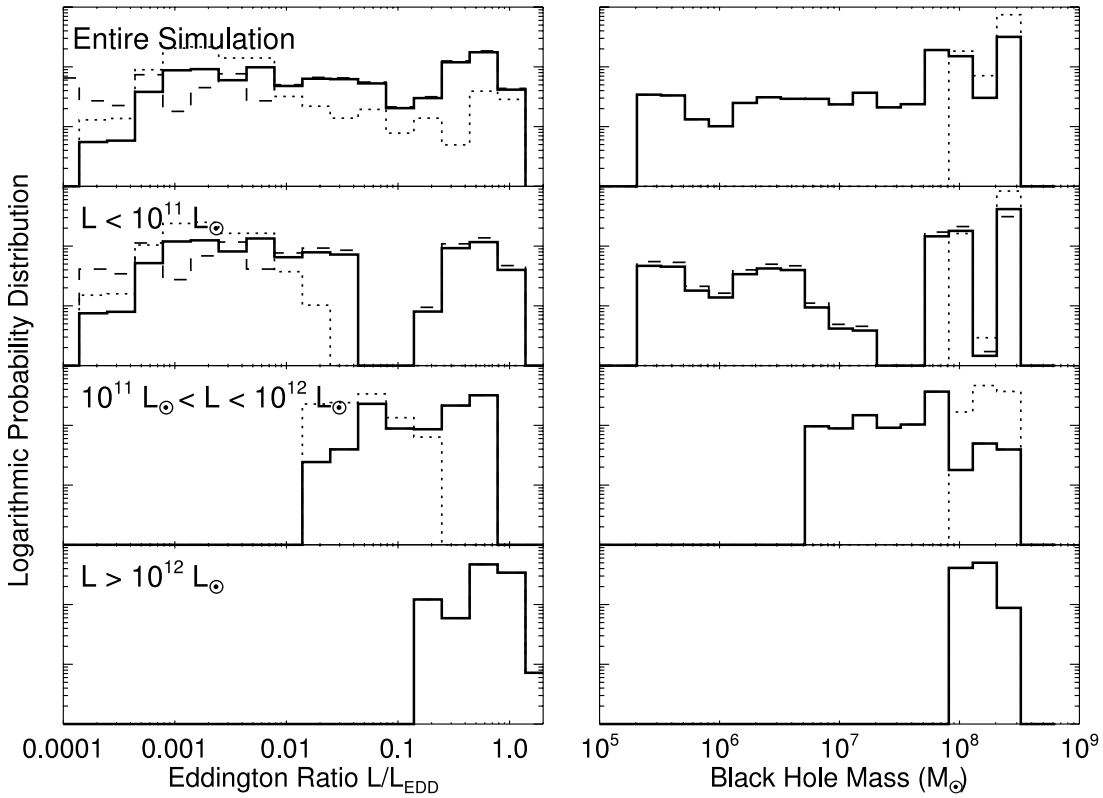

FIG. 19.—Distribution of Eddington ratios (*left*) and instantaneous black hole mass (*right*) as a function of quasar bolometric luminosity for our fiducial Milky Way–like A3 simulation, with $V_{\rm vir} = 160$ km s^{-1} and $L_{\rm peak} \sim 5 \times 10^{13}\ L_{\odot}$. The trend of an increasingly narrow Eddington ratio and mass distribution (concentrated at higher values) with increasing luminosity is clear. The result of applying an ADAF-type radiative efficiency correction at low accretion rates is shown (*dashed line*), as well as the result of considering only times after the final merger, with $M_{\rm BH} \sim M_{\rm BH}^{f}$ (*dotted line*).

supermassive black holes is uncertain, observations of black hole binaries (Maccarone 2003) as well as theoretical extensions of accretion models (e.g., Meyer et al. 2000) suggest $\dot{m}_{\rm crit} \sim 0.01$. We can examine whether this has a large impact on our predictions for the luminosity function and $\dot{n}(L_{\rm peak})$ distribution by determining whether the distribution of Eddington ratios is significantly changed by such a correction. Because we assume a constant radiative efficiency $L = \epsilon_r \dot{M} c^2$ with $\epsilon_r = 0.1$, we account for this effect by multiplying the simulation luminosity at all times by an additional "efficiency factor" $f_{\rm eff}$ that depends on the Eddington ratio $l = L/L_{\rm Edd}$,

$$f_{\rm eff} = \begin{cases} 1 & l > 0.01, \\ 100\, l & l \leq 0.01. \end{cases} \tag{32}$$

This choice for the efficiency factor follows from ADAF models (Narayan & Yi 1995) and ensures that the radiative efficiency is continuous at the critical Eddington ratio $l_{\rm crit} = 0.01$. Applying this correction and then examining the distribution of Eddington ratios as a function of luminosity (Fig. 19, *dashed histograms*), we see that the distribution of Eddington ratios is essentially identical, with only a slightly higher probability of observing extremely low Eddington ratios $l \lesssim 0.001$. Of course, our modeling of accretion processes does not allow us to accurately describe ADAF-like accretion at these low Eddington ratios, but such low values are not relevant for the observed luminosity functions and quantities with which we make our comparisons. This implies that such a transition in the radiative efficiency with accretion rate should not alter our conclusions regarding the luminosity function and the $\dot{n}(L_{\rm peak})$ distribution [essentially, the corrections are important only at luminosities well below those relevant in constructing the observed luminosity functions; see also Hopkins et al. 2005b for a calculation of the effects of such a correction on the fitted quasar lifetime and $\dot{n}(L_{\rm peak})$ distributions, which leads to the same conclusion].

Despite the broad range of Eddington ratios in the simulations, this entire distribution is unlikely to be observable in many samples. The effect of this can be predicted based on the behavior seen in Figure 19. For example, we consider the distribution of Eddington ratios that would be observed if the *B*-band luminosity $L_{B,\,\rm obs} \geq 10^{11}\ L_{\odot}$, comparable to the selection limits at high redshift of many optical quasar samples. As expected from the change in l with luminosity, this restricts the observed range of Eddington ratios to large values $l \sim 0.1-1$, in good agreement with the range of Eddington ratios actually observed in such samples. Essentially, it has reduced the observed range to a bolometric luminosity $L \gtrsim 10^{12}\ L_{\odot}$ in the case shown, giving a distribution similar to that seen in the bottom panels of the figure.

We compare our predicted distribution of Eddington ratios to observations in Figure 20. Using the distribution of peak luminosities $\dot{n}(L_{\rm peak})$ determined from the luminosity function, we can integrate over all luminosities to infer the observed Eddington ratio distribution,

$$P(l) \propto \int d\log L \int d\log L_{\rm peak} \times P(l\,|\,L,\,L_{\rm peak}) \frac{dt(L,\,L_{\rm peak})}{d\log L} \dot{n}(L_{\rm peak}). \tag{33}$$

As our estimate of $P(l|L,\,L_{\rm peak})$ above is rough, we do this by binning in $L_{\rm peak}$ and averaging the binned $P(l\,|\,L,\,L_{\rm peak})\,dt/d\log L$ for each simulation in the range of $L_{\rm peak}$, then weighting by

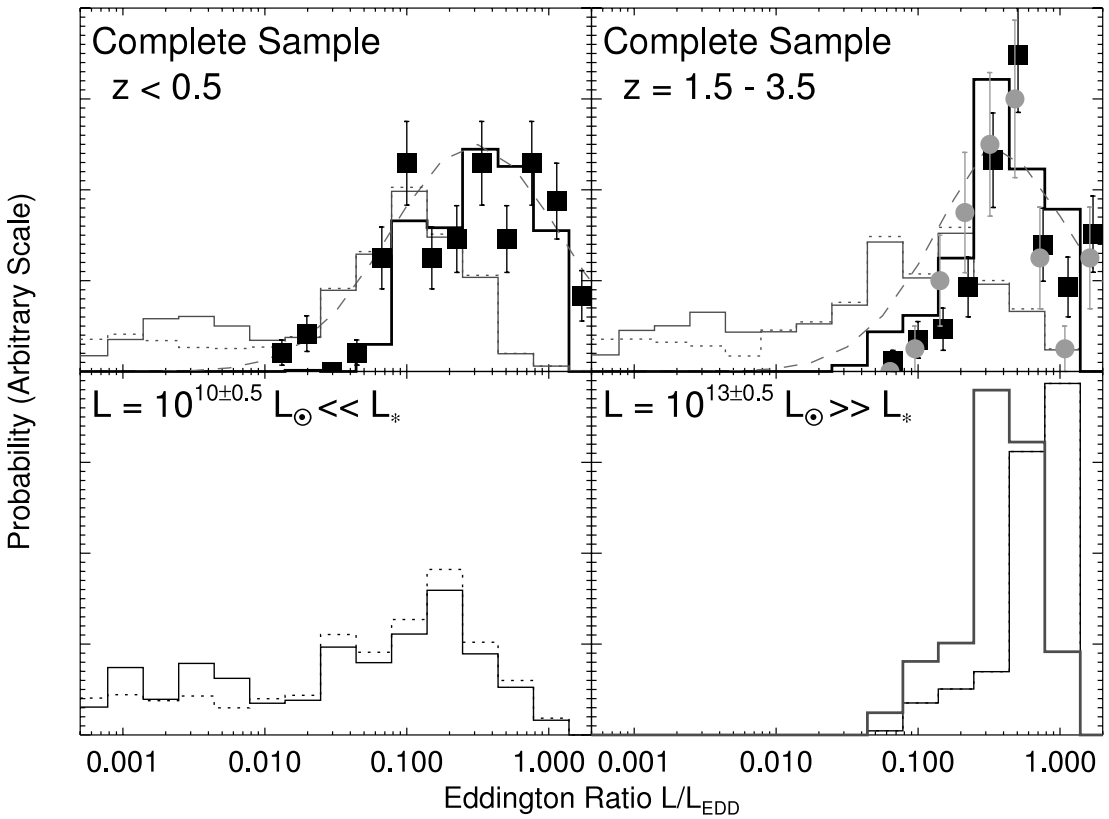

FIG. 20.—Predicted distribution of Eddington ratios based on the luminosity function and the quasar evolution in our simulations, in two redshift intervals $z < 0.5$ (*top left*) and $1.5 < z < 3.5$ (*top right*). The observed distributions for radio-loud (*squares*) and radio-quiet (*circles*) quasars are shown from Vestergaard (2004) with Poisson errors. Darker histograms show the predicted distribution given the same minimum observed luminosity as the observed sample. Lighter histograms show the predicted distributions for a sample extending to arbitrarily faint luminosities; dotted lines show the same, with the ADAF correction of § 5 applied at low accretion rates. Dashed lines show the prediction for a fixed (luminosity independent) Eddington ratio distribution in a light-bulb or exponential light curve model, fitted to the $z < 0.5$ data and used to predict the $1.5 < z < 3.5$ Eddington ratio distribution given the observational luminosity limit. Bottom panels show the predicted distributions for $z \lesssim 1$ in two luminosity intervals, above and below the break luminosity in the observed luminosity function (lighter histogram in bottom right panel corresponds to an observed [attenuated] B-band luminosity $L_{B,\,\rm obs} > 10^{11}\ L_{\odot}$). [*See the electronic edition of the Supplement for a color version of this figure.*]

$\dot{n}(L_{\rm peak})$ and integrating. We consider both the entire distribution that would be observed in the absence of selection effects (*lighter histograms*) and the distribution observed demanding a B-band luminosity above some reference value, $L_{B,\,\rm obs} > L_{\rm min}$ (*darker histograms*). The results are shown for redshifts $z < 0.5$ and $z = 1.5-3.5$, along with the observed distribution from Vestergaard (2004), with assumed Poisson errors. The observations should be compared to the darker histograms, which have luminosity thresholds $L = 10^{10}$ and $10^{11}\ L_{\odot}$ for $z < 0.5$ and 1.5–3.5, respectively, corresponding approximately to the minimum observable luminosities in the observed samples in each redshift interval.

The agreement is good, given the observational uncertainties, and it suggests that the observed Eddington ratio distribution can be related to the nontrivial nature of quasar lifetimes and light curves we model, rather than some arbitrary distribution of fixed l across sources. However, the selection effects in the observed samples are quite significant—the complete distribution of Eddington ratios is similar in both samples, implying that the difference in the observed Eddington ratio distribution is primarily a consequence of the higher luminosity limit in the observed samples—and a more detailed test of this prediction requires fainter samples.

Still, there is a systematic offset in the observed samples at $z < 0.5$ and 1.5–3.5 that is not due to selection effects. At progressively lower redshifts, more quasars with luminosities further below the break in the luminosity function are observed, and therefore the observed Eddington ratio is broadened to lower Eddington ratios $l \sim 0.1$, whereas at high redshift the distribution is more peaked at slightly higher Eddington ratios. This difference, although not dramatic, is a prediction of our model not captured in light-bulb or fixed-Eddington-ratio models, even when allowing for a distribution of Eddington ratios, if such a distribution is static. We demonstrate this by fitting the low-redshift Eddington ratio distribution to a Gaussian (Fig. 20, *top left, dashed line*) and then assuming that this distribution of accretion rates is unchanged with redshift, giving (after applying the same selection effects that yield the darker histograms plotted) the dashed line in the top right panel. Although the agreement may appear reasonable, the difference is significant; such a fit overpredicts the fraction of high-redshift objects at Eddington ratios $\lesssim 0.1$ and underpredicts the fraction at ~ 0.3, giving a somewhat poor fit overall ($\chi^2/\nu = 2.7$, but with typical $\gtrsim 3\ \sigma$ overpredictions for Eddington ratios $\lesssim 0.1$).

Furthermore, without being modified to allow for a distribution of Eddington ratios, such models are clearly inconsistent with the observations, as they would predict a single, constant Eddington ratio. However, models that fit the observed evolution in the quasar luminosity function with a nonstatic distribution of accretion rates do recover the broadening of the Eddington ratio distribution at low redshift, as long as strong evolution in the distribution of accretion rates for systems of a given black hole mass is not allowed (Steed & Weinberg 2003), giving a qualitatively similar picture of the evolution we model. Regardless of the evolution in accretion rates, an advantage of our modeling is

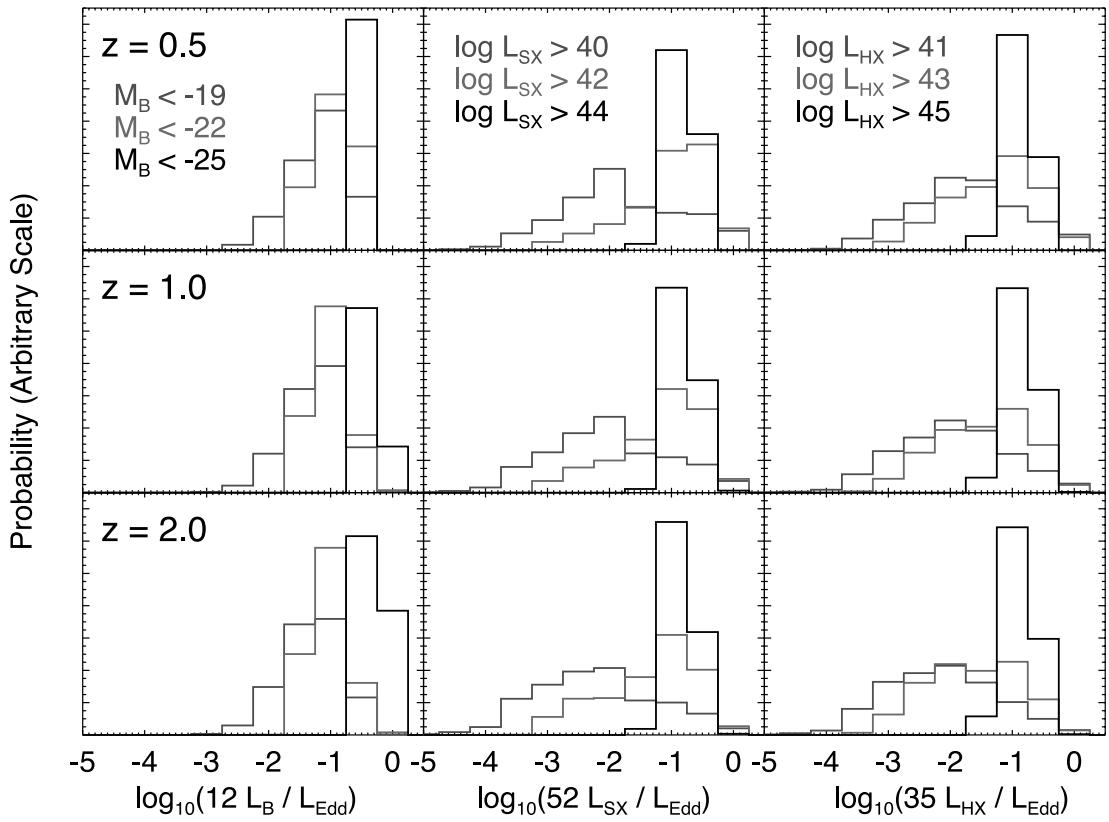

FIG. 21.—Predicted distribution of Eddington ratios based on the luminosity function and the quasar evolution in our simulations, at three redshifts $z = 0.5$ (*top*), $z = 1.0$ (*middle*), and $z = 2.0$ (*bottom*). The inferred distribution of Eddington ratios, adopting a constant bolometric correction from the observed (attenuated) luminosity in each of three bands, is shown, i.e., assuming $L = 12L_B^{\rm obs}$ (4400 Å; *left*), $L = 52L_{\rm SX}^{\rm obs}$ (0.5–2 keV; *middle*), and $L = 35L_{\rm HX}^{\rm obs}$ (2–10 keV; *right*). For each wave band, results are shown for three reference luminosities. In *B* band, $M_B < -19$ (*leftmost light line*), -22 (*rightmost light line*), and -25 (*dark line*). In soft X-rays, $\log [L_{\rm SX}(\rm ergs\ s^{-1})] > 40$ (*leftmost light line*), 42 (*rightmost light line*), and 44 (*dark line*). In hard X-rays, $\log [L_{\rm HX}(\rm ergs\ s^{-1})] > 41$ (*leftmost light line*), 43 (*rightmost light line*), and 45 (*dark line*). [*See the electronic edition of the Supplement for a color version of this figure.*]

that it provides a physically motivated predicted distribution of accretion rates, as opposed to being forced to adopt the distribution of accretion rates from observational input (which can be, as demonstrated in the figure, significantly biased by observational selection effects). The dotted histograms show the distribution if we apply our ADAF correction to the intrinsic distribution and demonstrate that this does not significantly change the result. We note that our model for black hole accretion employs the Eddington limit as a maximum accretion rate; if we remove this restriction, we find that the simulations spend some small but nonnegligible time with $l \sim 1-2$, which is also consistent with the observations.

Furthermore, we can make a prediction of this model that can be falsified, namely, that the Eddington ratio distribution at luminosities well below the break in the luminosity function should be broader and extend to lower values than the distribution at luminosities above the break luminosity. We quantify this in the bottom panels of Figure 20, for the distribution at low redshifts $z \lesssim 1$. Here we consider two bins in luminosity, $L = 10^{9.5}-10^{10.5}$ and $10^{12.5}-10^{13.5}\,L_\odot$, for redshifts where the break in the luminosity function is at approximately $L \sim 10^{11}-10^{12}\,L_\odot$. Clearly, the distribution is broader and extends to lower Eddington ratios in the former luminosity interval, whereas in the latter it is strongly peaked about $l \sim 0.2-1$, for both the complete distribution (*darker histograms*) and that with $L_{B,\,\rm obs} \geq 10^{11}\,L_\odot$ (*lighter histograms*). The distribution obtained applying the ADAF correction described above is shown as dotted histograms. Despite the fact that the Eddington ratio distribution at low luminosities will be strongly biased by selection effects, a reasonably complete sample should be able to test this prediction, at least qualitatively.

We illustrate the effects of changing observed wave band, redshift, and luminosity thresholds on the observed Eddington ratio distribution in Figure 21. Here we plot the predicted distribution of Eddington ratios determined as in Figure 20, from our fitted $\dot{n}(L_{\rm peak})$ distribution at each redshift and the distribution of Eddington ratios as a function of instantaneous and peak luminosity in each of our simulations (specifically, these are drawn from the Monte Carlo realizations of the quasar population described in § 8). We show the predictions at three redshifts $z = 0.5$ (*top panels*), $z = 1.0$ (*middle*), and $z = 2.0$ (*bottom*). For each redshift, results are shown in three wave bands, and with three reference luminosities. In *B* band, we require $M_B < -19$ (*leftmost light line*), $M_B < -22$ (*rightmost light line*), and $M_B < -25$ (*dark line*). In soft X-rays, $\log [L_{\rm SX}(\rm ergs\ s^{-1})] > 40$ (*leftmost light line*), 42 (*rightmost light line*), 44 (*dark line*). In hard X-rays, $\log[\,L_{\rm HX}(\rm ergs\ s^{-1})] > 41$ (*leftmost light line*), 43 (*rightmost light line*), 45 (*dark line*). The observationally inferred distribution of Eddington ratios at each redshift is loosely estimated by adopting a constant bolometric correction from the observed (attenuated) luminosity in each of three bands shown, i.e., assuming $L = 12L_B^{\rm obs}$ (4400 Å; *left*), $L = 52L_{\rm SX}^{\rm obs}$ (0.5–2 keV; *middle*), and $L = 35L_{\rm HX}^{\rm obs}$ (2–10 keV; *right*). This follows common practice in many observational estimates of the Eddington ratio distribution and allows for the effects of attenuation, but we caution that it can be misleading.

If we instead use the luminosity-dependent bolometric corrections of Marconi et al. (2004) that we adopt throughout, even

given that we are calculating from the observed (attenuated) luminosities, we do not see the large population of highly sub-Eddington (Eddington ratios $\lesssim 10^{-3}$) quasars in soft and hard X-ray samples with low luminosity thresholds. This is because these are actually reasonably high Eddington ratio quasars, but our bolometric corrections imply that a larger fraction of the bolometric luminosity is radiated in the X-ray at low bolometric luminosity, meaning that assuming a constant bolometric correction will underestimate the Eddington ratios of high bolometric luminosity sources. Regardless, the figure illustrates both the importance of different wavelengths (i.e., the ability to observe more low Eddington ratio sources in X-ray as compared to optical samples) and luminosity/magnitude limits on the inferred distribution of Eddington ratios. For example, even for relatively deep B-band quasar samples complete to $M_B < -23$ (i.e., complete to essentially all objects traditionally classified as having "quasar-like" luminosities), the expected observed Eddington ratio distribution at $z \sim 0.5-2$ is quite sharply peaked about $\sim$0.1–0.3, in good agreement with recent observational results (Kollmeier et al. 2006).

We do not compare to the $z = 0$ distribution of black hole accretion rates, as this is dominated by objects at extremely low Eddington ratios $l \sim 10^{-5}$ to 10^{-4} (e.g., Ho 2002; Marchesini et al. 2004; Jester 2005), which are well below the range we model and are not likely to be driven by merger activity (many of these objects are quiescent, low-luminosity Seyferts in normal spiral galaxy hosts); furthermore, many of these objects are not accreting at the Bondi rate (Fabian & Canizares 1988; Blandford & Begelman 1999; Di Matteo et al. 2000, 2001; Narayan et al. 2000; Quataert & Gruzinov 2000; Loewenstein et al. 2001; Bower et al. 2003), clearly showing that our simulations must incorporate more sophisticated models for accretion in quiescent, low-luminosity states (when gravitational torques cannot provide a mechanism to drive large amounts of gas to the central regions of the galaxy) in order to describe such phases.

However, it has been suggested that the rapid blowout phase and subsequent decay in accretion rates seen in our simulations, coupled with spectral modeling of radiatively inefficient accretion modes, can explain the apparently bimodal distribution of low-redshift accretion rates (Cao & Xu 2005). Moreover, present-day, relaxed ellipticals are observed to have mass accretion rates of $\sim 10^{-4}$ implying a long relaxation time at moderate and low accretion rates, qualitatively similar to that seen after the blowout in our modeling (Hopkins et al. 2006a). A pure exponential decay in accretion rate after the peak quasar phase would give $\dot{m} = \dot{M}/\dot{M}_{\rm Edd} \sim \exp(-t_{\rm H}/t_Q)$ at present, where $t_{\rm H}$ is the Hubble time and t_Q is the quasar lifetime of order, for example, the Salpeter time $t_{\rm S} = 4 \times 10^7$ yr, yielding an unreasonably low expected accretion rate $\dot{m} \sim 10^{-145}$. Even assuming an order-of-magnitude larger quasar lifetime, this gives $\dot{m} \sim 10^{-15}$, far below observed values, implying that regardless of the fueling mechanisms at low luminosities, the basic key point of our modeling must be true to some extent, namely, that quasars spend long times relaxing at moderate to low Eddington ratios.

6. THE MASS FUNCTION OF RELIC SUPERMASSIVE BLACK HOLES FROM QUASARS

From the $M_{\rm BH}$-σ relation and other host galaxy–black hole scalings, estimates of bulge and spheroid velocity dispersions have been used to determine the total mass density ($\rho_{\rm BH}$) and mass distribution of local, primarily inactive supermassive black holes (e.g., Salucci et al. 1999; Marconi & Salvati 2002; Yu & Tremaine 2002; Ferrarese 2002; Aller & Richstone 2002; Marconi et al. 2004; Shankar et al. 2004). These estimates, along with others based on X-ray background synthesis (e.g., Fabian & Iwasawa 1999; Elvis et al. 2002), have compared these quantities to those expected based on the mass distribution of "relic" black holes grown in quasars. It appears that most, and perhaps nearly all, of the present-day black hole mass density was accumulated in bright quasar phases, and the $M_{\rm BH}$-σ and $M_{\rm BH}$-$L_{\rm bulge}$ correlations yield estimates of the local mass function in good agreement with those from hard X-ray AGN luminosity functions (Marconi et al. 2004).

However, this modeling is dependent on several assumptions. Namely, the average radiative efficiency ϵ_r, Eddington ratio l, and average quasar lifetime t_Q are generally taken to be constants and either input into the model or constrained by demanding agreement with the local mass function. In our simulations, we find the quasar lifetime and Eddington ratio to be complex functions of both luminosity and host system properties (as opposed to being constants). We also find that quasars spend a large fraction of their lives in obscured growth phases, suggesting some mass gain outside of the bright quasar phase. It is thus of interest to determine the relic black hole mass function expected from our model for quasar evolution.

Using our estimate for the birthrate of quasars with a given peak luminosity at a particular redshift, $\dot{n}(L_{\rm peak})$, obtained from the luminosity function in § 3.2, we can estimate the total number density of relic quasars accumulated by a particular redshift that were born with a given $L_{\rm peak}$ (per logarithmic interval in $L_{\rm peak}$) from

$$n(L_{\rm peak}) = \int \dot{n}(L_{\rm peak})\, dt = \int \frac{\dot{n}(L_{\rm peak}, z)\, dz}{(1+z)H(z)}. \quad (34)$$

By redshift $z = 0$, most of these quasars will be "dead," with only a small residual fraction having been activated in the recent past.

Using our lognormal form for $\dot{n}(L_{\rm peak})$, with normalization $\dot{n}_*$ and dispersion σ_* held constant and only the median $L_* = L_*^0 \exp(k_L \tau)$ evolving with redshift, this integral can be evaluated numerically to give the space density of relic quasars $n(L_{\rm peak})$. Finally, we use $M_{\rm BH}^f(L_{\rm peak})$, roughly the Eddington mass of the given peak luminosity (but determined more precisely in § 2.4) to convert from $dn(L_{\rm peak})/d\log L_{\rm peak}$ to $dn(M_{\rm BH})/d\log M_{\rm BH}$. This formulation implicitly assumes that black holes do not undergo subsequent mergers after the initial quasar-producing event. However, this effect should be small (a factor $\lesssim$2) as subsequent mergers would be dry (gas-poor). We explicitly calculate the effects of dry mergers on the spheroid mass function (essentially a rescaling of the black hole mass function calculated here) in Hopkins et al. (2006b) and show that this is a small effect (significantly less than the uncertainties owing to our fit to the quasar luminosity function) even assuming the maximum dry merger rates of, e.g., van Dokkum (2005).

This mass function can then be integrated over $dM_{\rm BH}$ to give the total present-day black hole mass density, $\rho_{\rm BH}$. Neglecting temporarily the small corrections to $M_{\rm BH}^f(L_{\rm peak})$ from § 2.4, we expect

$$M_{\rm BH}^f \approx M_{\rm Edd}(L_{\rm peak}) = \frac{L_{\rm peak} t_{\rm S}}{\epsilon_r c^2}, \quad (35)$$

where $t_{\rm S}/\epsilon_r c^2 \approx 2.95 \times 10^{-5}\ M_\odot/L_\odot$, so therefore,

$$\rho_{\rm BH} = \frac{t_{\rm S}}{\epsilon_r c^2} \int L_{\rm peak} n(L_{\rm peak})\, d\log L_{\rm peak}. \quad (36)$$

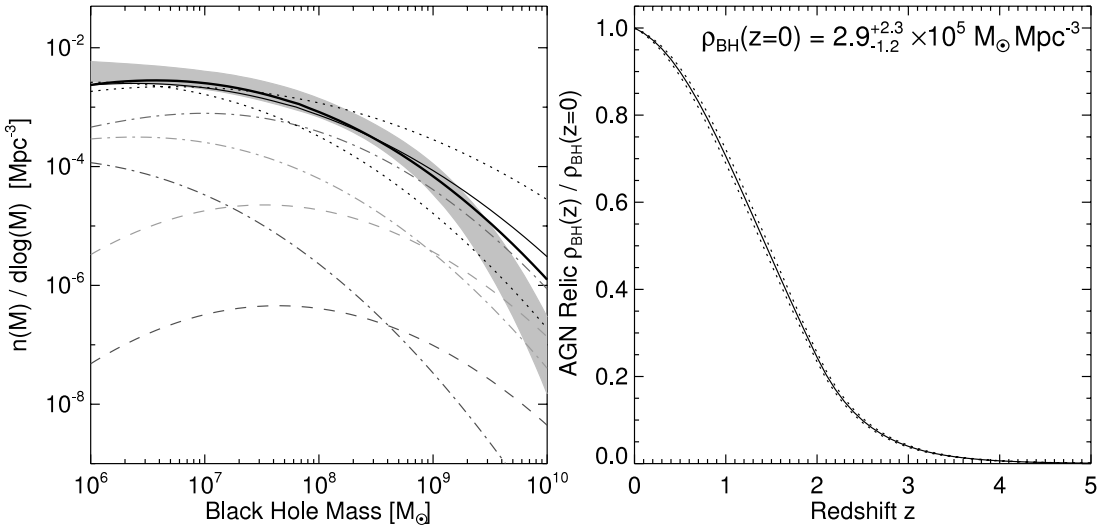

FIG. 22.—*Right*: Total predicted quasar relic black hole mass density and evolution of the fractional black hole mass density with redshift. Dotted lines show the difference resulting from 1 σ deviation in fitted $\dot{n}(L_{\rm peak})$ from the luminosity function. *Left*: Predicted present $z = 0$ relic mass function (*thick black line*), for comparison with the 1 σ range (*shaded region*) of the inferred supermassive black hole mass function from Marconi et al. (2004). Also shown are the results given 1 σ errors in the fitted $\dot{n}(L_{\rm peak})$ distribution (*dotted lines*), or ignoring the small corrections to $M^f_{\rm BH}(L_{\rm peak})$ from § 2.4 (*thin black line*). Dot-dashed lines show the predicted mass function at $z = 1.5$, 3.0, and 5.0 (*top to bottom*). The extensions to $z > 2$ includes the turnover (PPLE form) in the quasar space density above $z = 2$ from high-redshift luminosity functions described in § 3.2, except for the upper and lower dashed lines, which use the PDE form. [*See the electronic edition of the Supplement for a color version of this figure.*]

This can be combined with the integral over redshift for $n(L_{\rm peak})$, giving at each z a pure Gaussian integral over $\log(L_{\rm peak})$ in the form

$$\rho_{\rm BH} = \frac{L^0_* t_{\rm S}}{\epsilon_r c^2}\frac{\dot{n}_*}{H_0} e^{(1/2)(\sigma_*\ln 10)^2}\int \frac{e^{k_L\tau}\,dz}{(1+z)\hat{H}(z)} = \frac{L^0_* t_{\rm S}}{k_L\epsilon_r c^2}\frac{\dot{n}_*}{H_0} e^{(1/2)(\sigma_*\ln 10)^2}\left(e^{k_L\tau_f} - e^{k_L\tau}\right), \quad (37)$$

where $\hat{H}(z) \equiv H(z)/H_0$ and τ_f is the fractional look-back time at some upper limit. We must modify this integral above $z \sim 2$ to account for the decreasing space density of bright quasars, applying either our density or peak-luminosity evolution turnover from § 3.2, but quasars at these high redshifts contribute only a small fraction to the present-day density. Thus, in this formulation, the evolution of the total supermassive black hole mass density, i.e., $\rho_{\rm BH}(z)/\rho_{\rm BH}(z = 0)$, is given approximately by the dimensionless integral above and depends only on how L_* evolves, essentially the rate at which the break in the quasar luminosity function shifts. Although this is not strictly true if we include corrections to $M^f_{\rm BH}(L_{\rm peak})$ based on $L_{\rm peak}$, the difference is small and this behavior is essentially preserved. Note that the total supermassive black hole mass density is independent of corrections from subsequent dry mergers, which (being gas-poor) conserve total black hole mass.

Figure 22 shows our prediction for the mass distribution of supermassive black holes, as well as the total density $\rho_{\rm BH}$ and its evolution with redshift. We find a total relic black hole mass density of $\rho_{\rm BH} = 2.9^{+2.3}_{-1.2} \times 10^5\ M_\odot\ {\rm Mpc}^{-3}$, in agreement with the observational estimate of $\rho_{\rm BH} = (2.9 \pm 0.5) h^2_{0.7} \times 10^5\ M_\odot\ {\rm Mpc}^{-3}$ by Yu & Tremaine (2002) ($h_{0.7} \equiv H_0/70$ km s^{-1} Mpc^{-1}; their result is converted from $h = 0.65$) and within 1 σ of the value $\rho_{\rm BH} = 4.6^{+1.9}_{-1.4} h^2_{0.7} \times 10^5\ M_\odot\ {\rm Mpc}^{-3}$ of Marconi et al. (2004), based on the observations of Marzke et al. (1994), Kochanek et al. (2001), Nakamura et al. (2003), Bernardi et al. (2003), and Sheth et al. (2003). The fractional evolution of $\rho_{\rm BH}$ with redshift is quite well constrained, and we find, as with previous estimates, that most of the present-day black hole mass density accumulates at moderate to low redshifts $z \approx 0.5-2.5$. The 1 σ errors are shown as dotted lines in the figure and are close to our best-fit estimate, as we have demonstrated that this quantity depends only on k_L, the rate of evolution of the break in the luminosity function with redshift, which is fairly well constrained by observations (from our fitting to the luminosity functions, $k_L = 5.61 \pm 0.28$). The difference in $\rho_{\rm BH}$ if we include or neglect the small corrections to $M^f_{\rm BH}$ is negligible compared to our errors (~5%).

Our estimate for the relic black hole mass distribution (*thick black line*) also agrees well with observational estimates, with all observations within the range allowed by the 1 σ errors of our fitting to the luminosity function (*dotted lines*). The observations shown are again from Marconi et al. (2004), based on the combination of observations by Marzke et al. (1994), Kochanek et al. (2001), Nakamura et al. (2003), Bernardi et al. (2003), and Sheth et al. (2003). The high-mass end of the black hole mass function $M_{\rm BH} > 10^9\ M_\odot$ is relatively sensitive to whether or not we apply the $M^f_{\rm BH}(L_{\rm peak})$ corrections of § 2.4, instead of taking $M^f_{\rm BH} = M_{\rm Edd}(L_{\rm peak})$ (*thin line*), as well as to our fitting procedure. However, the agreement is still good, and this is also where the observational estimates of the mass distribution are most uncertain, as they are generally extrapolated to these masses and are sensitive to the assumed intrinsic dispersions in the $M_{\rm BH}$-σ and $M_{\rm BH}$-$L_{\rm bulge}$ relations (Yu & Tremaine 2002).

If instead we adopt a light-bulb, constant-Eddington-ratio, or exponential light curve model for quasar evolution, we would have $M^f_{\rm BH} \propto L_{\rm peak}$, and thus the prediction would be similar to the thin black line shown, a somewhat worse fit at high black hole masses. However, in these models this can be remedied by adjusting the typical Eddington ratios, quasar lifetimes, or radiative efficiencies. We do not show the range of predictions of these models for the relic supermassive black hole mass function, as they have been examined in detail previously (e.g., Salucci et al. 1999; Marconi & Salvati 2002; Yu & Tremaine 2002; Ferrarese 2002; Aller & Richstone 2002; Marconi et al. 2004; Shankar et al. 2004). These works demonstrate that the observed quasar luminosity functions are consistent with the relic supermassive black hole mass function, given typical radiative efficiencies $\epsilon_r \sim 0.1$ and Eddington ratios ~0.5–1.0, and that most of the mass of black

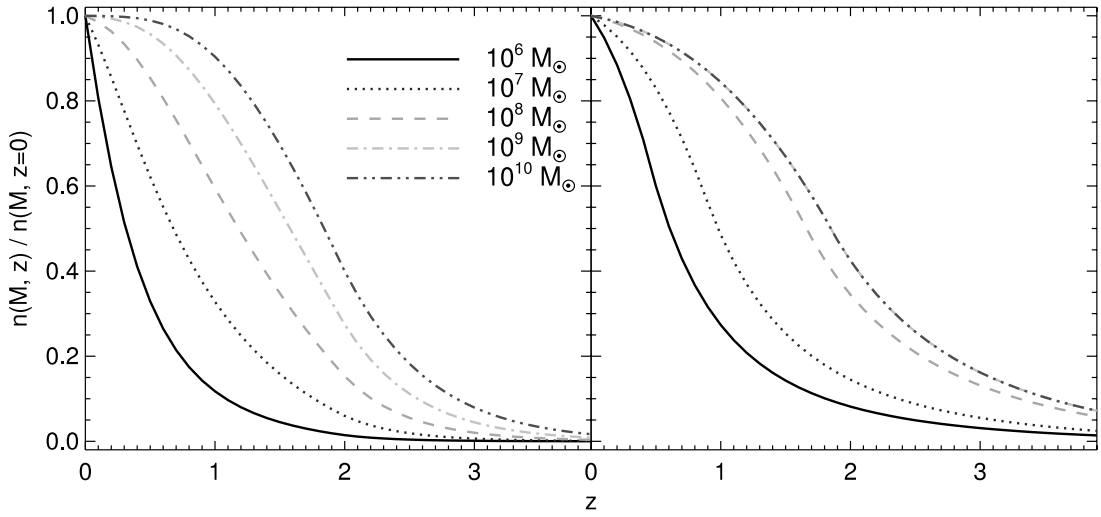

FIG. 23.—Fractional number density $n(M, z)/n(M, z = 0)$ of black holes of a given mass as a function of redshift, for several different black hole masses as shown. For $z > 2$ this includes the turnover (PDE form) in the quasar space density above $z = 2$ from high-redshift luminosity functions described in § 3.2. Left panel shows the results using our full model of quasar lifetimes, right panel assuming a light-bulb or exponential (constant Eddington ratio) light curve model. The dot-dashed (10^9 $M_\odot$) and double-dot–dashed (10^{10} $M_\odot$) curves are nearly identical in the right panel. [*See the electronic edition of the Supplement for a color version of this figure.*]

holes is accumulated in bright, observed phases, or else the required radiative efficiency would violate theoretical limits.

That our model of quasar lifetimes and obscuration reproduces the observed $z = 0$ supermassive black hole mass function explicitly demonstrates that we are consistent with these constraints. By choice, the radiative efficiency in our simulations is $\epsilon_r = 0.1$ and accretion rates are not allowed to exceed Eddington. As noted in § 4, most of the black hole mass is accumulated and radiant energy released in the final, blowout phase of quasar evolution, and here our black hole mass function and cumulative black hole mass density demonstrate that our modeling is consistent with integrated energy and mass arguments such as that of Soltan (1982), despite the fact that quasars spend more *time* in obscured phases than they do in bright optical quasar phases. In fact, comparison of our predicted total black hole mass density with estimates from the $z = 0$ black hole mass distribution allows some latitude for significant mass gain in radiatively inefficient growth or black holes in small, disky spheroids, although we emphasize that this is mainly because the uncertainty in our prediction is large; it is not inherent or necessary in our modeling.

The antihierarchical nature of black hole formation, where less massive black holes are formed at lower redshift, is reflected in our modeling by the shift of the break in the quasar luminosity function to lower values with decreasing redshift. This can be seen in Figure 22, where the black hole mass distributions are shown at redshifts $z = 1.5$, 3.0, and 5.0, assuming either PPLE or PDE for $z > 2$ (*dot-dashed and dashed lines, respectively*). While the choice for the turnover in the $z > 2$ quasar density matters little for the $z < 2$ black hole mass functions, the low-$M_{\rm BH}$ distribution at high redshift [where observations do not constrain $\dot{m}(L_{\rm peak})$ well] is quite different between the two models. Figure 23 plots the fractional number density of black holes of a given mass as a function of redshift, i.e., $n(M, z)/n(M, z = 0)$, where $n(M) = dn/d\log(M)$ is just the number density at mass M. This figure demonstrates that higher mass black holes originated over a larger range of redshifts and that they mostly formed at higher redshift, compared to lower mass black holes.

The right panel of Figure 23 compares our prediction to that of a light-bulb or exponential light curve model for quasar lifetimes. In these models, the antihierarchical nature of black hole assembly is dramatically suppressed. At the high-mass end, there is no measurable difference in the distribution of formation redshifts (i.e., the $M_{\rm BH} = 10^9$ and 10^{10} $M_\odot$ curves are indistinguishable), and there is little change in the formation times at $M_{\rm BH} = 10^8$ $M_\odot$. The shift in formation redshift at lower masses, although significant, is smaller than that predicted in our model. If spheroids and black holes are produced together, as in our picture, these models of the quasar lifetime would imply that spheroids of masses $M_{\rm vir} \sim 10^{11}$–10^{13} $M_\odot$ all formed over nearly identical ranges of redshifts, which is inconsistent with many observations indicating antihierarchical growth of the red, elliptical galaxy population (e.g., Treu et al. 2001, 2002, 2005; van Dokkum et al. 2001; van Dokkum & Stanford 2003; Gebhardt et al. 2003; Rusin et al. 2003; van de Ven et al. 2003; Wuyts et al. 2004; Holden et al. 2005; van der Wel et al. 2005; di Serego Alighieri et al. 2005; Nelan et al. 2005). Implications of our model for the red galaxy sequence are considered in Hopkins et al. (2006b), where we show that this weaker antihierarchical black hole (and correspondingly, spheroid) evolution is inconsistent with observed luminosity functions, color-magnitude relations, and mass-to-light ratios of elliptical galaxies.

Our modeling reproduces the observed total density and mass distribution of supermassive black holes at $z = 0$ with black holes accreting at the canonical efficiency $\epsilon_r = 0.1$ expected for efficient accretion through a Shakura & Sunyaev (1973) disk. Presumably, a large change in ϵ_r would give a significantly different relation between peak luminosity and black hole mass (for the same $L_{\rm peak}$, $M_{\rm BH}^f \propto 1/\epsilon_r$), and thus if the quasar lifetime remained similar as a function of peak luminosity, this would translate to a shift in the black hole mass function. The long obscured stage in black hole evolution does not generate problems in reproducing the black hole mass density, and the final phases of growth are still in bright optical quasar stages. However, a large Compton-thick population of black holes at all luminosities (or even at some range of luminosities at or above the break in the luminosity function) (e.g., Gilli et al. 2001; Ueda et al. 2003), or a large population accreting in a radiatively inefficient ADAF-type solution, as invoked to explain discrepancies in the X-ray background produced by synthesis models (Di Matteo et al. 1999), would result in a significant overprediction of the present-day supermassive black hole density. As we demonstrate in § 7.2, invoking such populations is unnecessary, as our picture for

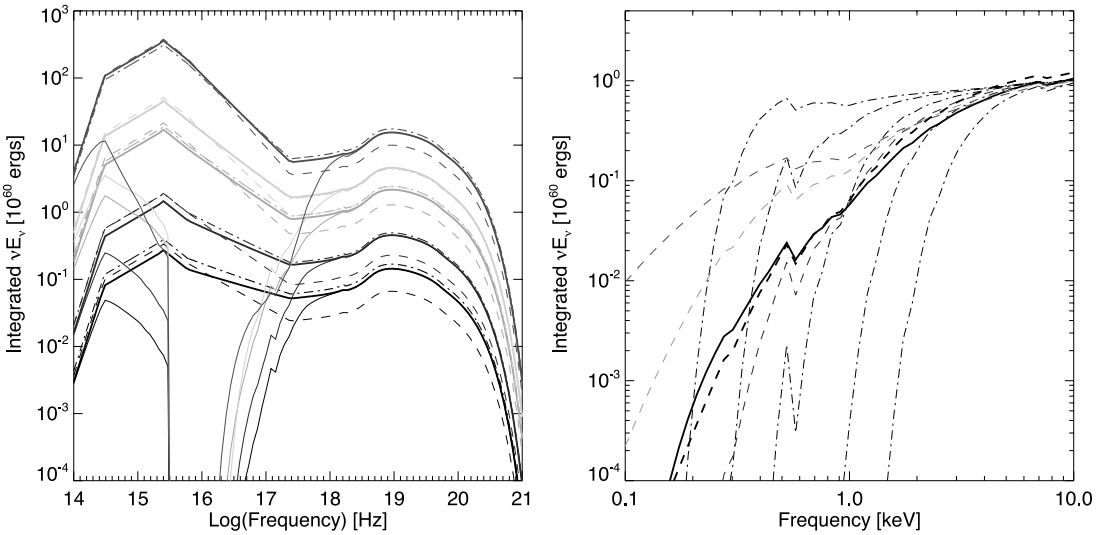

FIG. 24.—*Left*: Integrated intrinsic spectra (*thick solid lines*) from simulations A1, A2, A3, A4, and A5 (*bottom to top*), with virial velocities $V_{\rm vir} = 80, 113, 160, 226,$ and 320 km s^{-1}. The predicted integrated spectra from our model for quasar lifetimes are shown as dot-dashed lines, and the prediction of a light-bulb model, where the same total energy is radiated at $L = L_{\rm peak}$, as dashed lines. Integrated observed spectra are shown as thin solid lines. *Right*: Integrated observed X-ray spectrum from the A3 simulation (*thick solid line*), compared with the integrated intrinsic spectrum, reddened by various column density distributions: our fitted $N_{\rm H}$ distributions from § 2.3 (*thick dashed line*), constant (luminosity independent) lognormal $N_{\rm H}$ distribution with $\bar{N}_{\rm H} = 10^{22}$ cm^{-2} and $\sigma_{N_{\rm H}} = 0.4,\ 0.7,$ and 1.0 (*light dashed lines, bottom to top*), and constant $N_{\rm H} = 10^{21},\ 10^{21.5},\ 10^{22},\ 10^{22.5},$ and 10^{23} cm^{-2} (*thin dot-dashed lines*). [*See the electronic edition of the Supplement for a color version of this figure.*]

quasar lifetimes and evolutionary obscuration self-consistently reproduces the observed X-ray background.

Finally, we note that we reproduce the $z = 0$ distribution of black hole masses *inferred* from the distribution of spheroid velocity dispersions (Sheth et al. 2003) and luminosity functions (Marzke et al. 1994; Kochanek et al. 2001; Nakamura et al. 2003), based on the observed $M_{\rm BH}$-σ relation and fundamental plane for galaxy properties (e.g., Bernardi et al. 2003; Gebhardt et al. 2003). Therefore, since our modeling also reproduces the observed $M_{\rm BH}$-σ (Di Matteo et al. 2005; Robertson et al. 2005b) and fundamental plane (Robertson et al. 2006) relations, we implicitly reproduce the $z = 0$ distribution of spheroid velocity dispersions and spheroid luminosity functions, given our basic assumption that the mergers that produce these spheroids also give rise to luminous quasar activity.

7. THE COSMIC X-RAY BACKGROUND

7.1. *The Integrated Spectra of Individual Quasars*

Unresolved extragalactic sources, specifically obscured AGNs, have been invoked to explain the cosmic X-ray background (e.g., Setti & Woltjer 1989). This picture has been confirmed as deep surveys with *Chandra* and *XMM-Newton* have resolved most or all of the X-ray background into discrete sources, primarily obscured and unobscured AGNs (Brandt et al. 2001; Hasinger et al. 2001; Rosati et al. 2002; Giacconi et al. 2002; Baldi et al. 2002). The X-ray background, however, has a harder X-ray spectrum than typical quasars, with a photon index $\Gamma \sim 1.4$ in the 1–10 keV range (Marshall et al. 1980). Therefore, obscured AGNs are important in producing this shape, as absorption in the ultraviolet and soft X-rays hardens the observed spectrum. Indeed, population synthesis models based on observed quasar luminosity functions and involving large numbers of obscured AGNs have been successful at matching both the X-ray background intensity and spectral shape (Madau et al. 1994; Comastri et al. 1995; Gilli et al. 1999, 2001). However, these models make arbitrary assumptions about the ratio of obscured to unobscured sources and its evolution with redshift, choosing these quantities to reproduce the X-ray background. Furthermore, as X-ray surveys have been extended to higher redshifts, it has become clear that both the observed redshift distribution of X-ray sources and the ratio of obscured to unobscured sources is inconsistent with that required by these models (Hasinger 2002; Barger et al. 2003). Even synthesis models based on higher redshift X-ray surveys and using observationally derived ratios of obscured to unobscured sources (e.g., Ueda et al. 2003) have invoked ad hoc assumptions about additional populations of obscured sources to reproduce the X-ray background shape and intensity.

We can test our model by examining whether the quasar luminosity function, relic AGN mass distribution, and X-ray background can be simultaneously reproduced in a self-consistent manner. Because our formulation describes the birthrate of quasars with a peak luminosity $L_{\rm peak}$, it is most useful to consider the integrated energy spectrum of such a quasar over its lifetime,

$$\nu E_\nu = \int dt\, \nu L_\nu(t) = \int \nu f_\nu(L) L \frac{dt(L,\, L_{\rm peak})}{d\log L}\, d\log L, \quad (38)$$

where $f_\nu(L)$ is the bolometric correction ($L_\nu \equiv f_\nu L$). As an example, Figure 24 shows the integrated intrinsic spectra (*thick solid lines*) from the simulations A1, A2, A3, A4, and A5, described in § 2.1. The final black hole masses for these simulations are $M_{\rm BH}^f = 7\times10^6, 3\times10^7, 3\times10^8, 7\times10^8,$ and $2\times10^9\ M_\odot$, respectively. The integrated spectral shape in the X-ray, in particular, is ultimately determined by the observationally motivated bolometric corrections of Marconi et al. (2004), with a reflection component in the X-ray determined following Magdziarz & Zdziarski (1995), and, in the case of the observed spectrum, the distribution of column densities calculated from the simulations. Using our fits to the lifetime $dt/d\log L$ as a function of instantaneous and peak luminosities, we can calculate the expected νE_ν from the integral above. These integrated spectra are shown as the dot-dashed lines in the figure and agree well with the actual integrated spectra of the simulations, demonstrating the self-consistency of our model and applicability of our fitted lifetimes.

This can be compared to idealized models for the quasar lifetime, where we allow the quasar to radiate just at its peak luminosity $L_{\rm peak} \approx L_{\rm Edd}(M^f_{\rm BH})$ for some fixed lifetime t^0_Q. We determine t^0_Q by demanding that the total energetics be correct, $L_{\rm peak} t^0_Q = \epsilon_r M^f_{\rm BH} c^2$. The predicted integrated energy spectra are shown as the dashed lines and underpredict the soft and hard X-ray energy output by a factor of $\sim$1.5–2. This is because higher luminosity quasars tend to have a larger fraction of their energy radiated in the UV–optical rather than the X-ray (e.g., Wilkes et al. 1994; Green et al. 1995; Vignali et al. 2003; Strateva et al. 2005), reflected in our bolometric corrections. Thus, assuming that the quasar spends all its time at $L_{\rm peak}$ does not account for extended times at lower luminosity, where the ratio of X-ray to total luminosity is higher, which would generate an integrated spectrum with a larger fraction of its energy in the X-ray. Assuming that the quasar undergoes pure Eddington-limited growth to its peak luminosity produces an almost identical integrated spectrum to this light-bulb model, as it is similarly dominated by $L \sim L_{\rm peak}$.

Of course, the intrinsic integrated energy spectrum of the simulations is not what determines the X-ray background, but rather the integrated *observed* spectrum is the critical quantity. This is shown as the thin lines in the left panel of Figure 24, and in detail for our fiducial A3 simulation in the right panel of the figure (*thick solid line*). Along a given sight line, the observed integrated spectrum will be

$$\nu \frac{dE_\nu}{d\Omega} = \int dt\, \nu \frac{L_\nu(t)}{4\pi} e^{-\tau_\nu(\Omega,\, t)}, \tag{39}$$

where τ_ν is the optical depth at a given frequency. We can integrate over solid angle and obtain

$$\nu E_{\nu,\,\rm obs} = \int \nu f_\nu \langle e^{-\tau_\nu} \rangle L \frac{dt(L,\, L_{\rm peak})}{d\log L}\, d\log L, \tag{40}$$

where $\langle e^{-\tau_\nu} \rangle$ is the averaged $e^{-\tau_\nu}$ over the column density distribution $P(N_{\rm H}|L,\, L_{\rm peak})$. Using our fits to the column density distribution and quasar lifetimes and calculating $\nu E_{\nu,\,\rm obs}$ as above, we reproduce the integrated observed spectrum quite well (*thick dashed line*). For comparison, we show that it is not a good approximation to redden the spectrum with a constant $N_{\rm H}$, giving the results for $N_{\rm H} = 10^{21}$, $10^{21.5}$, 10^{22}, $10^{22.5}$, and 10^{23} cm^{-2} (*thin dot-dashed lines*). Even allowing for a distribution of $N_{\rm H}$ values, the resulting spectrum is a poor match to the observed one if that distribution is taken to be static (i.e., luminosity-independent, as in traditional torus models, for example). We show the results of reddening the intrinsic spectrum by such a (Gaussian) distribution, varying the dispersion $\sigma_{N_{\rm H}} = 0.4,\ 0.7,$ and 1.0 (*light dashed lines, bottom to top*) for a median column density $\bar{N}_{\rm H} = 10^{22}$ cm^{-2}, the median column density expected around $L_{\rm peak}$ in this simulation. Therefore, the luminosity and host system property dependence of both quasar lifetimes and the column density distribution must be accounted for in attempting to properly predict the X-ray background spectrum from observations of the quasar luminosity function. Finally, note that the hard cutoff in the observed UV spectra at 912 Å is due to our calculated cross sections being incomplete in the extreme-UV. Properly modeling the escape fraction and observed emission at these frequencies, while not important for the X-ray background, is critical to calculating the contribution of quasars to reionization and requires a more detailed modeling of scattering and absorption, especially in the bright optical quasar phase.

7.2. *The Integrated X-Ray Background*

Given the volume emissivity $j_\nu(z)$ (per unit *comoving* volume) of some isotropic process at a given frequency at redshift z, the resulting background specific intensity at frequency ν_0 at $z = 0$ is (Peacock 1999)

$$I_{\nu_0} = \frac{c}{4\pi} \int \frac{j_\nu[(1+z)\nu_0,\ z]}{(1+z)H(z)}\, dz. \tag{41}$$

If we were to consider the emissivity j_ν per unit physical volume, there would be an extra factor of $(1+z)^{-3}$ in the integral above. In § 7.1 we determined the integrated observed energy $E_{\nu,\,\rm obs}(L_{\rm peak})$ produced by a quasar with peak luminosity $L_{\rm peak}$. We have also inferred $\dot{n}(L_{\rm peak})(z)$ in § 3.2, the rate at which quasars of peak luminosity $L_{\rm peak}$ are created per unit comoving volume per unit cosmological time. Therefore, the comoving volume emissivity is just

$$j_\nu(z) = \int E_{\nu,\,\rm obs}(L_{\rm peak}) \dot{n}(L_{\rm peak})\, d\log L_{\rm peak}, \tag{42}$$

or, expanding $E_{\nu,\,\rm obs}$,

$$j_\nu(z) = \int d\log L_{\rm peak} \int d\log L \times f_\nu \langle e^{-\tau_\nu} \rangle L \frac{dt(L,\ L_{\rm peak})}{d\log L} \dot{n}(L_{\rm peak}). \tag{43}$$

If the column density distribution were independent of $L_{\rm peak}$, as is assumed in even luminosity-dependent torus models or observationally determined $N_{\rm H}$ functions used for X-ray background synthesis (e.g., Ueda et al. 2003), then we could combine terms in $L_{\rm peak}$ and integrate over them. This simplification, along with the definition of the luminosity function in terms of $L_{\rm peak}$, gives the more traditional formula for the X-ray background in terms of only the observed column density distribution and luminosity function,

$$j_\nu(z) = \int d\log L \frac{d\Phi}{d\log L} L_\nu \langle e^{-\tau_\nu} \rangle. \tag{44}$$

However, as we showed in §§ 3.3 and 7.1, neglecting the dependence on $L_{\rm peak}$ is not a good approximation at all luminosities and gives an inaccurate estimate of the integrated quasar spectrum; therefore, "purely observation-based" synthesis models of the X-ray background will be inaccurate in a similar manner to synthesis models with an inappropriate model for the quasar lifetime. Essentially, this "averages out" the varying distribution of column densities with $L_{\rm peak}$, which changes the shape of the spectrum in a nonlinear manner, especially when integrated over varying bolometric corrections as shown above.

Figure 25 (*top panel*) shows the predicted X-ray background spectrum from our full modeling of quasar lifetimes and obscuration (*solid black lines*). We use our analytical fits to the quasar lifetime and column density distributions as in § 7.1 above, as Figure 24 demonstrates that they accurately reproduce the actual integrated quasar X-ray spectra of the simulations, and the analytical forms are integrated over all luminosities and redshifts. The dotted lines show the deviation resulting from shifting the parameters describing our fitted $\dot{n}(L_{\rm peak})$ distribution by 1 σ in either direction, although degeneracies in the parameters suggest that the actual uncertainty in the background prediction is smaller. The dashed line shows the predicted X-ray background if we

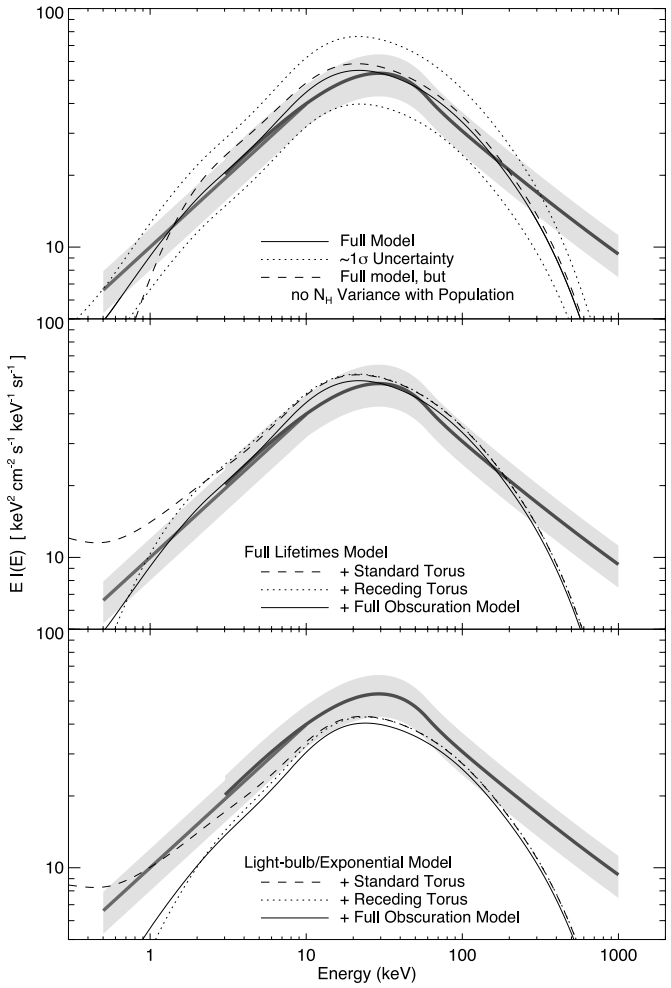

FIG. 25.—Predicted integrated X-ray background spectrum (*solid black line*) from our model of quasar lifetimes and attenuation, with the peak luminosity distribution $\dot{n}(L_{\rm peak})$ determined from the luminosity function. Lighter and darker thick lines show the observed spectrum from Barcons et al. (2000) and Gruber et al. (1999), respectively. The light gray shaded area illustrates the uncertainty in normalization between both samples (alternatively, 2 σ errors in the Barcons et al. 2000 normalization). The predictions given 1 σ deviations in the fitted $\dot{n}(L_{\rm peak})$ distribution (*dotted lines*) and given the $\dot{n}(L_{\rm peak})$ distribution determined from hard X-ray data only (*dashed line*) are shown in the top panel. Middle panel shows the prediction using our modeling of quasar lifetimes but different models of obscuration, bottom panel the prediction with a light-bulb or exponential (constant Eddington ratio) model and different obscuration models. [*See the electronic edition of the Supplement for a color version of this figure.*]

ignore the broadening of the $N_{\rm H}$ distribution across simulations ($\sigma_{N_{\rm H}} = 1.2$) and instead consider only the dispersion of an individual simulation at a given luminosity ($\sigma_{N_{\rm H}} = 0.4$).

These can be compared to the observations of Gruber et al. (1999) (*darker thick curve*; for $E \geq 3$ keV) and Barcons et al. (2000) (*lighter thick curve*; for $E \leq 10$ keV). We increase the normalization of the Gruber et al. (1999) spectrum to match that of the best estimate from Barcons et al. (2000) over the range of overlap, determined from combined *ASCA, BeppoSAX*, and *ROSAT* data to be $10.0^{+0.6}_{-0.9}$ keV cm^{-2} s^{-1} sr^{-1} keV^{-1} at 1 keV. The uncertainty in the normalization between the two samples, $\sim$20%, is shown as the gray shaded range (alternatively, this represents the $\sim$2 σ errors in the *ROSAT* normalization).

In the middle panel of the figure, we calculate the predicted X-ray background using our full model of the quasar lifetime, but with different models for quasar obscuration. The solid black line shows the prediction using our full model of quasar obscuration and is identical to the solid black line in the top panel. The observations are likewise shown in a manner identical to the top panel. The dashed black line is the prediction adopting the standard torus model for quasar obscuration, and the dotted line adopts the receding (luminosity dependent) torus model. These models produce the same overall $\sim$30 keV normalization, as this is relatively unaffected by obscuration, but they predict a slightly ($\sim$20%) higher background at low energies, giving a slightly softer spectrum. This may appear counterintuitive, given that in Figure 12 these models tend to overpredict the number of high column density sources, but this is because these models predict a strongly *bimodal* column density distribution, with unobscured sight lines encountering negligible column densities. These unobscured sight lines dominate the soft X-ray integrated spectrum, where the large column densities through the torus attenuate the quasar spectrum heavily. However, this net offset in the predicted background spectrum is generally within the range of the systematic theoretical and observational uncertainties and can further be alleviated by tuning the parameters of the torus model to fit the X-ray background spectrum (e.g., Treister & Urry 2005, although their fits require a larger fraction of Compton-thick $N_{\rm H} \sim 10^{25}$ cm^{-2} sources than shown for even the receding torus model in Fig. 12). The feature at $\lesssim$5 keV in the standard torus model prediction is a consequence of assuming that "unobscured" lines of sight encounter negligible column density, and does not appear if such sight lines encounter moderate ($\sim 10^{21}$ cm^{-2}) columns.

The bottom panel of the figure shows the predicted X-ray background spectrum if we instead consider a light-bulb or exponential light curve (fixed Eddington ratio) model for the quasar lifetime, again with various descriptions of quasar obscuration. In such models, the predicted X-ray background spectrum is independent of the quasar lifetime or characteristic Eddington ratio assumed (see eq. [44]). However, as shown in Figure 24, these models do imply a different integrated spectrum for quasars, i.e., different effective bolometric corrections for predicting the X-ray background. In particular, in this model, the observed quasar spectrum at a given luminosity (averaged over the quasar population at that luminosity) is the same as the "effective" quasar spectrum one would use to calculate the *total* contribution to the X-ray background from quasars of the corresponding observed or peak luminosity, whereas this is not true in our model of quasar lifetimes. The observations are shown in the same manner as the preceding panels. The black solid line shows the prediction with this simplified model for the quasar lifetime, but still adopting our full model for obscuration as a function of instantaneous and peak luminosity, the dashed line assumes instead a standard torus model for obscuration, and the dotted line assumes a receding torus for the obscuration. The variations among different obscuration models are relatively small at most energies and similar to those discussed above adopting our full model of quasar lifetimes.

In all three cases, however, this model of the quasar lifetime significantly underpredicts the X-ray background, particularly at the $\sim$30 keV peak. This shortfall is well known, and earlier attempts (e.g., Madau et al. 1994; Comastri et al. 1995; Gilli et al. 1999, 2001; Pompilio et al. 2000; Ueda et al. 2003) have generally had to invoke additional assumptions about large obscured populations or a strong increase in the obscured fraction with redshift, neither of which is consistent with observations (e.g., Hasinger 2002; Barger et al. 2003, 2005; Ueda et al. 2003; Szokoly et al. 2004). The difference between the predictions of various quasar lifetime models is, as explained above, attributable to the difference between the integrated quasar spectrum produced in our full model of the quasar lifetime (in which quasars spend

long periods of time at low luminosities, with harder X-ray spectra) and the integrated spectrum in these simplified quasar lifetime models, which is proportional to the *instantaneous* quasar spectrum and therefore underpredicts the hard X-ray portion of the spectrum by as much as ~50%.

Our prediction of the X-ray background agrees well with the observed spectrum over the range ~1–100 keV. (At energies above 100 keV it is likely that processes we have not included, such as those involving magnetic fields, contribute significantly to the background.) Unlike previous synthesis models for the X-ray background, we are able to do so without invoking assumptions about large Compton-thick populations or larger obscured populations at different redshifts. In part, this is because our modeling allows us to predict, based on $\dot{n}(L_{\rm peak})$ and our column density formulation, the population of Compton-thick sources (see Fig. 12). However, as we have demonstrated, it is primarily because the deficit in previous synthesis models can be attributed to their inability to properly account for the dependence of quasar lifetimes and attenuation on both the instantaneous quasar luminosity and the host system properties (peak luminosity). Our picture, on the other hand, yields an estimate for the X-ray background spectrum that is simultaneously consistent with the observed supermassive black hole mass distribution and total density, as well as the "luminosity-dependent density evolution" observed in X-ray samples (Hopkins et al. 2006a). The background is primarily built up from $z \sim 2.5$ to 0.5, as is evident from the evolution of the black hole mass density in Figure 22, although a harder spectrum at low luminosities will weight this slightly toward lower redshifts (where more low-luminosity quasars are forming). Compton-thick and relaxing, low-luminosity sources are accounted for, not as large, independent populations, but as evolutionary phenomena continuously connected to the "normal" quasar population.

8. DISCUSSION

8.1. *General Implications of Our Model*

Our modeling suggests two important paradigm shifts in interpreting quasar populations and evolution:

1. First, as proposed in Hopkins et al. (2005b), a proper accounting of the luminosity dependence of quasar lifetimes (as opposed to models in which quasars grow in a pure exponential fashion or turn on and off as "light bulbs") implies a novel interpretation of the luminosity function. The steep bright end (luminosities above the break in the luminosity function) consists of quasars radiating near their Eddington limits and is directly related to the distribution of intrinsic peak luminosities (or final black hole masses), as has been assumed previously. However, the shallow, faint end of the luminosity function describes black holes either growing in early stages of activity or in extended, quiescent states going into or coming out of a peak bright quasar phase, with Eddington ratios generally between $l \sim 0.01$ and 1. The break luminosity in the luminosity function corresponds directly to the *peak* in the birthrate of quasars as a function of peak luminosity $\dot{n}(L_{\rm peak})$.

This interpretation resolves inconsistencies in a number of previous theoretical studies. For example, semianalytical models of the quasar luminosity functions (e.g., Kauffmann & Haehnelt 2000; Haiman & Menou 2000; Wyithe & Loeb 2003) assume, based on simplified models for the quasar lifetime, that quasars at the faint end of the luminosity function correspond to low-final-mass black holes (low $L_{\rm peak} \sim L$), presumably in small halos. Consequently, these models overpredict the number of active low-mass black holes (as estimated from radio source counts), especially at high redshift, by orders of magnitude (Haiman et al. 2004) and overpredict the number of low-mass spheroids and red galaxies observed (Hopkins et al. 2006b).

Moreover, both observations (McLure & Dunlop 2004) and comparison of the present-day black hole mass function with radio and X-ray luminosity functions (e.g., Merloni 2004) suggest antihierarchical evolution for the growth of supermassive black holes, where the most massive black holes were produced mainly at high ($z \gtrsim 2$) redshift, and low-mass black holes mostly formed later, which does not follow from idealized descriptions of quasar lifetimes and the luminosity function (for a review, see, e.g., Combes 2006).

A one-to-one correspondence between observed luminosity and black hole mass does produce antihierarchical behavior in some sense at the high-mass end, because the most massive black holes are formed at $z \sim 2-3$ during the peak of bright quasar activity and the quasar luminosity function evolves to lower luminosities at lower redshifts (as is also the case for our model because the bright end of the luminosity function is dominated by sources near their peak luminosities). However, at black hole masses equal to or below $\sim 10^8\ M_\odot$ (i.e., galaxies of stellar mass $\lesssim 10^{11}\ M_\odot$), the evolution in the quasar luminosity function implies a roughly constant production of black holes with these masses at all redshifts, which is inconsistent with observations of galaxy spheroids indicating that typical ages increase with mass, ruling out a large population of low-mass spheroids with ages equal to or older than those of high-mass spheroids (e.g., Treu et al. 2001, 2002, 2005; van Dokkum et al. 2001; van Dokkum & Stanford 2003; Gebhardt et al. 2003; Rusin et al. 2003; van de Ven et al. 2003; Wuyts et al. 2004; Holden et al. 2005; van der Wel et al. 2005; di Serego Alighieri et al. 2005; Nelan et al. 2005). As demonstrated in Figure 23 such a model does not produce antihierarchical growth or any age gradients within the high-mass spheroid population, also inconsistent with observations. Even given observed luminosity-dependent density evolution (e.g., Page et al. 1997; Miyaji et al. 2000, 2001; La Franca et al. 2002; Cowie et al. 2003; Ueda et al. 2003; Fiore et al. 2003; Hunt et al. 2004; Cirasuolo et al. 2005; Hasinger et al. 2005), implying that the densities of lower redshift quasars peak at lower redshift, the inferred antihierarchical evolution, if observed luminosity directly corresponds to black hole mass (i.e., as in light-bulb or fixed-Eddington-ratio models), is not strong enough to account for observed antihierarchical growth of the corresponding galaxy spheroids (Hopkins et al. 2006b).

Furthermore, in these earlier models a break in the luminosity function is not necessarily reproduced (Wyithe & Loeb 2003) and the faint-end slope has no direct physical motivation. The break may be caused by feedback mechanisms that set a characteristic turnover in both the galaxy mass function and quasar luminosity function (e.g., Scannapieco & Oh 2004; Dekel & Birnboim 2006), as in our modeling. The $\dot{n}(L_{\rm peak})$ distributions in our model and light-bulb or fixed-Eddington-ratio models are comparable at and above the break in the quasar luminosity function and therefore make similar predictions for some observations at these luminosities. However, the *faint-end* slope has a different physical motivation in our model. Unlike the bright-end slope, which is determined directly by the active final black hole mass function or peak luminosity distribution (in essentially all models of the quasar lifetime), the faint-end slope in our modeling is a consequence of the quasar lifetime as a function of luminosity, and is a prediction of our simulations and modeling almost independent of the underlying faint-end slope of the active black hole mass function or peak luminosity distribution. In Hopkins et al. (2006a) we examine this in more detail, and

demonstrate that it predicts well the evolution in the faint-end quasar luminosity function slope with redshift and the observed luminosity-dependent density evolution in many samples (Page et al. 1997; Miyaji et al. 2000, 2001; La Franca et al. 2002; Cowie et al. 2003; Ueda et al. 2003; Fiore et al. 2003; Hunt et al. 2004; Cirasuolo et al. 2005; Hasinger et al. 2005).

Other observational evidence for our picture exists, for example in the observed distribution of Eddington ratios (see § 5), the distribution of low-redshift, active black hole masses (see § 4.3), and the turnover in the expected distribution of black hole masses in early-type galaxies at $\sim 10^8\ M_\odot$ (e.g., Sheth et al. 2003). Total (integrated) quasar lifetimes estimated from observations are inferred to increase with increasing black hole mass as we predict (Yu & Tremaine 2002), and furthermore, the Eddington ratios of observed quasar samples are seen to increase systematically with redshift, as the sample becomes increasingly dominated by luminosities above the break in the luminosity function (McLure & Dunlop 2004).

Moreover, observations show that the evolution of the luminosity function with decreasing redshift is driven by a decrease in the characteristic mass scale of actively accreting black holes (e.g., Heckman et al. 2004), which can be explained in our model by the relation of the observed luminosity function to the *peak* in the distribution of active black hole masses $\dot{n}(L_{\rm peak})$. This observation, however, has caused considerable confusion, as observations of both radio-quiet (Woo & Urry 2002) and radio-loud (O'Dowd et al. 2002) local (low redshift) AGNs indicate that nuclear and host luminosities are uncorrelated, implying that nuclear luminosity does not depend on black hole mass (Heckman et al. 2004), and therefore that the primary variable determining the nuclear luminosity is the Eddington ratio, with the luminosity function spanning a broad range in Eddington ratios (Hao et al. 2005). Furthermore, observations show that this is *not* true of high-redshift quasars, as both direct estimates of accretion rates (e.g., Vestergaard 2004; McLure & Dunlop 2004) and the fact that their high luminosities would yield unreasonably large black hole masses rule out substantially sub-Eddington accretion rates for most objects. Many previous empirical and semianalytical models could not simultaneously account for these observations. To explain just the low-redshift observations, such models adopt tunable distributions of Eddington ratios fitted to the data. However, both these observations are consequences of our interpretation of the luminosity function, as observations of local AGNs and the low-redshift luminosity function are dominated by quasars below the break in the luminosity function, which are undergoing sub-Eddington growth and span a wide range of Eddington ratios, while observations at high redshift are dominated by bright objects at or above the break in the luminosity function, which are undergoing Eddington-limited (or near Eddington-limited) growth near their peak luminosity (see § 5).

2. The second paradigm shift indicated by our modeling is that quasar obscuration is not a static or quasi-static geometric effect, but is primarily an *evolutionary* effect. The physical reasoning for this is simple: the massive gas inflows required to fuel quasar activity produce large obscuring columns, and so column densities are correlated with quasar luminosity. The basic picture of buried quasar activity associated with the early growth of supermassive black holes and starburst activity has been proposed previously and studied for some time (e.g., Sanders & Mirabel 1996; Fabian 1999), but our modeling allows us to describe the evolution of obscuration in a self-consistent manner, defining obscured and unobscured phases appropriately and identifying dynamical correlations between the column density distribution and instantaneous and peak luminosities.

There is substantial observational support for this picture. Pointlike X-ray sources have been observed in many bright submillimeter or infrared and starburst sources, with essentially all very luminous infrared galaxies showing evidence of buried quasar activity (e.g., Sanders & Mirabel 1996; Komossa et al. 2003; Ptak et al. 2003), indicating simultaneous buried black hole growth and star formation at redshifts corresponding to peak quasar activity ($z \gtrsim 1$) (Alexander et al. 2005a, 2005b). The buried black holes in high-z starbursting galaxies appear to be active but undermassive compared to the quiescent galaxy black hole–stellar mass relation (Borys et al. 2005), implying that they are rapidly growing in the starburst but have not yet reached their final masses, presumably set in the subsequent blowout phase. Similarly, observations suggest that obscured AGNs are significantly more likely to exhibit strong submillimeter emission characteristic of star formation, implying both that obscured black hole growth and star formation are correlated and that obscuration mechanisms (responsible for reradiation in the submillimeter and IR) may be primarily isotropic in at least some cases (e.g., Page et al. 2004; Stevens et al. 2005). Evidence from quasar emission line structure (e.g., Kuraszkiewicz et al. 2000; Tran 2003), directly related to the inner broad-line region, suggests that isotropic obscuration of quasars can be important, in contradiction to angle-dependent models. Finally, many observations (e.g., Steffen et al. 2003; Ueda et al. 2003; Hasinger 2004; Grimes et al. 2004; Sazonov & Revnivtsev 2004; Barger et al. 2005; Simpson 2005) indicate that the fraction of broad-line or obscured quasars is a function of luminosity, which cannot be accounted for in traditional static "torus" models (e.g., Antonucci 1993) or reproduced even by modified luminosity-dependent torus models (Lawrence 1991), an observation that is explained by our model (see § 4 for a detailed discussion).

Much of the obscuration in our modeling comes from large scales, arising from the inner regions of the host galaxy on scales of $\sim$50 pc or larger. While our resolution limits prevent our ruling out the possibility of gas collapse to a dense, approximately parsec scale torus surrounding the black hole, during the peak obscured phases of the final merger, our simulations indicate that these large scales dominate the contribution to the column density, with quite large columns, which should be observationally testable. Indeed, this is suggested by the typical scales of obscuration in starbursting systems (e.g., Soifer et al. 1984a, 1984b; Sanders et al. 1986, 1988a, 1988b; for a review, see, e.g., Soifer et al. 1987), given that, as discussed above, the dominant obscured phase of growth is closely associated with a starburst as implied observationally (Alexander et al. 2005a, 2005b; Borys et al. 2005).

Observations of polarized light in intrinsically bright type II AGNs with unobscured luminosities typical of quasars (as opposed to local, dim Seyfert II objects in relaxed hosts) show scattering on large scales, approximately kiloparsecs, and in some cases obscuration clearly generated over scales extending beyond the host galaxy in the form of distortions, tidal tails, and streams from interactions and major mergers (Zakamska et al. 2004, 2005). The angular structure seen in these observations is consistent with our modeling. Moreover, in optically faint X-ray quasars (e.g., Donley et al. 2005) it appears that obscuration is generated by the host galaxies and is directly related to host galaxy morphologies and line-of-sight distance through the host. The critical point is that regardless of the angular structure of obscuration, typical column densities are strongly evolving functions of time, luminosity, and host system properties, and the observed distribution of column densities is dominated by these effects, not by differences in viewing angle across a uniform

population. This is the case in our modeling, as the lognormal dispersion (across different lines of sight) in column densities is $\sigma_{N_H} \sim 0.4$ for a given simulation at some instant, whereas typical column densities across simulations, as a function of instantaneous and peak luminosities, span several orders of magnitude from $N_H \sim 10^{18}$ to 10^{26} cm^{-2}.

8.2. *Specific Predictions of Our Model*

Our predictions include the following:

1. *Quasar lifetimes*.—We find that for a particular source, the quasar lifetime depends sensitively on luminosity, with the observed lifetime in addition depending on the observed wave band. Intrinsic quasar lifetimes vary from $t_Q \sim 10^6$ to 10^8 yr, with observable lifetimes of $\sim 10^7$ yr in optical B band (Hopkins et al. 2005a, 2005d), in good agreement with observational estimates (for a review, see Martini 2004).

2. *Luminosity functions*.—Using a parameterization of the intrinsic distribution of peak luminosities (final quasar black hole masses) at a given redshift, our model of quasar lifetimes allows us to reproduce the observed luminosity function at all luminosities and redshifts $z = 0-6$. Although this is an empirical determination of the peak luminosity distribution, it implies a new interpretation of the luminosity function (Hopkins et al. 2005b), which provides a physical basis for the observed break corresponding to the peak in the peak luminosity distribution. Moreover, the faint-end slope is not determined by our empirical fitting procedure, but instead by the dependence of the quasar lifetime on luminosity, with its value and redshift evolution predicted by our modeling (Hopkins et al. 2006a). The evolution of typical column densities in different stages of merger activity produces a significant population of obscured quasars, accounting for the difference between hard X-ray (e.g., Ueda et al. 2003), soft X-ray (e.g., Miyaji et al. 2001), and optical B band (e.g., Croom et al. 2004) luminosity functions (§ 3.2).

3. *Column density distributions*.—The evolution of the column densities in our simulations reproduces the observed distribution of columns in optically selected quasar samples, when the appropriate selection criteria are applied (Hopkins et al. 2005a), as well as complete column distributions in hard X-ray selected samples (§ 3.3). Column density evolution over the course of a merger yields a wider observed distribution of columns than that produced across different viewing angles at a given point in a merger.

4. *Broad-line luminosity function and fraction*.—Using our simulations to estimate when quasars will be observable as broad-line objects (either based on the ratio of quasar to host galaxy optical B-band luminosity or the obscuring column density), we reproduce the luminosity function of broad-line quasars in hard X-ray selected samples as well as in optical broad-line quasar surveys, and the fraction of broad-line quasars in a given sample as a function of luminosity, to better precision than traditional or luminosity-dependent (but nondynamical) torus models that are fitted to the data (§ 4.2). By providing an a priori prediction of the broad-line fraction as a function of luminosity and redshift that depends systematically on the typical quasar host galaxy gas fraction, we propose that observations of the broad-line fraction at different redshifts can be used to constrain the gas fraction of quasar hosts and its evolution with redshift.

5. *Active black hole mass functions*.—Using our prescription for deciding when objects will be visible as broad-line quasars, we predict the distribution of low-redshift, broad-line and non-broad-line active quasar masses, in good agreement with observations from the SDSS, with expected incompleteness in the observed sample at low ($M_{BH} \lesssim 10^6\ M_\odot$) black hole masses (§ 4.3). This is a new prediction that can be tested in greater detail by future observations, and our calculations allow us to model the differences in active black hole mass functions of the type I and type II populations. The width of the expected broad-line black hole mass function depends significantly on the model of quasar lifetimes, enabling such measurements to probe the statistics of quasar evolution.

6. *Eddington ratios*.—We determine Eddington ratio distributions from our simulations, given the peak luminosity distribution implied by the observed quasar luminosity function. The predicted distribution, once the appropriate observed magnitude limit is imposed, agrees well with observations at both low ($z < 0.5$) and high ($1.5 < z < 3.5$) redshifts (§ 5). As noted above, our interpretation of the luminosity function explains seemingly contradictory observations of Eddington ratios at different redshifts. There is even a suggestion (Cao & Xu 2005) that the evolution of quasars seen in our simulations (with bright phases in mergers and extended relaxation after) can account for observations of bimodal Eddington ratio distributions at $z \sim 0$ (Marchesini et al. 2004) when coupled with an appropriate description of radiatively inefficient accretion phases, although it is possible that many of these low-redshift black holes are not fueled by mergers, especially in, for example, low-luminosity Seyfert galaxies.

7. *Relic black hole mass function*.—With our model for quasar lifetimes, the luminosity function at a given redshift implies a birthrate of sources with given peak luminosities, $\dot{n}(L_{peak})$, which translates to a distribution in final black hole masses. Integrating this over redshift, we predict the present-day mass distribution and total mass density of supermassive black holes. They agree well with observational estimates inferred from local populations of galaxy spheroids. In our picture, these spheroids are produced simultaneously with the supermassive black holes they harbor (§ 6). We demonstrate that the integrated supermassive black hole density, quasar flux density, and number counts in different wave bands can be reconciled with a radiative efficiency $\epsilon_r = 0.1$, satisfying the constraints of counting arguments such as that of Soltan (1982). Further, we show in §§ 2.4 and 4.1 that the corrections to such observational arguments based on optical quasar samples are small (order unity) when we account for the luminosity dependence of quasar lifetimes, despite an extended obscured phase of quasar growth. In other words, although a quasar spends more time obscured than it does as a bright optical source, the total mass growth and radiated energy are dominated by the final blowout stage visible as a bright optical quasar.

8. *X-ray background*.—The integrated quasar spectrum from our models of quasar lifetimes and column densities as a function of instantaneous and peak luminosities can be combined with the birthrate of quasars with a given peak luminosity to give the integrated cosmic background in any frequency range. We predict both the normalization and shape of the X-ray background from $\sim$1 to 100 keV, with our modeling accounting for quasar obscuration as an evolutionary process (with a corresponding population of Compton-thick objects), avoiding any need for arbitrary assumptions about additional obscured populations (§ 7.2). For any model in which the quasar spectrum depends on luminosity or accretion rate, we demonstrate that a proper modeling of the quasar lifetime is critical to reproducing observed backgrounds.

9. *Correlation functions*.—In Lidz et al. (2006), we predict the quasar correlation function and bias as a function of redshift and luminosity using our model, and compare it to that expected using light-bulb or exponential light curve models. As most

quasars in our modeling have a characteristic peak luminosity or final black hole mass corresponding to the peak of the $\dot{n}(L_{\rm peak})$ distribution, they reside in hosts of similar mass, and there is little change in bias with luminosity at a given redshift, in contrast to idealized models for the quasar lifetime and luminosity function. Our predicted bias agrees well with the observations of Croom et al. (2005), who also find no evidence for a dependence of the correlation on quasar luminosity at a given redshift, as we expect. In fact, Porciani et al. (2004) and Croom et al. (2005) find that their observations can be explained if quasars lie in hosts with a constant characteristic mass of $\sim 2\times 10^{12}\ M_{\odot}$ ($h = 0.7$). If we consider their redshift range $z\sim 1-2$, we predict the quasar population will be dominated by sources with $L_{\rm peak} = L_{*}(z)$, which given $M_{\rm BH}^{f}(L_{\rm peak})$ and using the $M_{\rm BH}$-$M_{\rm halo}$ relation of Wyithe & Loeb (2003) yields a nearly constant characteristic host halo mass of $\sim(1-2)\times 10^{12}\ M_{\odot}$, in good agreement. Similarly, Adelberger & Steidel (2005) find that the quasar-galaxy cross-correlation function does not vary with luminosity, implying with $\sim$90% confidence that faint and bright quasars reside in halos with similar masses and that fainter AGNs are longer lived, strongly disfavoring traditional light-bulb and exponential light curve models. Furthermore, Hennawi et al. (2006) find an order-of-magnitude excess in quasar clustering at small scales, $\lesssim 40\ h^{-1}$ kpc, with the correlation function becoming progressively steeper at sub-Mpc scales, suggesting that quasar activity is triggered by interactions and mergers.

10. *Host galaxy properties*.—Because black hole growth and spheroid formation occur together in our picture, our modeling allows us to describe relationships between black hole and galaxy properties. For example, we reproduce both the observed $M_{\rm BH}$-σ relation (Di Matteo et al. 2005; Robertson et al. 2005b) and the fundamental plane of elliptical galaxies (Robertson et al. 2006). Since we also reproduce the distribution of relic black holes inferred from the $z = 0$ distribution of spheroid velocity dispersions or luminosity functions using the observed versions of these relations, our match to these relations indicates that we also reproduce these distributions of host spheroid properties. We consider this in detail in Hopkins et al. (2006b) and find that we are able to account for a wide range of host galaxy properties, including luminosity and mass functions, color-magnitude relations, mass-to-light ratios, and ages as a function of size, mass, and redshift. With our modeling of the quasar lifetime as motivated by our simulations, the evolution and distribution of properties of red-sequence galaxies and the quasar population are shown to be self-consistent, which is not the case for idealized models of quasar evolution.

Aside from an empirical estimate of the distribution of peak quasar luminosities $\dot{n}(L_{\rm peak})$, we determine all of the quantities summarized above self-consistently from the input physics of our simulations, including a physically motivated dynamic accretion and feedback model in which black holes accrete at the Bondi rate determined from the surrounding gas, and $\sim$5% of the radiant energy couples thermally to that gas. Beyond this, our simulations enable us to calculate the various predictions above a priori, without the need for additional assumptions or tunable parameters.

We compare each of these predictions to those obtained using idealized descriptions of the quasar lifetime, i.e., light-bulb and exponential light curve (constant Eddington ratio) models, and the column density distribution, i.e., standard and "receding" (luminosity dependent) torus models. We fit all these (along with our full model) to the observed luminosity function in the same manner (allowing the same degree of freedom to ensure that they all yield the same observed luminosity function), and we fit the free parameters of these tunable models (e.g., typical Eddington ratios and quasar lifetimes for the light-bulb or exponential models, typical column densities and torus scalings for the torus models) *independently* to each observation to maximize their ability to reproduce observations. However, we still find better agreement between our model (with no parameters tuned to match observations) and the observations in nearly every case where the tunable phenomenological model is not guaranteed to reproduce the observation by construction. The one exception is the relic supermassive black hole mass function, for which the predictions of our modeling and idealized lifetime models are essentially identical, reflecting the fact that in both cases black hole growth is dominated by bright, optically observable, high Eddington ratio phases.

Moreover, the best-fit parameters for the idealized models, when fitted independently to each observation, are not self-consistent. For example, calculations of the black hole mass function imply high Eddington ratios $l\sim 0.5-1$ (e.g., Yu & Tremaine 2002), and our fit to the active black hole mass function (Heckman et al. 2004) suggests $l\sim 1$, but the observed distribution of accretion shows a typical $l\sim 0.3$ (Vestergaard 2004), and fitting to the broad-line fraction as a function of luminosity with our full obscuration model but these lifetime models implies a lower $l\sim 0.05$. Likewise, fitting torus models to the X-ray background suggests typical column densities through the torus of $N_{\rm H}\sim 10^{25}\ {\rm cm}^{-2}$ (e.g., Treister & Urry 2005), while fitting to the observed column density distributions (Treister et al. 2004; Mainieri et al. 2005) suggests equatorial columns $N_{\rm H}\lesssim 10^{24}\ {\rm cm}^{-2}$. Clearly then, reproducing the observations listed above, and in particular doing so self-consistently, is not implicit in any model that successfully reproduces the quasar luminosity function, even at multiple frequencies.

8.3. *Further Testable Predictions of Our Model*

Our model for quasar evolution makes a number of observationally testable predictions:

1. Quasar lifetimes are only weakly constrained by observations (e.g., Martini 2004), but future studies may be able to measure both the lifetime of individual quasars and the statistical lifetimes of quasar populations as a function of luminosity. We describe in detail our predictions for the evolution of individual quasars and quantify their lifetimes in § 2, and further predict the distribution of both integrated and differential lifetimes in an observed sample as a function of luminosity. This should provide a basis for comparison with a wide range of observations, with the most important prediction being that the quasar lifetime should increase with decreasing luminosity.
2. For a reasonably complete, optically selected sample above some luminosity, the distribution of observed column densities should broaden to both larger and smaller $N_{\rm H}$ values as the minimum observed luminosity is decreased, as both intrinsically faint periods with low column density and intrinsically bright periods with high column density become observable.
3. Similarly, the Eddington ratio distribution should be a function of observed luminosity, with a broad distribution of Eddington ratios down to $l\sim 0.01-0.1$ at luminosities well below the break in the observed luminosity function, and a more strongly peaked distribution about $l\sim 0.2-1$ for luminosities above the break (Fig. 20).
4. In our interpretation, the bright and faint ends of the luminosity function correspond statistically to similar mixes of galaxies, but in various stages of evolution; whereas in all other competing scenarios, the quasar luminosity is directly related to

the mass of the host galaxy. Therefore, any observational probe that differentiates quasars based on their host galaxy properties such as, for example, the dependence of the clustering of quasars on luminosity or the host stellar mass and size as a function of luminosity (although we caution that this is somewhat dependent of the modeling of star formation in mergers), can be used to discriminate our picture from older models. We present a detailed prediction of the quasar correlation function based on our modeling for comparison with observations in Lidz et al. (2006).

5. Our distribution $\dot{n}(L_{\rm peak})$ directly translates to a black hole merger rate, as a function of mass, in our modeling, allowing a detailed prediction of the gravitational wave signal from black hole mergers as a function of redshift.

6. The broad-line fraction as a function of luminosity, defined by requiring that broad-line objects have an observed B-band luminosity above a fraction $f_{\rm BL}$ of that of their host galaxy, is a prediction of our model quasar and galaxy light curves. However, the uncertainties are large, primarily because different observational samples have varying sensitivity to quasar versus host galaxy optical light. Furthermore, the host galaxy gas fraction and $f_{\rm BL}$ are degenerate in these predictions; a well-defined observational sample complete to some $f_{\rm BL}$ can constrain our modeling of quasar fueling and the relation between quasar and host galaxy light curves. In particular, such observations, either by measuring the faint-end shape of the broad-line quasar luminosity function or the mean broad-line fraction at a given luminosity as a function of redshift, can constrain the gas fractions of quasar host galaxies and their evolution, essentially a free parameter in our empirical modeling.

7. We also predict the distribution of active, low-redshift black hole masses in § 4. These predictions can be compared to mass functions for active black holes from numerous quasar surveys, which should include improved mass functions of the entire quasar population complete to lower luminosities as well as future mass functions for the population of active broad-line AGNs. We provide predictions for the black hole mass function of all active quasars, and for just the broad-line population (as a function of the survey selection).

8. Because the evolution of spheroids and supermassive black holes is linked in our modeling, with each affecting the evolution of the other, we can also use the distribution of observed quasar properties to predict galaxy properties such as number counts, spheroid masses and luminosities, and colors as a function of redshift. For the calculation and discussion of these predictions, see Hopkins et al. (2006b).

9. In our model, the growth of supermassive black holes is dominated by galaxy mergers. Therefore, at any given redshift, the mass (and as a consequence, luminosity) function of galaxy mergers should have a shape similar to our distribution of quasar birthrates, $\dot{n}(L_{\rm peak})$, distinct from the shapes of either the quasar or total galaxy luminosity functions. Indeed, preliminary observational estimates of both the merger luminosity function (e.g., Xu et al. 2004; Conselice et al. 2003; Wolf et al. 2005) and quasar host galaxy luminosity function (Bahcall et al. 1997; Hamilton et al. 2002), primarily at low redshifts, appear be consistent with this expectation. Theoretically, it may be possible to predict the merger luminosity function using either cosmological simulations or semianalytical models; we discuss this further in § 9.

8.4. *Mock Quasar Catalogs*

In principle, our modeling can be used to predict the distributions of quasar luminosities in various wave bands, column densities, active black hole masses, and peak luminosities for a wide range of observational samples, but it is impractical for us to plot predictions of these quantities for all possible sample selection criteria. To enable comparison with a wider range of observations, we have used our modeling and the conditional probability distributions for these quantities from our simulations to generate Monte Carlo realizations of quasar populations, which we provide publicly via ftp.[4]

At a particular redshift, we use our fitted $\dot{n}(L_{\rm peak})$ distribution and our suite of simulations to generate a random population of mock "quasars." We first generate the peak luminosities of each "quasar" according to the fitted $\dot{n}(L_{\rm peak})$ at that redshift. For each object, we then use the probability of being at a given instantaneous luminosity in simulations with a similar peak luminosity to generate a current bolometric luminosity. In practice, we calculate the $P(L|L_{\rm peak})$ distribution by summing $w(L_{\rm peak},\ L_{{\rm peak},\,i}) \times P(L|L_{{\rm peak},\,i})$, where $L_{\rm peak}$ is the mock quasar peak luminosity, $L_{{\rm peak},\,i}$ is the peak luminosity of each simulation, and $w(L_{\rm peak}, L_{{\rm peak},\,i})$ is a Gaussian weighting factor ($\propto \exp[-\log^2(L_{\rm peak}/L_{{\rm peak},\,i})/2(0.05)^2]$). Knowing the instantaneous bolometric luminosity L and peak luminosity $L_{\rm peak}$, we then follow an identical procedure to determine the joint distribution $P(X|L,\ L_{\rm peak})$ of each subsequent quantity X, from simulations with similar L and $L_{\rm peak}$. We have compared this with Monte Carlo realizations based on our fitted probability distributions in this paper, and find that essentially identical results are achieved for, e.g., the distribution of L and $L_{\rm peak}$, and column densities in phases of growth not near peak luminosity. However, this modeling is not identical for, e.g., the distribution of Eddington ratios and column densities around $L \sim L_{\rm peak}$, which reflects the fact that our fits to the Eddington ratio distribution (§ 5) are rough and that our fits to the column density distribution do not apply to the final blowout phase of quasar evolution (as discussed in detail in § 4).

For each mock quasar, we generate a peak luminosity, final (postmerger) black hole mass, instantaneous bolometric luminosity, intrinsic (unattenuated) B-band (νL_ν at $\nu = 4400$ Å), soft X-ray (0.5–2 keV), and hard X-ray (2–10 keV) luminosities, observed (attenuated using the generated column density and the reddening/dust extinction modeling described in § 2.2, with SMC-like reddening curves and extinction following, e.g., Pei 1992; Morrison & McCammon 1983) B-band, soft X-ray, and hard X-ray luminosities, column density of neutral hydrogen, column density of neutral+ionized hydrogen, and instantaneous black hole mass. The intrinsic luminosities in each band are calculated using the bolometric corrections described in Marconi et al. (2004), which account for the luminosity dependence of the optical–to–X-ray luminosity ratio $\alpha_{\rm ox}$ (as discussed in § 3.2), and then attenuated to give the observed luminosities. We also provide intrinsic and attenuated luminosities in each wave band using the constant bolometric corrections of Elvis et al. (1994), but we caution that these are not calculated in a completely self-consistent manner, as our assumed bolometric luminosity function to which we fit the $\dot{n}(L_{\rm peak})$ distribution is based on using the luminosity-dependent bolometric corrections. We do not directly calculate Eddington ratios, as these are defined differently in many observed samples, but they should be calculable with the given luminosities and black hole masses.

We calculate these quantities for a mock sample of $\sim 10^9$ quasars at each redshift $z = 0.2$, 0.5, 1, 2, and 3. Most of these quasars are at luminosities orders of magnitude below those observed; therefore for space considerations and because our predictions become uncertain at low luminosities, we retain only the 10^6 quasars with brightest bolometric luminosities at each redshift.

[4] See ftp://cfa-ftp.harvard.edu/pub/phopkins/qso_catalogs.

TABLE 1
Mock Quasar Distribution Examples Mock Quasar Distribution Examples

$L_{\rm peak}$ (1)	$M^f_{\rm BH}$ (2)	L (3)	$M_{\rm BH}$ (4)	$N_{\rm H}$ (5)	$N_{\rm H\,I}$ (6)	L^i_B (7)	$L^i_{\rm SX}$ (8)	$L^i_{\rm HX}$ (9)	$L^{\rm obs}_B$ (10)	$L^{\rm obs}_{\rm SX}$ (11)	$L^{\rm obs}_{\rm HX}$ (12)
10.6.....	6.1	8.5	6.1	20.5	20.1	7.2, 7.5	7.4, 6.8	7.6, 7.0	7.2, 7.4	7.4, 6.8	7.6, 7.0
10.4.....	6.4	8.7	6.0	22.2	22.0	7.4, 7.6	7.6, 7.0	7.8, 7.2	5.8, 6.0	7.5, 6.9	7.8, 7.2
...	...	...	...	...	...	...	...	...	...	...	...

Notes.—Col. (1): Peak quasar bolometric luminosity, $\log(L_{\rm peak}/L_\odot)$. Col. (2): Final (postmerger) black hole mass, $\log(M^f_{\rm BH}/M_\odot)$. Col. (3): Current (at time of "observation") intrinsic (no attenuation) bolometric luminosity, $\log(L/L_\odot)$. Col. (4): Current black hole mass, $\log(M_{\rm BH}/M_\odot)$. Col. (5): Total (neutral+ionized) hydrogen column density along the "observed" sight line, $\log(N_{\rm H}/{\rm cm}^{-2})$. Col. (6): Neutral hydrogen column density along the "observed" sight line, $\log(N_{\rm H\,I}/{\rm cm}^{-2})$. Col. (7): Intrinsic (no attenuation) B-band luminosity, $\log(L^i_B/L_\odot)$, where $L_B = \nu_B L_{\nu_B}$ at $\nu_B = 4400$ Å. Calculated with the luminosity-dependent bolometric corrections from Marconi et al. (2004; left) and constant (luminosity independent) $L = 11.8L_B$ (Elvis et al. 1994; right). Col. (8): Intrinsic soft X-ray (0.5–2 keV) luminosity, $\log(L^i_{\rm SX}/L_\odot)$, calculated with the luminosity-dependent bolometric corrections from Marconi et al. (2004; left) and constant (luminosity independent) $L = 52.5L_{\rm SX}$ (Elvis et al. 1994; right). Col. (9): Intrinsic hard X-ray (2–10 keV) luminosity, $\log(L^i_{\rm HX}/L_\odot)$, calculated with the luminosity-dependent bolometric corrections from Marconi et al. (2004; left) and constant (luminosity independent) $L = 35.0L_{\rm HX}$ (Elvis et al. 1994; right). Col. (10): "Observed" (with attenuation) B-band luminosity, $\log(L^{\rm obs}_B/L_\odot)$. Left and right use luminosity-dependent and luminosity-independent bolometric corrections, respectively, as L^i_B. Col. (11): "Observed" soft X-ray luminosity, $\log(L^{\rm obs}_{\rm SX}/L_\odot)$. Left and right use luminosity-dependent and luminosity-independent bolometric corrections, respectively, as $L^i_{\rm SX}$. Col. (12): "Observed" hard X-ray luminosity, $\log(L^{\rm obs}_{\rm HX}/L_\odot)$. Left and right use luminosity-dependent and luminosity-independent bolometric corrections, respectively, as $L^i_{\rm HX}$. The complete tables can be downloaded at ftp://cfa-ftp.harvard.edu/pub/phopkins/qso_catalogs.

This introduces some uncertainty in our statistics at the lowest luminosities in any given band, but these luminosities are generally still well below those observed in most samples. At any luminosity, but especially at the brightest luminosities, there is also a significant amount of effective "noise" owing to our incomplete sampling of the enormous parameter space of possible mergers, and decreasing total time across simulations spent at large luminosities, which can be estimated from, e.g., Figures 8 and 17. Finally, at each redshift, we generate two distributions, reflecting the $\sim 1\ \sigma$ range in $\dot{n}(L_{\rm peak})$ and roughly parameterizing the degeneracies in our fit to the observed luminosity functions and uncertainty in the faint end of $\dot{n}(L_{\rm peak})$; "fit 1" has a lower L_* [lower peak in $\dot{n}(L_{\rm peak})$], with a larger σ_* [broader $\dot{n}(L_{\rm peak})$ distribution], and "fit 2" has a higher L_* and smaller σ_* [more narrowly peaked $\dot{n}(L_{\rm peak})$ distribution]. We show a few example "quasars" from our $z = 0.2$ mock distribution in Table 1, to demonstrate the format and units used.

8.5. *Starburst Galaxies*

Although we do not yet model the reradiation of absorbed light by dust or the contribution of stellar light to quasar host IR luminosities, including these in our picture for quasar evolution will enable us to predict luminosity functions in the IR and submillimeter and their evolution with redshift. We can at this point, however, estimate whether our model of quasar lifetimes and merger-driven evolution with $\dot{n}(L_{\rm peak})$ is consistent with the observed distribution of ultraluminous infrared galaxies. Naively, we might expect that since the obscured quasar phase has a duration up to $\sim$10 times that of the optically observable quasar phase, there should be $\sim$10 times as many ULIRGs as bright optical QSOs. But, this neglects the complicated, luminosity-dependent nature of quasar lifetimes.

Given that the bright quasars we simulate attain, during their peak growth phase, an intrinsic luminosity comparable to that of the host starburst, and that this period of peak growth has a similar duration to the starburst phase (see Fig. 13 and Di Matteo et al. 2005; Springel et al. 2005b), we can estimate (roughly) the ULIRG bolometric luminosity function from our bolometric quasar luminosity function. Thus, the more accurate comparison to the ULIRG luminosity function is with the hard X-ray quasar luminosity function, as this recovers (and at some luminosities can be dominated by) "buried" quasars in starburst phases. This is only applicable *above* the break in the luminosity function, where quasars are undergoing peak quasar growth. Below the break, quasars are, on average, sub-Eddington and can have luminosities well below that of their star-forming hosts (see Fig. 13), so we expect our quasar luminosity function to be significantly shallower than the ULIRG luminosity function at these luminosities. Note also that this does not imply that ULIRGs are all AGN-dominated, as the starburst and peak AGN activity can be (and generally are) somewhat offset, but only says that the lifetime curves at the bright end should be similar.

Considering the luminosity function at $z = 0.15$, then, we expect ULIRG densities $d\Phi/dM_{\rm bol} \sim 3\times10^{-7}$ and 9×10^{-8} Mpc^{-3} mag^{-1} at $L \sim 1.6\times10^{12}$ and $2.5\times10^{12}\ L_\odot$, respectively. These estimates are consistent with the observed density in the *IRAS* 1 Jy Survey (Kim & Sanders 1998) at a mean redshift $z = 0.15$, with $d\Phi/dM_{\rm bol} \sim 5\times10^{-7}$ and 7×10^{-8} Mpc^{-3} mag^{-1} (rescaled to our cosmology), and as expected, our quasar luminosity function slope becomes significantly shallower than the observed 1 Jy survey luminosity function slope below $L \sim 10^{11}$–$10^{12}\ L_\odot$, roughly the break luminosity of the luminosity function. We predict these densities to change with redshift according to the evolution of $\dot{n}(L_{\rm peak})$, decreasing by a factor of $\sim$1.5 at $z = 0.04$, in good agreement with the evolution of *IRAS* ULIRG luminosity functions (Kim & Sanders 1998). Likewise, at $z \sim 1$–3, we predict a mean space density $\Phi(L > 10^{11}\ L_\odot) \sim (1$–$3)\times10^5$ Mpc^{-3}, in agreement with the $\sim 5\times10^5$ Mpc^{-3} density of such sources expected to reproduce the observed cumulative source density 4×10^4 deg^{-2} of 1 mJy 850 μm SCUBA sources (Barger et al. 1999). Furthermore, our prediction of the fraction of buried AGNs and its evolution with redshift agrees well with determinations from X-ray samples (Barger et al. 2005) and recent *Spitzer* results in the mid- and near-infrared at $z \sim 2$ (Martinez-Sansigre et al. 2005).

8.6. *AGNs Not Triggered by Mergers*

Some low-redshift quasars (e.g., Bahcall et al. 1996) and many nearby, low-luminosity Seyferts appear to reside in ordinary, relatively undisturbed galaxies. Our picture for quasar evolution does not immediately account for these objects because we suppose that nuclear activity is mainly triggered by tidal torques during a merger.

This work is primarily concerned with the origin of the majority of the mass in spheroids and supermassive black holes, and as a consequence, the relation of this to the abundance and

evolution of quasars and the cosmic X-ray background. Based on our present analysis, we believe that a merger-driven picture can account for the main part of each of these and, as described earlier, that the most relevant phase in the history of the universe to these phenomena appears to have been at moderate redshifts, $z \sim 2.5-0.5$.

Our model does not exclude other mechanisms for triggering AGNs, and it is likely that a variety of stochastic or continuous processes are relevant to nuclear activity in undisturbed disks and residual low-level accretion in relaxed systems. This is not contrary to our picture because most of the total black hole mass density in the universe is in spheroid-dominated systems. The principal requirement of our model is that AGN activity in undisturbed galaxies should not contribute a large fraction of the black hole mass density in the universe, to avoid spoiling tight correlations between the black hole and host galaxy properties and producing too large a present-day black hole mass density in violation of the Soltan (1982) constraint.

For example, if a molecular cloud passed through the center of our Galaxy near Sgr A*, it is possible that the Milky Way would resemble a Seyfert for some period of time. Alternatively, it has long been recognized that mass loss from normal stellar evolution of bulge stars or stellar clusters near the centers of galaxies can provide a continuous supply of fuel for low-level accretion (e.g., McMillan et al. 1981; MacDonald & Bailey 1981; Shull 1983). Typical galactic stellar mass-loss rates [$\dot{M} \sim 1\ M_\odot\ \mathrm{yr}^{-1}\ (10^{11}\ M_\odot)^{-1}$] yield Bondi-Hoyle accretion rates $\sim 10^{-5}$ to 10^{-4} of Eddington in relaxed, dynamically hot systems; and mass-loss rates from O and W-R stars [$\dot{M} \sim 10^{-6}\ M_\odot\ \mathrm{yr}^{-1}\ (10\ M_\odot)^{-1}$] in young, dense star clusters near the centers of galaxies with sufficient cold gas for continued star formation can yield rates as high as $\sim 10^{-2}$ of Eddington.

Even though these fueling mechanisms do not involve mergers, the scenario that we have discussed might still be relevant to the origin of these black holes. Of course, the black holes and spheroids in disk-dominated systems may have been produced in a manner that did not involve mergers. Alternatively, most of the black hole mass in these objects (which is small compared to that in spheroid-dominated galaxies) could have been assembled long ago in mergers with bright quasar phases and then these "dead" quasars are resurrected sporadically by other fueling mechanisms.

Independently of how these black holes were formed, elements of our modeling may still account for certain observed properties of Seyfert galaxies. The observed Seyfert luminosity function appears to join smoothly onto the quasar luminosity function (Hao et al. 2005). It is not obvious that this would be the case if the two types of objects are produced by entirely distinct mechanisms. In addressing this, it is useful to separate the process by which gas is delivered to the black hole from the subsequent evolution that determines the observed activity. In our picture, gas is delivered to the black hole by gravitational torques during a merger, but other mechanisms, such as bar-induced fueling, may be important for objects such as Seyfert galaxies. Regardless, the induced activity may be generic, if black hole growth is self-regulated in the way we describe it in our simulations.

In Hopkins et al. (2006a) we show that the faint-end slope of the quasar luminosity function in our model is partly determined by the time dependence of the blowout phase of black hole growth. We derive an analytical model for this using a Sedov-Taylor type analysis and show that the impact of this feedback depends on the mass of the host. This analysis does not depend on the fueling mechanism, only on the subsequent evolution. If this self-regulated growth applies to Seyfert galaxies as well (for example, if Seyfert growth is regulated by a balance between accretion feedback and the spheroid potential, as expected if these objects obey a similar $M_{\rm BH}$-σ relation), we would expect the Seyfert luminosity function to smoothly join onto the quasar one, even if the fuel is delivered in a different manner.

9. CONCLUSIONS

We have studied the evolution of quasars in simulations of galaxy mergers spanning a wide region of parameter space. In agreement with earlier work (Hopkins et al. 2005d), we find that the lifetime of a particular source depends on luminosity and increases at lower luminosities, and that quasar obscuration is time-dependent. Our new, large set of simulations shows that the lifetime and obscuration can be expressed in terms of the instantaneous and peak luminosities of a quasar and that these descriptions are robust, with no systematic dependence on simulation parameters. We have combined these results with a semiempirical method to describe the cosmological distribution of quasar properties, allowing us to predict a large number of observables as a function of, for example, luminosity and redshift. This approach also makes it possible to compare our picture to simpler models for quasar lifetimes and obscuration.

In the model we examine, quasars are triggered by mergers of gas-rich galaxies, which produce inflows of gas through gravitational torquing, fueling starbursts and rapid black hole growth. The large gas densities obscure the central black hole at optical wavelengths until feedback energy from the growth of the black hole ejects gas and rapidly slows further accretion ("blowout"). Quasar lifetimes and light curves are nontrivial, with strong accretion activity during first passage of the merging galaxies and extended quiescent (sub-Eddington) phases leading into and out of the phase of peak quasar activity associated with the final merger. The blowout phase, in which the quasar is visible as a bright, near-Eddington optical source, has a lifetime related to the dynamical time in the inner regions of the merging galaxies, which characterizes the timescale over which obscuring gas and dust are expelled, but the quasar spends a longer time at lower luminosities before and after this stage. These evolutionary processes have important consequences which cannot be captured in models of pure exponential or "on/off" quasar growth.

Our work emphasizes several goals for quasar and galaxy observations and theory. Observationally, it is important to constrain the faint end of the *peak* luminosity distribution, i.e., the low-mass active black hole distribution. Unfortunately, our modeling of quasar lifetimes implies that the faint-end quasar luminosity function is dominated by quasars with peak luminosities around the break in the luminosity function, and can provide only weak constraints on the faint-end $L_{\rm peak}$ distribution. However, there is still hope, as for example broad-line quasar activity is more closely associated with near-peak luminosities, and thus probing the faint-end of broad-line luminosity functions may in particular improve the estimates. Moreover, studies of the black hole mass distribution (or the distribution of galaxy spheroids) as a function of redshift, extending to small spheroid masses/velocity dispersions, probes the faint end of $\dot{n}(L_{\rm peak})$. Other techniques, such as studies of faint radio sources at high redshift (Haiman et al. 2004), can similarly constrain these populations. Furthermore, the calculations in this paper can be combined to better determine $\dot{n}(L_{\rm peak})$, as, given a model for the quasar lifetime and obscuration, they all derive from this fundamental quantity. Additional observational tests of the modeling we have presented will provide an important means of constraining models for AGN accretion and feedback; for example, the faint-end slope of the quasar lifetime depends on how the blowout phase occurs and could provide a

sensitive probe of feedback models, enabling the adoption of more realistic and sophisticated feedback prescriptions than we have thus far employed. Of course, improved constraints on the luminosity function at all luminosities at high redshift remains a valuable means of testing theories of quasar evolution.

Our simulations are based on isolated galaxy mergers and thus do not provide a cosmological prediction for the distribution of peak luminosities $\dot{n}(L_{\rm peak})$, merger rates, or mass functions; we instead have adopted a semiempirical model, in which we use our modeling of quasar evolution to determine these distributions from the observed luminosity function. While this allows us to predict a large number of observables and to demonstrate that a wide range of quasar and galaxy properties are self-consistent in a model of merger-driven quasar activity with realistic quasar lifetimes, future theoretical work in these areas should predict the distribution of peak luminosities $\dot{n}(L_{\rm peak})$ and its evolution with redshift. These quantities are to be distinguished from the distribution of observed luminosities, as the two are not trivially related in our model or any other with a nontrivial quasar lifetime.

Although the quasar birthrate as a function of peak luminosity will be, in general, a complicated function of galaxy merger rates, gas fractions, morphologies, and other factors, we have parameterized it for comparison with the results of future cosmological simulations and semianalytical models. This distribution is particularly valuable as a theoretical quantity because it is more directly related to physical galaxy properties than even the complete (intrinsic) luminosity function, and additionally because theoretical modeling that successfully reproduces this $\dot{n}(L_{\rm peak})$ distribution is guaranteed to reproduce the large number of observable quantities we have discussed in detail in this work. We cannot determine the cosmological context in our detailed simulations of the relatively small-scale physics of galaxy mergers, and conversely, cosmological simulations and semianalytical models cannot resolve the detailed physics driving quasar activity in mergers. However, our determination of quasar evolution as a function of peak luminosity or final black hole mass can be grafted onto these cosmological models to greatly increase the effective dynamic range of such calculations. Combined with our modeling, this would remove the one significant empirical element we have adopted and allow for a complete prediction of the above quantities from a single theoretical framework.

In these efforts, we emphasize that the mergers relevant to our picture are of a specific type. First, the merging galaxies must contain a supply of cold gas in a rotationally supported disk. Hot, diffuse gas, as in the halos of elliptical galaxies, will not be subject to the gravitational torques that drive gas into galaxy centers and fuel black hole growth. Clearly, gas-poor mergers are also not important for this process. Second, the mergers will likely involve galaxies of comparable, although not necessarily equal, mass, so that the gravitational torques excited are strong enough and penetrate deep enough into galaxy centers to drive substantial inflows of gas. The precise requirement for the mass ratio is somewhat ill-defined because it also depends on the orbit geometry, but mergers with a mass ratio larger than 10 : 1 are probably not generally important to our model. Simulations of minor mergers involving galaxies with mass ratios $\lesssim$10 : 1 (e.g., Hernquist 1989; Hernquist & Mihos 1995) have shown that for particular orbital geometries, these events can produce starbursts similar to those in major mergers, leaving behind disturbed remnants with dynamically heated disks (e.g., Quinn et al. 1993; Mihos et al. 1995; Walker et al. 1996). It is of interest to establish whether black hole growth can also be triggered in minor mergers, as these events may be relevant to weak AGN activity such as that in some Seyfert galaxies or LINERs.

In summary, the work presented here supports the conjecture that many aspects of galaxy formation and evolution can be understood in terms of the "cosmic cycle" in Figure 1. To be sure, much of what is summarized in Figure 1 has been proposed elsewhere, either in the context of observations or theoretical models. Our modeling of galaxy formation and evolution emphasizes the possibility that supermassive black holes could be *responsible* for much of what goes on in shaping galaxies, rather than being bystanders, closing the loop in Figure 1. In this sense, black holes may be the "prime movers" driving galaxy evolution, as has been proposed earlier for extragalactic radio sources (e.g., Begelman et al. 1984; Rees 1984). It may seem counterintuitive that compact objects with masses much smaller than those of galaxies could have such an impact, but it is precisely the concentrated nature of matter in black holes that makes this idea plausible.

Consider a black hole of mass $M_{\rm BH}$ at the center of a spherical galaxy of mass $M_{\rm sph}$ with a characteristic velocity dispersion σ. The energy available to affect the galaxy through the growth of the black hole will be some small fraction of its rest mass, $E_{\rm feed} \sim \epsilon_f M_{\rm BH} c^2$. This can be compared with the binding energy of the galaxy, $E_{\rm bind} \sim M_{\rm sph}\sigma^2$. Observations indicate that $M_{\rm BH}$ and $M_{\rm sph}$ are correlated and that, roughly, $M_{\rm BH} \sim (0.002-0.005)M_{\rm sph}$ (Magorrian et al. 1998; Marconi & Hunt 2003). Therefore, the ratio of the feedback energy to the binding energy of the galaxy is $E_{\rm feed}/E_{\rm sph} > 10\epsilon_{f,-2}\,\sigma_{300}^{-2}$ for an assumed efficiency of 1%, $\epsilon_{f,-2} \equiv \epsilon/0.01$, and scaling the velocity dispersion to $\sigma_{300} \equiv \sigma/300$ km s^{-1}, as for relatively massive galaxies. This result demonstrates that the supermassive black holes in the centers of spheroidal galaxies are by far the largest supply of potential energy in these objects, exceeding even the galaxy binding energy. When viewed in this way, if even a small fraction of the black hole radiant energy can couple to the surrounding ISM, then black hole growth is not an implausible mechanism for regulating galaxy formation and evolution; in fact, it appears almost inevitable that it should play this role.

We thank our referee, David Weinberg, for many comments and suggestions that greatly improved this paper. We thank Paul Martini for helpful discussion and Gordon Richards and Alessandro Marconi for generously providing data for observational comparisons. This work was supported in part by NSF grants ACI 96-19019, AST 00-71019, AST 02-06299, and AST 03-07690, and NASA ATP grants NAG5-12140, NAG5-13292, and NAG5-13381. The simulations were performed at the Center for Parallel Astrophysical Computing at the Harvard-Smithsonian Center for Astrophysics.

REFERENCES

Adelberger, K. L., & Steidel, C. C. 2005, ApJ, 630, 50
Alexander, D. M., Bauer, F. E., Chapman, S. C., Smail, I., Blain, A. W., Brandt, W. N., & Ivison, R. J. 2005a, ApJ, 632, 736
Alexander, D. M., et al. 2005b, Nature, 434, 738
Allen, D. A., Roche, P. F., & Norris, R. P. 1985, MNRAS, 213, 67P
Aller, M. C., & Richstone, D. 2002, AJ, 124, 3035
Antonucci, R. 1993, ARA&A, 31, 473
Armus, L., Heckman, T. M., & Miley, G. 1987, AJ, 94, 831
Bahcall, J. N., Kirhakos, S., Saxe, D. H., & Schneider, D. P. 1997, ApJ, 479, 642
Bahcall, J. N., Kirhakos, S., & Schneider, D. P. 1994, ApJ, 435, L11
———. 1995, ApJ, 450, 486
———. 1996, ApJ, 457, 557
Bajtlik, S., Duncan, R. C., & Ostriker, J. P. 1988, ApJ, 327, 570

Baldi, A., Molendi, S., Comastri, A., Fiore, F., Matt, G., & Vignali, C. 2002, ApJ, 564, 190
Barcons, X., Mateos, S., & Ceballos, M. T. 2000, MNRAS, 316, L13
Barger, A. J., Cowie, L. L., Mushotzky, R. F., Yang, Y., Wang, W.-H., Steffen, A. T., & Capak, P. 2005, AJ, 129, 578
Barger, A. J., Cowie, L. L., & Sanders, D. B. 1999, ApJ, 518, L5
Barger, A. J., et al. 2003, AJ, 126, 632
Barnes, J. E. 1988, ApJ, 331, 699
———. 1992, ApJ, 393, 484
———. 1998, in Galaxies: Interactions and Induced Star Formation, ed. D. Friedli, L. Martinet, & D. Pfenniger (Berlin: Springer), 275
Barnes, J. E., & Hernquist, L. 1991, ApJ, 370, L65
———. 1992, ARA&A, 30, 705
———. 1996, ApJ, 471, 115
Begelman, M. C., Blandford, R. D., & Rees, M. J. 1984, Rev. Mod. Phys., 56, 255
Begelman, M. C., & Nath, B. B. 2005, MNRAS, 361, 1387
Bernardi, M., et al. 2003, AJ, 125, 1849
Bisnovatyi-Kogan, G. S., & Silich, S. A. 1991, Soviet Astron., 35, 370
Blandford, R. D., & Begelman, M. C. 1999, MNRAS, 303, L1
Bondi, H. 1952, MNRAS, 112, 195
Bondi, H., & Hoyle, F. 1944, MNRAS, 104, 273
Borys, C., Smail, I., Chapman, S. C., Blain, A. W., Alexander, D. M., & Ivison, R. J. 2005, ApJ, 635, 853
Bouchet, P., Lequeux, J., Maurice, E., Prevot, L., & Prevot-Burnichon, M. L. 1985, A&A, 149, 330
Bower, G. C., Wright, M. C. H., Falcke, H., & Backer, D. C. 2003, ApJ, 588, 331
Boyle, B. J., Fong, R., Shanks, T., & Peterson, B. A. 1990, MNRAS, 243, 1
Boyle, B. J., Jones, L. R., & Shanks, T. 1991, MNRAS, 251, 482
Boyle, B. J., Shanks, T., Croom, S. M., Smith, R. J., Miller, L., Loaring, N., & Heymans, C. 2000, MNRAS, 317, 1014
Boyle, B. J., Shanks, T., & Peterson, B. A. 1988, MNRAS, 235, 935
Brandt, W. N., & Hasinger, G. 2005, ARA&A, 43, 827
Brandt, W. N., et al. 2001, AJ, 122, 2810
Brotherton, M. S., Tran, H. D., Becker, R. H., Gregg, M. D., Laurent-Muehleissen, S. A., & White, R. L. 2001, ApJ, 546, 775
Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000
Busha, M. T., Evrard, A. E., Adams, F. C., & Wechsler, R. H. 2005, MNRAS, 363, L11
Canalizo, G., & Stockton, A. 2001, ApJ, 555, 719
Cao, X., & Xu, Y.-D. 2005, ApJ, 629, 897
Cen, R., Miralda-Escudé, J., Ostriker, J. P., & Rauch, M. 1994, ApJ, 437, L9
Ciotti, L., & Ostriker, J. P. 1997, ApJ, 487, L105
———. 2001, ApJ, 551, 131
Cirasuolo, M., Shankar, F., Granato, G. L., & Danese, L. 2005, ApJ, 629, 816
Comastri, A., Setti, G., Zamorani, G., & Hasinger, G. 1995, A&A, 296, 1
Combes, F. 2006, in Astrophysics Update 2 (Berlin: Springer), in press (astro-ph/0505463)
Conselice, C. J., Bershady, M. A., Dickinson, M., & Papovich, C. 2003, AJ, 126, 1183
Cowie, L. L., Barger, A. J., Bautz, M. W., Brandt, W. N., & Garmire, G. P. 2003, ApJ, 584, L57
Cox, T. J., Di Matteo, T., Hernquist, L., Hopkins, P. F., Robertson, B., & Springel, V. 2005, ApJL, submitted (astro-ph/0504156)
Croft, R. A. C., Weinberg, D. H., Bolte, M., Burles, S. Hernquist, L., Katz, N., Kirkman, D., & Tytler, D. 2002, ApJ, 581, 20
Croft, R. A. C., Weinberg, D. H., Katz, N., & Hernquist, L. 1998, ApJ, 495, 44
Croft, R. A. C., Weinberg, D. H., Pettini, M., Hernquist, L., & Katz, N. 1999, ApJ, 520, 1
Croom, S. M., Smith, R. J., Boyle, B. J., Shanks, T., Miller, L., Outram, P. J., & Loaring, N. S. 2004, MNRAS, 349, 1397
Croom, S. M., et al. 2005, MNRAS, 356, 415
Croton, D. J., et al. 2006, MNRAS, 365, 11
Davé, R., Hernquist, L., Katz, N., & Weinberg, D. H. 1999, ApJ, 511, 521
Dekel, A., & Birnboim, Y. 2006, MNRAS, in press (astro-ph/0412300)
Di Matteo, T., Carilli, C. L., & Fabian, A. C. 2001, ApJ, 547, 731
Di Matteo, T., Esin, A., Fabian, A. C., & Narayan, R. 1999, MNRAS, 305, L1
Di Matteo, T., Quataert, E., Allen, S. W., Narayan, R., & Fabian, A. C. 2000, MNRAS, 311, 507
Di Matteo, T., Springel, V., & Hernquist, L. 2005, Nature, 433, 604
di Serego Alighieri, S., et al. 2005, A&A, 442, 125
Donley, J. L., Rieke, G. H., Rigby, J. R., & Perez-Gonzalez, P. G. 2005, ApJ, 634, 169
Ellison, S. L., Hall, P. B., & Lira, P. 2005, AJ, 130, 1345
Elvis, M., Risaliti, G., & Zamorani, G. 2002, ApJ, 565, L75
Elvis, M., et al. 1994, ApJS, 95, 1
Esin, A. A., McClintock, J. E., & Narayan, R. 1997, ApJ, 489, 865
Fabian, A. C. 1999, MNRAS, 308, L39
Fabian, A. C., & Canizares, C. R. 1988, Nature, 333, 829
Fabian, A. C., & Iwasawa, K. 1999, MNRAS, 303, L34
Fan, X., et al. 2001, AJ, 121, 54
———. 2004, AJ, 128, 515
Ferrarese, L. 2002, in Current High-Energy Emission around Black Holes, ed. C.-H. Lee & H.-Y. Chang (Singapore: World Scientific), 3
Ferrarese, L., & Merritt, D. 2000, ApJ, 539, L9
Fiore, F., et al. 2003, A&A, 409, 79
Francis, P. J., Drake, C. L., Whiting, M. T., Drinkwater, M. J., & Webster, R. L. 2001, Publ. Astron. Soc. Australia, 18, 221
Gebhardt, K., et al. 2000, ApJ, 539, L13
———. 2003, ApJ, 597, 239
George, I. M., Turner, T. J., Netzer, H., Nandra, K., Mushotzky, R. F., & Yaqoob, T. 1998, ApJS, 114, 73
Gerssen, J., et al. 2004, AJ, 127, 75
Giacconi, R., et al. 2002, ApJS, 139, 369
Gilli, R., Risaliti, G., & Salvati, M. 1999, A&A, 347, 424
Gilli, R., Salvati, M., & Hasinger, G. 2001, A&A, 366, 407
Granato, G. L., & Danese, L. 1994, MNRAS, 268, 235
Granato, G. L., De Zotti, G., Silva, L., Bressan, A., & Danese, L. 2004, ApJ, 600, 580
Grazian, A., Negrello, M., Moscardini, L., Cristiani, S., Haehnelt, M. G., Matarrese, S., Omizzolo, A., & Vanzella, E. 2004, AJ, 127, 592
Green, P. J., et al. 1995, ApJ, 450, 51
Gregg, M. D., Lacy, M., White, R. L., Glikman, E., Helfand, D., Becker, R. H., & Brotherton, M. S. 2002, ApJ, 564, 133
Grimes, J. A., Rawlings, S., & Willott, C. J. 2004, MNRAS, 349, 503
Gruber, D. E., Matteson, J. L., Peterson, L. E., & Jung, G. V. 1999, ApJ, 520, 124
Haehnelt, M. G., Natarajan, P., & Rees, M. J. 1998, MNRAS, 300, 817
Haiman, Z., & Cen, R. 2002, ApJ, 578, 702
Haiman, Z., & Loeb, A. 1998, ApJ, 503, 505
Haiman, Z., & Menou, K. 2000, ApJ, 531, 42
Haiman, Z., Quataert, E., & Bower, G. C. 2004, ApJ, 612, 698
Hamilton, T. S., Casertano, S., & Turnshek, D. A. 2002, ApJ, 576, 61
Hao, L., et al. 2005, AJ, 129, 1795
Hasinger, G. 2002, in New Visions of the X-Ray Universe in the *XMM-Newton* and *Chandra* Era, ed. F. Jansen (ESA SP-488; Noordwijk: ESA), in press (astro-ph/0202430)
———. 2004, Nucl. Phys. B Proc. Supp., 132, 86
Hasinger, G., Miyaji, T., & Schmidt, M. 2005, A&A, 441, 417
Hasinger, G., et al. 2001, A&A, 365, L45
Heckman, T. M., Bothun, G. D., Balick, B., & Smith, E. P. 1984, AJ, 89, 958
Heckman, T. M., Kauffmann, G., Brinchmann, J., Charlot, S., Tremonti, C., & White, S. D. M. 2004, ApJ, 613, 109
Hennawi, J. F., et al. 2006, AJ, 131, 1
Hernquist, L. 1989, Nature, 340, 687
———. 1990, ApJ, 356, 359
———. 1992, ApJ, 400, 460
———. 1993a, ApJ, 409, 548
———. 1993b, ApJ, 404, 717
Hernquist, L., Katz, N., Weinberg, D., & Miralda-Escudé, J. 1996, ApJ, 457, L51
Hernquist, L., & Mihos, J. C. 1995, ApJ, 448, 41
Hernquist, L., Spergel, D. N., & Heyl, J. S. 1993, ApJ, 416, 415
Hill, G. J., Goodrich, R. W., & DePoy, D. L. 1996, ApJ, 462, 163
Ho, L. C. 2002, ApJ, 564, 120
Holden, B. P., et al. 2005, ApJ, 620, L83
Hopkins, P. F., Hernquist, L., Cox, T. J., Robertson, B., Di Matteo, T., Martini, P., & Springel, V. 2005a, ApJ, 630, 705
Hopkins, P. F., Hernquist, L., Cox, T. J., Robertson, B., Di Matteo, T., & Springel, V. 2005b, ApJ, 630, 716
———. 2005c, ApJ, 632, 81
———. 2006a, ApJ, 639, 700
Hopkins, P. F., Hernquist, L., Cox, T. J., Robertson, B., & Springel, V. 2006b, ApJS, 163, 50
Hopkins, P. F., Hernquist, L., Martini, P., Cox, T. J., Robertson, B., Di Matteo, T., & Springel, V. 2005d, ApJ, 625, L71
Hopkins, P. F., et al. 2004, AJ, 128, 1112
Hoyle, F., & Lyttleton, R. A. 1939, Proc. Cambridge Philos. Soc., 35, 405
Hui, L., Burles, S., Seljak, U., Rutledge, R. E., Magnier, E., & Tytler, D. 2001, ApJ, 552, 15
Hunt, M. P., Steidel, C. C., Adelberger, K. L., & Shapley, A. E. 2004, ApJ, 605, 625
Hutchings, J. B., & Neff, S. G. 1992, AJ, 104, 1
Jakobsen, P., Jansen, R. A., Wagner, S., & Reimers, D. 2003, A&A, 397, 891
Jester, S. 2005, ApJ, 625, 667

Jogee, S. 2006, in AGN Physics on All Scales, ed. D. Alloin, R. Johnson, & P. Lira (Springer: Berlin), in press (astro-ph/0408383)
Jonsson, P., Cox, T. J., Primack, J. R., & Somerville, R. S. 2006, ApJ, 637, 255
Joseph, R. D., & Wright, G. S. 1985, MNRAS, 214, 87
Katz, N., Weinberg, D. H., & Hernquist, L. 1996a, ApJS, 105, 19
Katz, N., Weinberg, D. H., Hernquist, L., & Miralda-Escudé, J. 1996b, ApJ, 457, L57
Kauffmann, G., & Haehnelt, M. 2000, MNRAS, 311, 576
Kawata, D., & Gibson, B. K. 2005, MNRAS, 358, L16
Kennefick, J. D., Djorgovski, S. G., & De Carvalho, R. R. 1995, AJ, 110, 2553
Kim, D. C., & Sanders, D. B. 1998, ApJS, 119, 41
Kleinmann, S. G., et al. 1988, ApJ, 328, 161
Kochanek, C. S., et al. 2001, ApJ, 560, 566
Kollmeier, J. A., et al. 2006, ApJ, in press (astro-ph/0508657)
Komossa, S., Burwitz, V., Hasinger, G., Predehl, P., Kaastra, J. S., & Ikebe, Y. 2003, ApJ, 582, L15
Koo, B.-C., & McKee, C. F. 1990, ApJ, 354, 513
Koo, D. C., & Kron, R. G. 1988, ApJ, 325, 92
Kormendy, J., & Gebhardt, K. 2001, in AIP Conf. Proc. 586, 20th Texas Symposium on Relativistic Astrophysics, ed. J. C. Wheeler & H. Martel (Melville: AIP), 363
Kormendy, J., & Richstone, D. 1995, ARA&A, 33, 581
Kuraszkiewicz, J., Wilkes, B. J., Brandt, W. N., & Vestergaard, M. 2000, ApJ, 542, 631
La Franca, F., et al. 2002, ApJ, 570, 100
———. 2005, ApJ, 635, 864
Lawrence, A. 1991, MNRAS, 252, 586
Lidz, A., Hopkins, P. F., Cox, T. J., Hernquist, L., & Robertson, B. 2006, ApJ, in press (astro-ph/0507361)
Loewenstein, M., Mushotzky, R. F., Angelini, L., Arnaud, K. A., & Quataert, E. 2001, ApJ, 555, L21
Lynden-Bell, D. 1969, Nature, 223, 690
Maccarone, T. J. 2003, A&A, 409, 697
MacDonald, J., & Bailey, M. E. 1981, MNRAS, 197, 995
Madau, P., Ghisellini, G., & Fabian, A. C. 1994, MNRAS, 270, L17
Magdziarz, P., & Zdziarski, A. A. 1995, MNRAS, 273, 837
Magorrian, J., et al. 1998, AJ, 115, 2285
Mainieri, V., et al. 2005, A&A, 437, 805
Marano, B., Zamorani, G., & Zitelli, V. 1988, MNRAS, 232, 111
Marchesini, D., Celotti, A., & Ferrarese, L. 2004, MNRAS, 351, 733
Marconi, A., & Hunt, L. K. 2003, ApJ, 589, L21
Marconi, A., Risaliti, G., Gilli, R., Hunt, L. K., Maiolino, R., & Salvati, M. 2004, MNRAS, 351, 169
Marconi, A., & Salvati, M. 2002, in ASP Conf. Ser. 258, Issues in Unification of Active Galactic Nuclei, ed. R. Maiolino, A. Marconi, & N. Nagar (San Francisco: ASP), 217
Marshall, F. E., Boldt, E. A., Holt, S. S., Miller, R. B., Mushotzky, R. F., Rose, L. A., Rothschild, R. E., & Serlemitsos, P. J. 1980, ApJ, 235, 4
Martinez-Sansigre, A., Rawlings, S., Lacy, M., Fadda, D., Marleau, F. R., Simpson, C., Willott, C. J., & Jarvis, M. J. 2005, Nature, 436, 666
Martini, P. 2004, in Coevolution of Black Holes and Galaxies, ed. L. C. Ho (Cambridge: Cambridge Univ. Press), 170
Martini, P., & Schneider, D. P. 2003, ApJ, 597, L109
Marzke, R. O., Geller, M. J., Huchra, J. P., & Corwin, H. G. 1994, AJ, 108, 437
Max, C. E., et al. 2005, ApJ, 621, 738
McDonald, P., Miralda-Escudé, J., Rauch, M., Sargent, W. L., Barlow, T. A., Cen, R., & Ostriker, J. P. 2000, ApJ, 543, 1
McDonald, P., et al. 2005, ApJ, 635, 761
McKee, C. F., & Ostriker, J. P. 1977, ApJ, 218, 148
McLure, R. J., & Dunlop, J. S. 2002, MNRAS, 331, 795
———. 2004, MNRAS, 352, 1390
McMillan, S. L. W., Lightman, A. P., & Cohn, H. 1981, ApJ, 251, 436
Melnick, J., & Mirabel, I. F. 1990, A&A, 231, L19
Merloni, A. 2004, MNRAS, 353, 1035
Meyer, F., Liu, B. F., & Meyer-Hofmeister, E. 2000, A&A, 354, L67
Mihos, J. C., & Hernquist, L. 1994, ApJ, 431, L9
———. 1996, ApJ, 464, 641
Mihos, J. C., Walker, I. R., Hernquist, L., Mendes de Oliveira, C., & Bolte, M. 1995, ApJ, 447, L87
Miyaji, T., Hasinger, G., & Schmidt, M. 2000, A&A, 353, 25
———. 2001, A&A, 369, 49
Mo, J. J., Mao, S., & White, S. D. M. 1998, MNRAS, 295, 319
Morrison, R., & McCammon, D. 1983, ApJ, 270, 119
Mulchaey, J. S., Koratkar, A., Ward, M. J., Wilson, A. S., Whittle, M., Antonucci, R. R. J., Kinney, A. L., & Hurt, T. 1994, ApJ, 436, 586
Murray, N., Quataert, E., & Thompson, T. A. 2005, ApJ, 618, 569
Nakamura, O., Fukugita, M., Yasuda, N., Loveday, J., Brinkmann, J., Schneider, D. P., Shimasaku, K., & SubbaRao, M. 2003, AJ, 125, 1682
Narayan, R., Igumenshchev, I. V., & Abramowicz, M. A. 2000, ApJ, 539, 798
Narayan, R., & Yi, I. 1995, ApJ, 452, 710
Navarro, J. F., Frenk, C. S., & White, S. D. M. 1996, ApJ, 462, 563
Nelan, J. E., Smith, R. J., Hudson, M. J., Wegner, G. A., Lucey, J. R., Moore, S. A. W., Quinney, S. J., & Suntzeff, N. B. 2005, ApJ, 632, 137
O'Dowd, M., Urry, C. M., & Scarpa, R. 2002, ApJ, 580, 96
O'Shea, B. W., Nagamine, K., Springel, V., Hernquist, L., & Norman, M. L. 2005, ApJS, 160, 1
Ostriker, J. P., & Ciotti, L. 2005, Philos. Trans. R. Soc. London, A, 363, 667
Page, M. J., & Carrera, F. J. 2000, MNRAS, 311, 433
Page, M. J., Mason, K. O., McHardy, I. M., Jones, L. R., & Carrera, F. J. 1997, MNRAS, 291, 324
Page, M. J., Stevens, J. A., Ivison, R. J., & Carrera, F. J. 2004, ApJ, 611, L85
Peacock, J. A. 1999, Cosmological Physics (Cambridge: Cambridge Univ. Press)
Pei, Y. C. 1992, ApJ, 395, 130
Perlmutter, S., et al. 1999, ApJ, 517, 565
Perola, G. C., Matt, G., Cappi, M., Fiore, F., Guainazzi, M., Maraschi, L., Petrucci, P. O., & Piro, L. 2002, A&A, 389, 802
Pompilio, F., La Franca, F., & Matt, G. 2000, A&A, 353, 440
Porciani, C., Magliocchetti, M., & Norberg, P. 2004, MNRAS, 355, 1010
Ptak, A., Heckman, T., Levenson, N. A., Weaver, K., & Strickland, D. 2003, ApJ, 592, 782
Quataert, E., & Gruzinov, A. 2000, ApJ, 539, 809
Quinn, P. J., Hernquist, L., & Fullager, D. P. 1993, ApJ, 403, 74
Rees, M. J. 1984, ARA&A, 22, 471
Richards, G. T., et al. 2001, AJ, 121, 2308
———. 2003, AJ, 126, 1131
———. 2005, MNRAS, 360, 839
Richstone, D., et al. 1998, Nature, 395, A14
Riess, A., et al. 1998, AJ, 116, 1009
———. 2000, ApJ, 536, 62
Rines, K., Geller, M. J., Diaferio, A., Kurtz, M. J., & Jarrett, T. H. 2004, AJ, 128, 1078
Rines, K., Geller, M. J., Diaferio, A., Mahdavi, A., Mohr, J. J., & Wegner, G. 2002, AJ, 124, 1266
Rines, K., Geller, M. J., Diaferio, A., Mohr, J. J., & Wegner, G. A. 2000, AJ, 120, 2338
Rines, K., Geller, M. J., Kurtz, M. J., & Diaferio, A. 2003, AJ, 126, 2152
Risaliti, G., & Elvis, M. 2005, ApJ, 629, L17
Risaliti, G., Maiolino, R., & Salvati, M. 1999, ApJ, 522, 157
Robertson, B., Cox, T. J., Hernquist, L., Franx, M., Hopkins, P. F., Martini, P., & Springel, V. 2006, ApJ, in press (astro-ph/0511053)
Robertson, B., Hernquist, L., Bullock, J. S., Cox, T. J., Di Matteo, T., Springel, V., & Yoshida, N. 2005a, ApJL, submitted (astro-ph/0503369)
Robertson, B., Hernquist, L., Cox, T. J., Di Matteo, T., Hopkins, P. F., Martini, P., & Springel, V. 2005b, ApJ, submitted (astro-ph/0506038)
Robertson, B., Yoshida, N., Springel, V., & Hernquist, L. 2004, ApJ, 606, 32
Rosati, P., et al. 2002, ApJ, 566, 667
Rusin, D., et al. 2003, ApJ, 587, 143
Salpeter, E. E. 1964, ApJ, 140, 796
Salucci, P., Szuszkiewicz, E., Monaco, P., & Danese, L. 1999, MNRAS, 307, 637
Sanders, D. B., & Mirabel, I. F. 1996, ARA&A, 34, 749
Sanders, D. B., et al. 1986, ApJ, 305, L45
———. 1988a, ApJ, 325, 74
———. 1988b, ApJ, 324, L55
———. 1988c, ApJ, 328, L35
Sargent, A. I., Sanders, D. B., & Phillips, T. G. 1989, ApJ, 346, L9
Sargent, A. I., et al. 1987, ApJ, 312, L35
Sazonov, S. Y., Ostriker, J. P., Ciotti, L., & Sunyaev, R. A. 2005, MNRAS, 358, 168
Sazonov, S. Y., & Revnivtsev, M. G. 2004, A&A, 423, 469
Scannapieco, E., & Oh, S. P. 2004, ApJ, 608, 62
Schartmann, M., Meisenheimer, K., Camenzind, M., Wolf, S., & Henning, T. 2005, A&A, 437, 861
Schmidt, M., Schneider, D. P., & Gunn, J. 1995, AJ, 110, 68
Schweizer, F. 1998, in Galaxies: Interactions and Induced Star Formation, ed. R. C. Kennicutt, Jr., et al. (Berlin: Springer), 105
Scoville, N. Z., et al. 1986, ApJ, 311, L47
Setti, G., & Woltjer, L. 1989, A&A, 224, L21
Shakura, N. I., & Sunyaev, R. A. 1973, A&A, 24, 337
Shankar, F., Salucci, P., Granato, G. L., De Zotti, G., & Danese, L. 2004, MNRAS, 354, 1020
Sheth, R. K., et al. 2003, ApJ, 594, 225
Shull, J. M. 1983, ApJ, 264, 446
Silk, J., & Rees, M. J. 1998, A&A, 331, L1
Simpson, C. 1998, MNRAS, 297, L39
———. 2005, MNRAS, 360, 565

Simpson, C., & Rawlings, S. 2000, MNRAS, 317, 1023
Simpson, C., Rawlings, S., & Lacy, M. 1999, MNRAS, 306, 828
Small, T. A., & Blandford, R. D. 1992, MNRAS, 259, 725
Soifer, B. T., Houck, J. R., & Neugebauer, G. 1987, ARA&A, 25, 187
Soifer, B. T., et al. 1984a, ApJ, 278, L71
———. 1984b, ApJ, 283, L1
Soltan, A. 1982, MNRAS, 200, 115
Somerville, R. S., Primack, J. R., & Faber, S. M. 2001, MNRAS, 320, 504
Spergel, D. N., et al. 2003, ApJS, 148, 175
Springel, V. 2005, MNRAS, 364, 1105
Springel, V., Di Matteo, T., & Hernquist, L. 2005a, ApJ, 620, L79
———. 2005b, MNRAS, 361, 776
Springel, V., & Hernquist, L. 2002, MNRAS, 333, 649
———. 2003, MNRAS, 339, 289
———. 2005, ApJ, 622, L9
Springel, V., Yoshida, N., & White, S. D. M. 2001, NewA, 6, 79
Steed, A., & Weinberg, D. H. 2003, preprint (astro-ph/0311312)
Steffen, A. T., Barger, A. J., Cowie, L. L., Mushotzky, R. F., & Yang, Y. 2003, ApJ, 596, L23
Stevens, J. A., Page, M. J., Ivison, R. J., Carrera, F. J., Mittaz, J. P. D., Smail, I., & McHardy, I. M. 2005, MNRAS, 360, 610
Stockton, A. 1978, ApJ, 223, 747
Stockton, A., & MacKenty, J. 1987, ApJ, 316, 584
Stockton, A., & Ridgway, S. 1991, AJ, 102, 488
Strateva, I., Brandt, N., Schneider, D. P., Vanden Berk, D. G., & Vignali, C. 2005, AJ, 130, 387
Szokoly, G. P., et al. 2004, ApJS, 155, 271
Telfer, R. C., Zheng, W., Kriss, G. A., & Davidsen, A. F. 2002, ApJ, 565, 773
Toomre, A. 1977, in Evolution of Galaxies and Stellar Populations, ed. B. M. Tinsley & R. B. Larson (New Haven: Yale Univ. Obs.), 401
Toomre, A., & Toomre, J. 1972, ApJ, 178, 623
Tran, H. D. 2003, ApJ, 583, 632
Treister, E., & Urry, C. M. 2005, ApJ, 630, 115
Treister, E., et al. 2004, ApJ, 616, 123
Tremaine, S., et al. 2002, ApJ, 574, 740
Treu, T., Ellis, R. S., Liao, T. X., & van Dokkum, P. G. 2005, ApJ, 622, L5
Treu, T., Stiavelli, M., Bertin, G., Casertano, S., & Moller, P. 2001, MNRAS, 326, 237
Treu, T., Stiavelli, M., Casertano, S., Moller, P., & Bertin, G. 2002, ApJ, 564, L13
Tully, R. B., Somerville, R. S., Trentham, N., & Verheijen, M. A. W. 2002, ApJ, 569, 573
Ueda, Y., Akiyama, M., Ohta, K., & Miyaji, T. 2003, ApJ, 598, 886
Vanden Berk, D. E., et al. 2001, AJ, 122, 549
van der Wel, A., Franx, M., van Dokkum, P. G., Rix, H. W., Illingworth, G. D., & Rosati, P. 2005, ApJ, 631, 145
van de Ven, G., van Dokkum, P. G., & Franx, M. 2003, MNRAS, 344, 924
van Dokkum, P. G. 2005, AJ, 130, 2647
van Dokkum, P. G., Franx, M., Kelson, D. D., & Illingworth, G. D. 2001, ApJ, 553, L39
van Dokkum, P. G., & Stanford, S. A. 2003, ApJ, 585, 78
Vestergaard, M. 2004, ApJ, 601, 676
Viel, M., Haehnelt, M. G., & Springel, V. 2004, MNRAS, 354, 684
Viel, M., Matarrese, S., Theuns, T., Munshi, D., & Wang, Y. 2003, MNRAS, 340, L47
Vignali, C., Brandt, W. N., & Schneider, D. P. 2003, AJ, 125, 433
Vitvitska, M., Klypin, A. A., Kravtsov, A. V., Wechsler, R. H., Primack, J. R., & Bullock, J. S. 2002, ApJ, 581, 799
Volonteri, M., Haardt, F., & Madau, P. 2003, ApJ, 582, 559
Walker, I. R., Mihos, J. C., & Hernquist, L. 1996, ApJ, 460, 121
Warren, S. J., Hewett, P. C., & Osmer, P. S. 1994, ApJ, 421, 412
Webster, R. L., Francis, P. J., Peterson, B. A., Drinkwater, M. J., & Masci, F. J. 1995, Nature, 375, 469
White, R. L., Helfand, D. J., Becker, R. H., Gregg, M. D., Postman, M., Lauer, T. R., & Oegerle, W. 2003, AJ, 126, 706
White, S. D. M., & Rees, M. J. 1978, MNRAS, 183, 341
Wilkes, B. J., Tananbaum, H., Worrall, D. M., Avni, Y., Oey, M. S., & Flanagan, J. 1994, ApJS, 92, 53
Willott, C. J., Rawlings, S., Blundell, K. M., & Lacy, M. 2000, MNRAS, 316, 449
Wolf, C., Wisotzki, L., Borch, A., Dye, S., Kleinheinrich, M., & Meisenheimer, K. 2003, A&A, 408, 499
Wolf, C., et al. 2005, ApJ, 630, 771
Woo, J.-H., & Urry, C. M. 2002, ApJ, 579, 530
Wuyts, S., van Dokkum, P. G., Kelson, D. D., Franx, M., & Illingworth, G. D. 2004, ApJ, 605, 677
Wyithe, J. S. B., & Loeb, A. 2002, ApJ, 581, 886
———. 2003, ApJ, 595, 614
Xu, C. K., Sun, Y. C., & He, X. T. 2004, ApJ, 603, L73
Yu, Q., & Lu, Y. 2004, ApJ, 602, 603
———. 2005, ApJ, 620, 31
Yu, Q., & Tremaine, S. 2002, MNRAS, 335, 965
Zakamska, N. L., Strauss, M. A., Heckman, T. M., Ivezić, Ž., & Krolik, J. H. 2004, AJ, 128, 1002
Zakamska, N. L., et al. 2005, AJ, 129, 1212
Zel'dovich, Ya. B., & Novikov, I. D. 1964, Soviet Phys. Dokl., 158, 811
Zhang, Y., Anninos, P., & Norman, M. L. 1995, ApJ, 453, L57
Zitelli, V., Mignoli, M., Zamorani, G., Marano, B., & Boyle, B. J. 1992, MNRAS, 256, 349