The origins of active galactic nuclei obscuration: the ‘torus’ as a dynamical, unstable driver of accretion

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters

Citation
Hopkins, Philip F., Christopher C. Hayward, Desika Narayanan, and Lars Hernquist. 2011. “The Origins of Active Galactic Nuclei Obscuration: The ‘torus’ as a Dynamical, Unstable Driver of Accretion.” Monthly Notices of the Royal Astronomical Society 420 (1): 320–39. https://doi.org/10.1111/j.1365-2966.2011.20035.x.

Citable link
http://nrs.harvard.edu/urn-3:HUL.InstRepos:41381864

Terms of Use
This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA
The origins of active galactic nuclei obscuration: the ‘torus’ as a dynamical, unstable driver of accretion

Philip F. Hopkins,1⋆ Christopher C. Hayward,2 Desika Narayanan3 and Lars Hernquist2
1Department of Astronomy, University of California Berkeley, Berkeley, CA 94720, USA
2Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA
3Steward Observatory, University of Arizona, 933 N Cherry Ave, Tucson, AZ 85721, USA

Accepted 2011 October 19. Received 2011 October 18; in original form 2011 August 23

ABSTRACT
Recent multiscale simulations have made it possible to follow gas inflows responsible for high-Eddington ratio accretion on to massive black holes (BHs) from galactic scales to the BH accretion disc. When sufficient gas is driven towards a BH, gravitational instabilities generically form lopsided, eccentric discs that propagate inwards from larger radii. The lopsided stellar disc exerts a strong torque on the gas, driving inflows that fuel the growth of the BH. Here, we investigate the possibility that the same disc, in its gas-rich phase, is the putative ‘torus’ invoked to explain obscured active galactic nuclei (AGN) and the cosmic X-ray background. The disc is generically thick and has characteristic ∼1–10 pc sizes and masses resembling those required of the torus. Interestingly, the scale heights and obscured fractions of the predicted torii are substantial even in the absence of strong stellar feedback providing the vertical support. Rather, they can be maintained by strong bending modes and warps/twists excited by the inflow-generating instabilities. A number of other observed properties commonly attributed to ‘feedback’ processes may in fact be explained entirely by dynamical, gravitational effects: the lack of alignment between torus and host galaxy, correlations between local star formation rate (SFR) and turbulent gas velocities and the dependence of obscured fractions on AGN luminosity or SFR. We compare the predicted torus properties with observations of gas surface density profiles, kinematics, scale heights and SFR densities in AGN, and find that they are consistent in all cases. We argue that it is not possible to reproduce these observations and the observed column density distribution without a clumpy gas distribution, but allowing for simple clumping on small scales the predicted column density distribution is in good agreement with observations from \( N_{\text{H}} \sim 10^{20}–10^{27} \text{cm}^{-2} \). We examine how the \( N_{\text{H}} \) distribution scales with galaxy and AGN properties. The dependence is generally simple, but AGN feedback may be necessary to explain certain trends in obscured fraction with luminosity and/or redshift. In our paradigm, the torus is not merely a bystander or passive fuel source for accretion, but is itself the mechanism driving accretion. Its generic properties are not coincidence, but requirements for efficient accretion.

Key words: galaxies: active – galaxies: evolution – quasars: general – cosmology: theory.

1 INTRODUCTION

It has long been realized that bright, high-Eddington ratio accretion (i.e. a quasar) dominates the accumulation of mass in the supermassive black hole (BH) population (Soltan 1982; Salucci et al. 1999; Shankar et al. 2004; Hopkins, Narayan & Hernquist 2006d).

⋆E-mail: phopkins@astro.berkeley.edu

The discovery, in the past decade, of tight correlations between BH mass and host spheroid properties including mass (Kormendy & Richstone 1995; Magorrian et al. 1998), velocity dispersion (Ferrarese & Merritt 2000; Gebhardt et al. 2000) and binding energy or potential well depth (Aller & Richstone 2007; Hopkins et al. 2007b,c; Feoli & Mancini 2009) implies that BH growth is tightly coupled to the process of galaxy and bulge formation. Increasingly, models invoke feedback processes from active galactic nuclei (AGN) to explain a host of phenomena, from the origin of...
the $M_{\text{BH}}-\sigma$ relation, to rapid quenching of star formation in bulges, to the build-up of the colour–magnitude relation and resolution of the cooling flow problem (see e.g. Silk & Rees 1998; Di Matteo, Springel & Hernquist 2005; Springel, Di Matteo & Hernquist 2005a; Croton et al. 2006; Hopkins et al. 2008b; Cattaneo et al. 2009; Hopkins & Elvis 2010, and references therein).

Observations have demonstrated that most of the accretion luminosity in the Universe is obscured by large columns of gas and dust (e.g. Lawrence 1991; Hill, Goodrich & Depoy 1996; Risaliti, Maiolino & Salvati 1999; Simpson, Rawlings & Lacy 1999; Simpson & Rawlings 2000; Willott et al. 2000; Ueda et al. 2003; Hao et al. 2005, and references therein). This obscured AGN population dominates the population of X-ray sources (Miyaji, Hasagawa & Schmidt 2001; Steffen et al. 2003; Ueda et al. 2003; Grimes, Rawlings & Willott 2004; Hasinger 2004, 2008; Sazonov & Revnivtsev 2004; Bargere & Cowie 2005; Hasagawa, Miyaji & Schmidt 2005; Nandra, Laird & Steidel 2005; Gilli, Comastri & Hasinger 2007), and accounts for most of the observed X-ray background luminosity (Setti & Woltjer 1989; Madau, Ghisellini & Fabian 1994; Comastri et al. 1995; Treister & Urry 2006; Gilli, Comastri & Hasinger 2007). It may dominate the bright end of the infrared luminosity function as well (Sanders & Mirabel 1996; Komossa et al. 2003; Ptak et al. 2003; Daddi et al. 2007; Hickox et al. 2007; Alexander et al. 2008; Hopkins & Hasinger 2010).

The abundance of obscured quasars remains a major uncertainty in reconciling synthesis models of AGN populations with the BH mass function today, and (by implication) understanding the radiative efficiencies of quasars (Salucci et al. 1999; Yu & Tremaine 2002; Hopkins, Richards & Hasinger 2007a; Shankar, Weinberg & Miralda-Escude 2009). Various specific galaxy populations [for example extreme X-ray/optical sources (EXOs), X-ray bright optically normal galaxies (XBONGs), ultraluminous infrared galaxies (ULIRGs), submillimeter galaxies (SMGs)] commonly host obscured AGN (Yuan & Narayan 2004; Alexander et al. 2005, 2008; Georgantopoulos & Georgakakis 2005; Maihneri et al. 2005; Max et al. 2005; Daddi et al. 2007; Riechers et al. 2008; Georgakakis et al. 2009; Nardini et al. 2009; Trump et al. 2009). And theoretical models have long predicted that in violent events such as galaxy mergers, there should be a transition from an early, ‘buried’ accretion stage corresponding to e.g. ‘warm’ ULIRGs and similar galaxies, to an at least partially unobscured phase in which the AGN removes dust and gas and is visible as a bright quasar (Sanders et al. 1988a,b; King 2003; Di Matteo et al. 2005; Hopkins et al. 2006a,b,c,e, 2008a, 2010; Younger et al. 2009; Hopkins & Hasinger 2010).

Yet AGN obscuration remains poorly understood. The most popular models invoke a torus-shaped ‘donut’ of obscuring material, on scales anywhere from $\sim 0.1$ to $100$ pc, to explain most of the heavily obscured AGN population (Lawrence 1991; Antonucci 1993). If one empirically assumes unification of obscured and unobscured AGN, then a number of the properties of the torus can be inferred: scale radii somewhere in the range above, and scale heights $h/R$ of order $\sim 1/3$ (Risaliti et al. 1999). The observed distributions of quasar and AGN column densities, and their detailed spectral energy distribution (SED) properties, place strong constraints on the densities, structure and column densities within the obscuring material, with typical column densities as high as $\sim 10^{26}$ cm$^{-2}$ through the edge-on plane of the material. And direct observations are beginning to probe these scales, through combinations of diverse techniques such as adaptive optics and maser observations (Greenhill et al. 2003; Iaffe et al. 2004; Davies et al. 2006, 2007; Mason et al. 2006; Sanchez et al. 2006; Krips et al. 2007; Hicks et al. 2009; Ramos Almeida et al. 2009), giving constraints on the kinematics, gas and dynamical masses and star formation rates (SFRs) at these scales. Indeed, this simple model of obscuration has proven successful at explaining a large number of AGN observables, and the torus forms the basis of most models uniting type 1 and type 2 AGN.

These successes should not mask the fact that the torus remains a phenomenological model. The simple ‘donut’ picture is just a toy model – there are a large and growing number of unambiguous cases where it fails, whether in predicting detailed radiative transfer properties coming from the microphysical gas structure (Elitzur & Shlosman 2006; Mason et al. 2006; Mor, Netzer & Elitzur 2009), or where the implied torus properties would involve bizarre radii and/or dust temperatures (Kuraszkiewicz et al. 2000; Tran 2003; Page et al. 2004; Stevens et al. 2005; Ramos Almeida et al. 2009), or where it is simply clear that the dominant obscuration is isotropic, or time dependent, or comes from much larger scales (e.g. those associated with circumnuclear starbursts and/or the host galaxy; see Soifer et al. 1984; Scoville et al. 1986; Sanders et al. 1988a,b; Donley et al. 2005; Schartmann et al. 2005; Rigby et al. 2006; Zakamska et al. 2006; Hatzinikosoglou, Fritz & Jarrett 2009; Liu et al. 2009; Martínez-Sansigre et al. 2009; Rowan-Robinson, Valtchanov & Nandra 2009; Lagos et al. 2011).

Without a physical model for the origin and evolution of nuclear gas inflows, a large number of questions remain unanswered. Where do toroidal-like obscuring gas structures come from, in the first place? What determines their characteristic gas masses, radii and structure? Why are such structures ubiquitous around AGN? Are they, in fact? It is also usually assumed that the torus is simply an obscuring ‘bystander’ to the accretion event, or at most a passive fuel reservoir. But could the torus play some more critical role in the accretion process itself? A major long-standing puzzle is what drives and maintains the scale height of the torus – simple thermal pressure would be lost to cooling in a time much shorter than the local dynamical time. A large number of torus properties have been attributed to feedback from either young stars or the BH accretion itself – including the typical scale heights, clumping/phase structure, gaseous velocity dispersions, possible correlations between these quantities and star formation, and even the fueling rates on to the BH (e.g. Wada & Norman 2002; Schartmann et al. 2009, and references therein). However, it is important to recall that we do not yet understand the basic dynamics of gas and stars entirely in the absence of feedback!

There have been some attempts to address these from a physically motivated perspective, both in analytic and numerical work (Elvis 2000; Cattaneo et al. 2005; Elitzur & Shlosman 2006; Kawakatu & Wada 2008, Wada, Papadopoulos & Spaans 2009; Hopkins & Elvis 2010). However, analytic models are severely limited by the fact that the systems at these radii are highly non-linear, often chaotic, and not necessarily in steady state (with inflow, mass build-up, star formation and feedback processes all competing). If one wishes to simultaneously follow the torus itself and the chaotic, non-symmetric gas inflows that form it in the first place, simulations are necessary. But simulations of galactic scales used to follow inflows and AGN obscuration have resolution of $\sim 100$ pc, much larger than the relevant scales here (Cattaneo et al. 2005; Hopkins et al. 2005b,c).

Although progress has been made with zoom-in refinement techniques (see e.g. Escala 2007; Mayer et al. 2007; Levine et al. 2008), the computational expense involved means that these simulations have, thus far, only barely probed that scales of interest and, in doing so, have made restrictive assumptions (typically explicitly turning off cooling and/or star formation on small scales); moreover, they provide only a snapshot at one instant from the parent simulation – they cannot survey statistical properties of the nuclear
2 THE SIMULATIONS

The simulations described here are from a survey of multiscale ‘zoom-in’ runs which model gas inflows and star formation from large galactic scales to sub-pc scales, and have been discussed in a series of papers (Hopkins 2010; Hopkins & Quataert 2010a,b, 2011a,b). A detailed description and list of simulations is presented in Hopkins & Quataert (2010b); we briefly summarize the salient properties here.

The simulations were performed with the TreeSPH code GADGET-3 (Springel 2005); the detailed numerical methods are described there and in Springel & Hernquist (2002) and Springel, Di Matteo & Hernquist (2005b). The simulations include collisionless stellar discs and bulges, dark matter haloes, gas and BHs. For this study, we are interested in isolating the physics of gas inflow. As a result, we do not include explicit models for BH accretion feedback – the BH’s only dynamical role is via its gravitational influence on scales \( \lesssim 10\) pc.

Because of the large dynamic range in both space and time needed for the self-consistent simulation of galactic inflows and nuclear disc formation, we use a ‘zoom-in’ re-simulation approach. This begins with a large suite of simulations of galaxy–galaxy mergers, and isolated bar-(un)stable discs. These simulations have \( 0.5 \times 10^6 \) particles, corresponding to a spatial resolution of 50 pc. These simulations have been described in a series of previous papers (Di Matteo et al. 2005; Cox et al. 2006; Robertson et al. 2006a,b,c; Younger et al. 2008; Hopkins et al. 2009a). From this library of simulations, we select representative simulations of gas-rich major mergers of Milky Way mass galaxies (baryonic mass \( 10^{11} \) M\(_\odot\)), and their isolated but bar-unstable analogues, to provide the basis for our re-simulations. The dynamics on smaller scales does not depend critically on the details of the larger scale dynamics. Rather, the small-scale dynamics depends primarily on global parameters of the system, such as the total gas mass channelled to the centre relative to the pre-existing bulge mass.

Following gas down to the BH accretion disc requires much higher spatial resolution than is present in the galaxy-scale simulations. We begin by selecting snapshots from the galaxy-scale simulations at key epochs. In each, we isolate the central 0.1–1 kpc region, which contains most of the gas that has been driven in from large scales. Typically this is about \( 10^{10} \) M\(_\odot\) of gas, concentrated in a roughly exponential profile with a scalelength of \( \sim 0.3–0.5 \) kpc. From this mass distribution, we then repopulate the gas in the central regions at much higher resolution, and simulate the dynamics for several local dynamical times. These simulations involve \( 10^6 \) particles, with a resolution of a few pc and particle masses of \( \approx 10^4 \) M\(_\odot\). We have run \( \sim 50 \) such re-simulations, corresponding to variations in the global system properties, the model of star formation and feedback and the exact time in the larger scale dynamics at which the re-simulation occurs. Hopkins & Quataert (2010b) present a number of tests of this re-simulation approach and show that it is reasonably robust for this problem. This is largely because, for gas-rich discy systems, the central \( \sim 300 \) pc becomes strongly self-gravitating, generating instabilities that dominate the subsequent dynamics.
These initial re-simulations capture the dynamics down to \(\sim 10\) pc, still insufficient to quantitatively describe accretion onto a central BH. We thus repeat our re-simulation process once more, using the central \(\sim 10\)–\(30\) pc of the first re-simulations to initialize a new set of even smaller scale simulations. These typically have \(\sim 10^6\)–\(10^7\) particles, a spatial resolution of 0.1 pc and a particle mass \(\approx 100 M_\odot\). We carried out \(\approx 50\) such simulations to test the robustness of our conclusions and survey the parameter space of galaxy properties. These final re-simulations are evolved for \(\sim 10^7\) yr – many dynamical times at 0.1 pc, but short relative to the dynamical times of the larger scale parent simulations. We also carried out a few extremely high-resolution intermediate-scale simulations, which include \(\sim 5 \times 10^7\) particles and resolve structure from \(\sim 1\) kpc to \(\sim 0.3\) pc – these are slightly less high resolution than the net effect of our two zoom-ins, but they obviate the need for a second zoom-in iteration and ‘bridge’ the scales of the above simulation suites. The conclusions from these higher resolution simulations are identical.

Our simulations include gas cooling and star formation, with gas-forming stars at a rate motivated by the observed Kennicutt (1998) relation. Specifically, we use a SFR per unit volume \(\rho_* \propto \rho^{3/2}\) with the normalization chosen so that a Milky-Way-like galaxy has a total SFR of about \(1 M_\odot\) yr\(^{-1}\). Varying the exact slope or normalization of this relation has no qualitative effect on our conclusions. However, we caution that since we do not resolve the scales of individual bound star-forming cores in these simulations, the star formation is probably more uniform over the small radii than it would be in a more realistic ISM model. This is unlikely to be important for global properties here, but may have important consequences for e.g. detailed radiative transfer effects.

Because we cannot resolve the detailed processes of supernovae explosions, stellar winds and radiative feedback, the effect of feedback from stars is crudely modelled with an effective equation of state (Springel & Hernquist 2003). In this approach, feedback is assumed to generate a non-thermal (turbulent, in reality) sound speed that depends on the local SFR, and thus the gas density. Hopkins & Quataert (2010b) describe in detail the effects of different subgrid ISM sound speeds on angular momentum transport and inflow rates, and argue that observations favour effective turbulent speeds of \(\sim 10\)–\(50\) km s\(^{-1}\) for densities \(\sim 1\)–\(10^4\) cm\(^{-3}\), respectively. However, because the real physics and their effects are uncertain, it is important to vary this prescription and determine which of our conclusions are sensitive to the assumed subgrid properties.

Within the context of this model, we can interpolate between two extremes using a parameter \(q_{\text{eos}}\). At one end, the gas has an effective sound speed of 10 km s\(^{-1}\), motivated by e.g. the observed turbulent velocity in atomic gas in nearby spirals or the sound speed of low-density photoionized gas; this is the ‘no-feedback’ case with \(q_{\text{eos}} = 0\).\(^2\) This is broadly similar to what is assumed in Bournaud, Elmegreen & Elmegreen (2007) and Teysseier, Chapon & Bournaud (2010). The opposite extreme, \(q_{\text{eos}} = 1\), represents the ‘maximal feedback’ model of Springel et al. (2005b); in this case, 100 per cent of the energy from supernovae is assumed to stir up the ISM. This equation of state is substantially stiffer, with effective sound speeds as high as \(\sim 200\) km s\(^{-1}\). This is qualitatively similar to the near-adiabatic equations of state in the BH accretion studies of Mayer et al. (2007) and Dotti et al. (2009). The sound speed at scales we consider cannot meaningfully be much larger than this, since it is similar to the circular/escape velocity. By varying \(q_{\text{eos}}\), we examine a spectrum of intermediate cases: for example, equations of state similar to the ‘starburst’ model in Klessen, Spaans & Jappsen (2007) or the sub-Giant Molecular Cloud (GMC) equation of state in Spaans & Silk (2005). Most of our suite of simulations focuses on a wide range of subgrid sound speeds \(\sim 20\)–\(100\) km s\(^{-1}\), motivated by a variety of observations of dense, star-forming regions both locally and at high redshift (Downes & Solomon 1998; Bryant & Scoville 1999; Förster Schreiber et al. 2006; Iono et al. 2007), and recent numerical simulations (Hopkins, Quataert & Murray 2011a).

Within this range, we found little difference in the physics of angular momentum transport or in the resulting accretion rates, gas masses etc. on the scales we consider (Hopkins & Quataert 2010b). More detailed comparison with the explicit stellar feedback models presented in Hopkins et al. (2011a,b,c) will be the subject of future work. Here, we will focus on the effects on the obscuring gas near the BH. Because we are not explicitly accounting for or resolving feedback processes, we do not expect these models to accurately reflect the detailed dynamics of gas in response to strong feedback. Rather, we wish to use our suite of simulations to identify behaviour that is robust to the effective pressure or turbulent sound speed of the gas – i.e. to identify robust aspects of the system that are present even without feedback such as stellar winds.

3 FORMATION OF THE TORUS

Hopkins & Quataert (2010b) show that when large-scale inflows are sufficient, the build-up of gas in the central regions of the galaxy triggers a cascade of secondary instabilities, that drive rapid inflows to still smaller radii and ultimately on to the BH. Around the BH radius of influence, these instabilities generically take the form of an \(m = 1\) mode – a thick, eccentric, slowly precessing gas+stellar disc, in which the eccentric stellar pattern torques strongly on the gas, inducing shocks and inflows. The disc can then propagate gas inflows and the \(m = 1\) pattern down to small radii \(\lesssim 0.1\) pc, where it transitions to a traditional alpha-disc. This should be generic to any quasi-Keplerian potential in a dissipative system with shocks (Hopkins 2010; Hopkins & Quataert 2011b).

Fig. 1 shows some illustrative examples of the nuclear gas discs that form around the BH radius of influence in our simulations. We plot gas surface density maps, with colour encoding the gas effective sound speed, from scales of \(\gtrsim 10\) kpc to \(\lesssim 1\) pc. The initial large-scale simulation in this case is a fairly gas-rich major merger of two \(\sim L_\odot\) galaxies (with initial bulges of mass 1/3 the disc mass and BHs of mass \(10^7 M_\odot\)). The zoom-in simulations were carried out just after the coalescence of the two nuclei, which is near the peak of star formation activity, but when the system is still quite gas rich.\(^3\)

We show both the global structure, and edge-on (\(R, z\)) disc. The scales shown include the BH radius of influence, about \(10\) pc in

\(^3\)The specific properties of each simulation are given in Hopkins & Quataert (2010b), those shown here are (top-bottom, left-to-right): Nf1h1c0thin, Nf1h1c1thin, Nf1h1c1q1s, Nf1h1c1den, Nf1h1c0 (left); Nf1h1c1ICs, Nf3h1c1mid, Nf2h2b2, Nf8h2b2, Nf8h2b4 (right). They have (respectively) initial gas fractions \(f_{\text{gas}} = 0.5, 0.6, 0.8, 0.8, 0.8, 0.75, 0.26, 0.20, 0.8, 0.8;\) BH mass \(\sim 3 \times 10^7 M_\odot\) and disc mass \(\sim 1-2, 1.7, 3.0, 8.1, 0.25, 1.7, 4.6, 7.0, 3.5, 0.5 \times 10^7 M_\odot\) inside 10 pc, and subgrid sound speeds \(c_s \sim 35, 20, 40, 50, 10, 40, 30, 25, 25, 20\) km s\(^{-1}\).
these galaxies. In the face-on projection, the \( m = 1 \) modes that form at these scales are clearly evident. They drive large torques on the gas, driving inflow into \( \ll 0.1 \) pc at accretion rates as high as \( 10 \, M_\odot \, \text{yr}^{-1} \) in these simulations, sufficient to power the most luminous quasars (see figs 5 and 13 in Hopkins & Quataert 2010b).

Here, however, we note the broad resemblance of these nuclear discs to the canonical AGN ‘torus’. The discs are thick, with characteristic scale \( \ll 0.1–10 \) pc, gas masses \( \sim M_\text{BH} \) and scale heights of order unity. Of course, unlike in toy models of the torus, the gas is part of a continuous distribution at all radii, and its structure is non-trivial.

### 4 VERTICAL STRUCTURE: DEPENDENCE ON STELLAR FEEDBACK

#### 4.1 Overview

The major input parameter of our models is the parametrization of the effects of stellar feedback on the ISM. This is accomplished, here, with the parameter \( q_{\text{eos}} \) described in Section 2, that allows us to interpolate between a feedback-free ISM and one with large non-thermal internal gas velocities and pressures driven by stellar feedback.

The so-called torus is defined largely by its vertical structure, which determines the obscured fractions. To the extent that the amount of turbulent velocity and pressure support in the simulated gas is defined by a subresolution model, we must ask whether the vertical structure we see in our simulations is entirely a consequence of our model inputs, or whether there are robust statements and predictions we can make.

We therefore consider the vertical structure in detail, in a specific survey of \( q_{\text{eos}} \). This survey (\( N/8b2b4j \) in Hopkins & Quataert 2010b) is a typical, canonical set of conditions (\( 3 \times 10^7 \, M_\odot \, \text{BH} \), with disc-to-BH mass ratio of a few inside \( \sim 100 \) pc initially, and initial gas fraction \( \sim 50 \) per cent, typical of the simulations in Fig. 1). We re-simulate the identical cases, but with \( q_{\text{eos}} = 0.0, 0.018, 0.06, 0.10, 0.12, 0.15, 0.21, 0.25, 0.35, 0.60, 1.0 \). The spacing in \( q_{\text{eos}} \) is chosen such that the implied turbulent gas sound speeds are spaced over roughly equal logarithmic intervals from the minimum \( q_{\text{eos}} = 0 \) floor (10 km s\(^{-1}\)) to the maximum \( q_{\text{eos}} = 1 \) value (which is density dependent, but \( \sim 100 \) km s\(^{-1}\) at range of interest).

Fig. 2 shows the edge-on (\( R, z \)) gas structure, as a function of \( q_{\text{eos}} \). A few generic features stand out. The discs are generally thick. At the smallest radii (\( \lesssim 0.1–1 \) pc), they eventually become thin, since the gravity from the BH becomes arbitrarily strong. This gives a torus-like morphology. Flares (discussed below) and lopsidedness (reflecting the lopsided disc mode on these scales) are not uncommon. As a function of \( q_{\text{eos}} \), we see unsurprisingly that the gas distribution becomes more smooth and vertically extended at higher \( q_{\text{eos}} \). For \( q_{\text{eos}} \gtrsim 0.4 \), the system is no longer really a vertically
supported disc, but spherical – however, as discussed in Hopkins & Quataert (2010b), this is likely an unrealistically large implicit feedback efficiency.

The most surprising thing about Fig. 2 is how little change there is as a function of \( q_{\text{eos}} \). For \( q_{\text{eos}} \sim 0–0.35 \), there is a factor of \( \sim 5–25 \) change in \( c_s \), which leads to a naive expectation of a factor of \( \sim 5–25 \) change in \( h/R \). We see much weaker variation. We now consider this quantitatively.

### 4.2 General expectations

To inform our comparisons, consider a simple smooth, isothermal system, in which the self-gravity of the gas is negligible (i.e. the potential is dominated by the BH, stars and/or dark matter). The equation of vertical hydrostatic equilibrium,

\[
\frac{\partial P}{\partial z} = c_s^2 \frac{1}{\rho} \frac{\partial \rho}{\partial z} = - \frac{\partial \Phi}{\partial z},
\]

then has the trivial solution

\[
\rho(R, z) = \rho_0(R) \exp \left\{ -\frac{1}{2} \left( \frac{z}{h_s} \right)^2 \right\}.
\]

For large \( z/R \), this depends on the specific form of \( \Phi \), and so on the details of the global mass distribution. However, if the disc is thin, i.e. most of the mass is at \( z/R \ll 1 \), then this has a particularly simple expression. For any background spherical mass distribution, we have \( \frac{\partial \Phi}{\partial z} = \frac{\partial \Phi}{\partial r} \left( \frac{\partial r}{\partial z} \right) = (V_c^2/r)(z/r) \), where \( r^2 = R^2 + z^2 \) and \( V_c^2 = G M_{\text{enc}}(< R)/r \). So for \( z/R \ll 1 \), \( \Phi(R, 0) - \Phi(R, z) \approx G M_{\text{enc}}(< R) R^{-2} z^2 z / 2 \approx V_c^2 (z/R)^2 / 2 \).

Together, this gives the especially simple solution for the density for a quasi-spherical potential:

\[
\rho(R, z) \approx \rho_0(R) \exp \left\{ -\frac{1}{2} \left( \frac{z}{h_s} \right)^2 \right\},
\]

\[
\frac{h_s}{R} = c_s \frac{V_c}{c_s} = c_s \left( \frac{G M_{\text{enc}}(< R)}{R} \right)^{-1/2}.
\]

Of course, the \( c_s \) here does not need to be thermal. Non-thermal pressure sources such as turbulent motions will have the same effect. So for comparison with simulations we should take \( c_s \rightarrow c_s, \text{eff} \), where \( c_s, \text{eff} = c_s^2 + \sigma_z^2 \) includes both the thermal and/or subresolution effective sound speed \( (c_s) \) and resolved turbulent vertical motions \( (\sigma_z) \).

Fig. 3 compares this expectation for \( \rho(z) \) as a function of \( c_s, \text{eff} / V_c \) to the actual vertical mass distribution measured in narrow radial annuli from \( \sim 1 \) to \( 10 \) pc. We use the full \( c_s, \text{eff} \) as defined above. The distributions are reasonably described by the above scalings, a Gaussian core is typical, with a slightly broader (often more exponential) distribution at high-\( z \). Remember that at sufficiently large \( |z| \), the correct solution involves the full potential; if we account for this more accurately, we see similar agreement. The important point is that the gas does appear to be in vertical equilibrium.

### 4.3 Gravitational support

Given the gas dispersion \( c_{s, \text{eff}} \) in the simulations, the vertical structure is what we would expect. But are these dispersions...
primarily subresolution (set by the model), thermal, or gravitational in origin?

Fig. 4 again considers the vertical gas profile, but in our survey of different $q_{\text{eos}}$. We know from Fig. 3 that accounting for the full gas motion explains the observed scale heights. Therefore here we compare the expectation if the gas motions were purely thermal and/or subresolution -- i.e. $c_{\text{z,eff}} = c_s$, where $c_s$ is the sound speed and is dominated by the subresolution turbulent effective $c_s$ (since the explicit cooling time of the gas is $\sim 10^5$ times shorter than its dynamical time).

In the higher $q_{\text{eos}}$ (higher effective $c_s$) simulations, this explains most of the pressure support, i.e. the resolved turbulent dispersion $\sigma_r \ll c_s$ (subgrid). But at low $q_{\text{eos}}$, the scale heights and $c_{\text{z,eff}}$ do not drop nearly as quickly as the subgrid $c_s$ alone. There is some non-thermal, resolved gravitational process giving rise to minimum scale heights and vertical dispersions.

What, then, dominates the effective vertical ‘heating’ in the torus region?

### 4.3.1 Clump–clump encounters

It has been proposed that two-body scattering between dense molecular clumps in the gas could maintain the observed scale heights (Krolik & Begelman 1988; Nayakshin & King 2007; Hobbs et al. 2011). However, we find these effects are negligible in our simulations.

Consider clumps within the plane of a disc. Scattering a clump to large $V_z \sim V_c$ requires both (a) an encounter between two clumps with relative velocity $\gtrsim V_c$, and (b) an encounter within an impact parameter $b$ such that $G M_{\text{BH}} b V_c \gtrsim V_{\text{c,enc}}$. The mean time per clump between such encounters is just $\tau^{-1} \sim f(V_c/\sigma) n_{\text{cl}} b^2 V_c$, where $n_{\text{cl}}$ is the volume density of clumps and $f(V_c/\sigma)$ is the fraction of the clumps moving on orbits with large non-circular motions ($|V - V_{\text{c,enc}}| \gtrsim V_c$). If the system is sufficiently thin such that $b > h$, the disc thickness, then this becomes $\tau^{-1} \sim f(V_c/\sigma) \frac{dN_{\text{cl}}}{dB} b h V_c$. Using $n_{\text{cl}} = \frac{\rho_{\text{gas}}}{M_{\text{cl}}^3} = \frac{\Sigma_{\text{gas}}/h M_{\text{cl}}}{3}$, the maximum $b$ above and $V_{\text{c,enc}} \sim G M_{\text{BH}}/r$, this can be written as

$$\Omega \tau \sim \frac{1}{f(V_c/\sigma)} \left( \frac{M_{\text{BH}}}{M_{\text{gas}}} \right) (1 + Q N_{\text{cl}}),$$

where $Q$ is the Toomre $Q \sim (h/R)(M_{\text{gas}}/M_{\text{BH}})^{-1}$ and $N_{\text{cl}}$ the total number of clumps, and the expression shown interpolates between the extremely thin and thick-disc cases. For a Maxwellian velocity distribution, $f(V_c/\sigma) \sim \exp\left\{-(V_c/\sigma)^2/2\right\}$. Since both $f(V_c/\sigma)$ and $M_{\text{gas}}/M_{\text{BH}} \sim M_{\text{gas}}/M_{\text{BH}}$ are small at this radius, collisions require many dynamical times. But any induced vertical heating will relax away in just a single or couple dynamical times, since the cooling time is much shorter than the dynamical time. So without continuous energy input to drive large dispersions – which is essentially the problem we wished to solve in the first place – this mechanism fails.

© 2011 The Authors, MNRAS 420, 320–339

Monthly Notices of the Royal Astronomical Society © 2011 RAS
Moreover, if star formation occurs with some efficiency relative to the dynamical time ($\rho_* \sim \epsilon \rho \sqrt{G \rho}$, with $\epsilon \sim 1$–10 per cent), then using the fact that any clump must have $\rho \gtrsim M_{\text{enc}}/R^2$ to avoid tidal destruction, clump–clump gas heating must occur faster than the gas exhaustion time-scale in a clump, requiring

$$\left( \frac{M_{\text{gas}}}{M_{\text{enc}}} \right) \gtrsim \epsilon (1 + Q N_\text{cl}) \exp \left( \frac{1}{2} \left( \frac{V_*}{\sigma} \right)^2 \right).$$ (6)

Even for $\sigma \sim V_*$ (which begs the question) and an extremely thin disc $Q N_\text{cl} \lesssim 1$, this requires $M_{\text{gas}}/M_{\text{enc}} \gg 0.1$, which is not satisfied at the inner radii $\lesssim 10$ pc.

4.3.2 Twists and misalignment

Another possibility is that large covering factors are maintained by virtue of the fact that the nuclear disc is misaligned with the larger scale inflow/bar/disc. This is particularly interesting because observations find relatively little correlation between the axes of AGN (traced by jets or the torus) and the inclination of the host galaxy (e.g. Keel 1980; Lawrence & Elvis 1982; Schmitt et al. 1997; Simcoe et al. 1997; Kinney et al. 2000; Gallimore et al. 2006; Zhang et al. 2009; but see also Maiolino & Rieke 1995; Shen, Shao & Gu 2010 and references therein). In a companion paper, Hopkins et al. (2011d), we show that this lack of alignment is reproduced in our simulations owing to two processes. First, occasionally the central gas supply is strongly influenced by a single or couple large clumps that form at large radii, fragment and sink, realigning the central angular momentum vector (see e.g. Nayakshin & King 2007). Examples of this have also been seen in cosmological zoom-in simulations (Levine, Gnedin & Hamilton 2010). Secondly, even in smooth flows, the secondary and tertiary gravitational instabilities will tend to decouple their angular momentum from the primary (external) bar/spiral structure, and semichaoically precess or tumble in three dimensions (see Heller, Shlosman & Englmaier 2001; Shlosman & Heller 2002; El-Zant & Shlosman 2003; Englmaier & Shlosman 2004; Maciejewski & Athanassoula 2008).

In Fig. 5 we show how this can contribute to obscuration. It is straightforward to measure the axis of angular momentum of the disc in a radial annulus, and define the corresponding inclination $\Theta(r)$ (relative to the initial, uniform angular momentum axis of the entire initial disc). We also know the mass $\Delta M_{\text{gas}}(r)$ enclosed in

$$\text{Fraction of Simulations}$$

| Obscured Fraction [from Misalignment] |
|--------------------------------------|
| 0.00                                 |
| 0.05                                 |
| 0.10                                 |
| 0.15                                 |
| 0.20                                 |
| 0.25                                 |

Figure 5. Contribution of disc misalignments and twists (as opposed to disc thickness) to obscuration. For each simulation in the sample shown in Fig. 3, we calculate the (time-averaged) fraction of sightlines towards the BH which would be obscured ($N_{\text{H}} > 10^{22} \text{ cm}^{-2}$) assuming the disc is razor-thin at every radial annulus $\Delta r$ (using the angular momentum vector of gas in $\Delta r$ to define the disc plane). If the disc plane (angular momentum direction) were constant with radius, this would be zero, but if the plane tilts as a function of radius, it can be non-zero (a 180° flip at some radius, maintained for the duration of the simulation, would give an obscured fraction of unity). Although there are significant misalignments (see Hopkins et al. 2011d), and in a few systems they can account for obscured fractions $\gtrsim 50$ per cent, they would only give an integrated obscured fraction of $\gtrsim 25$ per cent if the disc were thin at all radii. And even simulated cases with no ‘twists’ still have large obscured fractions and $h/R$ in Fig. 3. The torus must actually be thick to match observations – so some process must explain large scale-heights in the gas.
If $\sigma_r$ is sufficiently large, the system is vulnerable to the so-called ‘firehose’ instability and bending modes will be self-excited. However, it is unusual to see such large $\sigma_r$ (and $\sigma_z \gg \sigma_\phi$) in discs. Even with large $\sigma_r$, the fact that $\sigma_r < V_c(r)$ for any meaningful ‘disc’ means that usually, when the self-gravity of the disc is small compared to the background potential, the system is stable. And in even in self-gravitating discs with large $\sigma_r$, it typically takes only a small $v_z$ to stabilize them, so the induced $h/R$ is not large.

However, consider the special case of interest here, where the disc is quasi-Keplerian and has a large lopsided mode driving accretion. The system is (initially) a thin disc in the quasi-Keplerian potential of a BH – i.e. to lowest order, the parameters are those of a pure Keplerian potential, with some correction terms that scale with $M_d/M_{\text{BH}}$ of $\mathcal{O}(\epsilon) \ll 1$. As discussed above, and in previous works, the disc develops a gravitational instability in the disc plane (the standard density waves of spiral/bar/etc. modes), which we can describe by e.g. the perturbed density field $\Sigma_1(R,t) = |a| \Sigma_0(R) \exp \{ i (\tilde{b} k_b(R)dR + \omega_p t - m_b \phi) \}$. Here $|a|$ is the effective mode amplitude in the density field at $R$, and the properties $\omega_p$, $m_b$ and $k_b$ refer to the frequencies and wavenumbers of this, in-plane mode (independent from the $\omega_n$, $m_n$ and $k_n$ of the bending mode).

The fact that the potential is quasi-Keplerian, i.e. has $\Omega \approx \kappa$, favours (and supports for long periods of time) global, ‘slow’ $m = 1$ modes – modes with $m_b = 1$, $\omega_p \sim \epsilon \Omega \ll \Omega$ and $|k_b| R \sim 1$. These are the lopsided/eccentric modes that we see above. The potential of the BH+disc system is

$$\Phi(r) = -G \frac{M_{\text{BH}}}{r} + \Phi_\delta(r),$$

and it is useful to define the parameter

$$\sigma \equiv \frac{\Omega^2 - k^2}{2 \Omega} = -\frac{1}{2 \Omega} \left( \frac{d}{dr} r^2 \frac{d}{dr} \right) \Phi_\delta.$$  

To first order in $\epsilon$, then, the Wentzel–Kramers–Brillouin (WKB) dispersion relation of such modes in a cold ($c_s \ll V_c$) disc is

$$\omega_p = \sigma + \pi G \Sigma_0 |k_b| \Omega^{-1}$$

(Tremaine 2001). The equations of motion for the perturbed velocity $v = R \Omega \dot{\phi} + v_r \dot{R} + v_\phi \dot{\phi}$ become, at this order,

$$v_r = -\frac{i}{2 (\omega_p - \sigma)} \left( \frac{d}{dr} \Phi_\delta \right) = -\frac{\Sigma_1}{\Sigma_0} \Omega |k_p|^{-1},$$

$$v_\phi = \frac{i}{2} v_r,$$

where we have used the WKB relation $\Phi_\delta \approx -2\pi G |k_p|^{-1} \Sigma_1$.

Since $\sigma_\phi^2 = |\langle \delta \phi \rangle|^2$ and $V_c = \Omega R$, this becomes just

$$\sigma_r = |a| |k_p R|^{-1} V_c,$$

with $\sigma_z^2 = 4 \sigma_\phi^2$. And recall for our simulations, the magnitude $|a|$ in the disc is order unity during the active phases of BH growth (Hopkins & Quataert 2010b).

This is the important point – because of cancellations that occur (essentially the entire system is near-resonance), the radial induced velocities from the $m = 1$ mode are quite large, $\sim V_c$, independent of how small the ratio $M_d/M_{\text{BH}}$ may be.

Now return to the dispersion relation for bending modes. For the non-disc (Keplerian) part of the potential, $\nu^2 = \Omega^2$. To leading (zeroth) order in $\epsilon \sim M_d/M_{\text{BH}}$, then, the dispersion relation

\[\text{Each annulus; we can simply integrate along all sightlines towards the BH, assuming the mass is in an axisymmetric razor-thin disc with inclination $\theta(r)$, to obtain the column density $\Sigma_\text{gas}$ along each sightline. If a sightline is covered by the disc at some radius, it is ‘obscured’.}

\[\text{We consider these assumptions because they effectively define a minimum obscured fraction stemming purely from twists and misalignments. This fraction can be considerable, but there is a broad range in different simulations – many systems have only $<20$ per cent covering fractions, but there is a long tail towards near full covering (anti-alignment of the central and outer discs). Integrated over all simulations and snapshots, the average covering fraction is $\approx 25$ per cent. A warped or twisted disc can therefore yield large covering angles towards the BH even when the disc itself is thin.}

\[\text{However, this is not the full story. First of all, the covering fractions of $\approx 25$ per cent are still significantly lower than the total covering fraction of obscuration in the simulations, by a factor of at least $\sim 2$. Moreover the cases with weak twists ($<20$ per cent covering in Fig. 5) still exhibit large obscured fractions and thick discs. The key point is that the vertical density distribution in Figs 3–4 shows that we must explain the actual thickness, not just the orientation of the discs. This is true for observations as well – empirical modelling of the hot dust continua indicates that the obscuring region must be geometrically thick, not just a misaligned larger scale thin disc (e.g. Deo et al. 2009, and references therein). A time-dependent twist can, in principle, ‘pump up’ vertical motions, but fast cooling times make it difficult to sustain a large scale height anywhere except close to the location of the twist (where the pumping occurs). Some mechanism that pumps vertical motion through the disc, on a time-scale comparable to the local dynamical time, is required.}

\[\text{4.3.3 Bending modes}

\[\text{Bending modes can provide an efficient channel for ‘heating’ the torus. Their behaviour is particularly interesting in response to ‘slow modes’ in a quasi-Keplerian potential. Consider a general bending mode}

\[h(R, \phi, t) = H(R) \exp \{ i (\omega_0 t - m_b \phi) \},

\[H(R) = h_0(R) \exp \left\{ i \int R k_b(R')dR' \right\}

in a system that includes some quasi-spherical component (BH+bulge+halo) and a thin disc with surface density $\Sigma_d$, angular (vertical) frequency $\Omega$ (v) and velocity dispersions in the radial, azimuthal and vertical directions $\sigma_r$, $\sigma_\phi$, $\sigma_z$. The value $k_b$ is the radial wavenumber of the bending mode, and $m_b$ is its azimuthal wavenumber. In the WKB regime, if $\sigma_r^2 \gg \sigma_\phi^2$, the dispersion relation can be written as

$$|\omega_0 - m_b \Omega| \Omega^2 = \nu^2 + 2\pi G \Sigma_d |k_b| + (\sigma_z^2 - \sigma_\phi^2) k_b^2$$

(Kulsrud & Mark 1970; Kulsrud, Mark & Caruso 1971; Mark 1971; Poliachenko 1977).\footnote{Technically, we require a column that translates to $N_{\text{H}} > 10^{22}$ cm$^{-2}$, but because of our razor-thin assumption, this is almost identical to being covered by the disc.}

\[\text{Note that the $\sigma_\phi^2$ that appears in equation (9) is not technically a dispersion (that being defined $\langle \nu^2 \rangle - \langle \nu \rangle^2$, but the mean $\langle \nu_\phi^2 \rangle$. Thus streaming/bulk motion in the radial direction is affected just as much as random motions about some mean $\nu_r$, important for our purposes, since gas parcels being collisional tend to move in coherent streaming motion.}
becomes
\[
\left( \frac{\omega_h}{\Omega} - m_p \right)^2 = 1 + \left[ \frac{h}{R} - |a|^2 \left| k_p R \right|^{-2} \right] \left| k_0 R \right|^2,
\]
(16)
where we have defined \( h/R = \sigma / V_c \).

Recall, \( |a| \left| k_p R \right|^{-2} \sim 1 \) for the lopsided disc mode, and \( |k_0 R| \gg 1 \) in the WKB limit – thus, whenever the disc is thin, the radial motions induced by the \( m_p = 1 \) eccentric disc excite bending modes. These bending modes will grow on the local dynamical time-scale, since \( \omega_h \sim \Omega \). Compare the slow modes in the plane, which have \( \omega_h \sim \epsilon \Omega \) and can be long-lived.

The bending mode, of course, creates some non-trivial vertical motion. The growth of the mode will saturate when it drives a \( \epsilon \) sufficiently large so as to reach the marginal stability condition of equation (16) (\( \text{Im}(\omega_h) = 0 \)), which we can write as
\[
\left( \frac{h}{R} \right) \rightarrow \left| k_0 R \right|^{-1} \left( \left| a \right|^2 \left| k_p R \right|^{-2} \left| k_0 R \right|^2 - 1 \right)^{1/2}
\]
\[
\approx \frac{|a|}{|k_0 R|} \sim |a|,
\]
(17)
where the second equality uses \( |k_0 R| \gtrsim 1 \). In short, tightly wound bending modes arise, and saturate \( v_z \) at a large fraction of \( V_c \), such that \( h/R \) is driven to order unity wherever the eccentric mode persists, independent of the degree of self-gravity of the disc!

In Fig. 6, we check whether this prediction at all describes our simulations. We compare the scale height \( h/R \) to the measured mode amplitude \( |a| \) of the in-plane \( m = 1 \) mode, at a random time during the active phase, for each of a subset of our simulations. We chose only the simulations for which \( q_{\text{gas}} \lesssim 0.1 \), where we can confirm that the subgrid assumed \( c_{z\text{eff}} \) does not dominate \( c_z \) or the vertical scale height on the scales we measure (see Fig. 4). We sample both quantities at even intervals in \( \log R \) from \( R = 0.3 \text{–} 10 \text{pc} \). There is, unsurprisingly, large scatter, but a correlation is significant at \( > 3 \sigma \) and consistent with \( h/R \sim |a| \) over most of the simulated range. That the relation is not exactly linear at the high-\( h/R \) end and shows considerable scatter is expected, both because of contributions from \( k_0 \) and \( k_p \) in the derivations above, non-linear effects (especially at \( h/R \) and/or \( |a| \gtrsim 0.1 \)), and some non-zero support from \( c_z \). But it is quite unlikely that this relation would arise accidentally – after all, for otherwise equal properties, a lower \( h/R \) disc is actually more gravitationally unstable, so if anything we would naively expect the inverse of the observed correlation.

5 BASIC DYNAMICAL PROPERTIES OF THE ‘TORUS’

Thus far, we have focused on the origin of torus structural properties in simulations. We now examine these properties in more detail and compare to observations. Fig. 7 shows a number of (azimuthally averaged) properties of the nuclear gas, as a function of radius. We plot the gas surface mass density, gas fraction, SFR, vertical gas velocity dispersion and gas inflow rate \( M \) (here defined so as positive is inflow). The velocity dispersion includes both resolved and subresolution components, i.e. \( \sigma^2 = \sigma_c^2 + \sigma_z^2 \), where \( \sigma_c \) is the subgrid implied sound speed (plus any thermal components) and \( \sigma_z \) is resolved vertical dispersion.

We show this for our suite of simulations from Fig. 4 in which we systematically vary the subgrid equation of state (via the parameter \( q_{\text{gas}} \)). For each, we select a random snapshot near the peak of inflow activity. Because the global properties – gas density profiles, inflow rates, circular velocities, etc. – are primarily set by global gravitational torques (see Hopkins & Quataert 2011b), the parameter \( q_{\text{gas}} \) does not appear to have a dramatic qualitative effect on these properties. The primary effect is to determine the efficiency of fragmentation, which in turn changes the variability and global efficiency of star formation and gas exhaustion. If we consider the wider range of simulations shown in Fig. 3, which vary the initial gas fractions, bulge-to-disc and BH-to-disc mass ratios, we find a similar range in the predicted properties.

In more detail, Hopkins & Quataert (2011a) show that the surface density profiles that arise are a natural consequence of the dynamics of tidal torques from the \( m = 1 \) lopsided disc instabilities. Specifically, the perturbation dynamics set a robust range of ‘quasi-equilibrium’ profiles in which the gas mass density remains quasi-steady state over the active phase so long as there is sufficient initial inflow to trigger the process. If the profile is a power law \( \Sigma \propto R^{-\alpha} \), then this range is \( 1/2 \lesssim \alpha \lesssim 1 \), similar to that seen in ‘cuspy’ ellipticals.

The SFR surface density follows simply from the assumed local relation between star formation efficiency and dynamical time: in the simulations, \( \rho_* \propto \rho^{5/2} \). Competition between gas inflows and star formation sets the gas fractions, although these evolve significantly via depletion.

There are some observations to which we can compare. Water masers have been observed and used to map the inner disc structure around AGN in a few nearby galaxies (Greenhill, Moran & Herrnstein 1997; Greenhill et al. 2003; Braatz et al. 2004; Henkel et al. 2005; Kondratko et al. 2006a, 2008; Kondratko, Greenhill & Moran 2006b). These are sensitive to densities \( \sim 10^5 \text{–} 10^6 \text{ cm}^{-3} \) (typically \( \sim 0.1\text{–}1 \text{ pc} \)). At larger radii, interferometry has also been used to image the molecular and H I gas in the nuclei of some nearby systems (Schinnerer et al. 2000, 2008; Lonsdale et al. 2003; Combes et al. 2004; García-Burillo et al. 2005). Complemented with adaptive optics imaging of nearby nuclei, this gives constraints on the

Figure 6. Comparison of the scale height \( h \) (as in Fig. 3) to the \( m = 1 \) eccentric disc mode amplitude \( |a| \equiv \left| \Sigma_1 / \Sigma \right| / \text{at different radii (each point samples radii evenly in \( \log R \) from \( R = 1\text{–}10 \text{pc} \) for different simulations (each colour denotes a different simulation, near its peak activity). All simulations shown are ‘cold’ (i.e. have weak stellar feedback assumed in their subresolution prescription, specifically \( q_{\text{gas}} < 0.1 \)), so that the scale height is dominated by resolved turbulent vertical motions, not the feedback model input. In these cases, there is a correlation of the form \( h/R \sim |a| \), corresponding to the prediction from an \( h/R \) pumped-up by bending modes, themselves excited by radial motions from the in-plane eccentric mode. The torus height can be continuously sustained by exchange of energy from the eccentric/lopsided disc mode that powers the accretion on the BH.

© 2011 The Authors, MNRAS 420, 320–339
Monthly Notices of the Royal Astronomical Society © 2011 RAS

The dynamical AGN torus

329

Downloaded from https://academic.oup.com/mnras/article-abstract/420/1/320/1045387 by guest on 18 September 2019
gas+stellar dynamics, and information on the star formation history (Kuntschner et al. 2001; Davies et al. 2006, 2007; Sánchez et al. 2006; Hicks et al. 2009).

We compile these observations and compare to our simulations in Fig. 7. Most of the observed systems have BHs with broadly similar masses to ours $\sim 3 \times 10^7 M_\odot$. We plot the observations at all radii available. The maser observations are shown as points with error bars for resolved properties of discs outside the minimum radius enclosing the BH. The larger scale surface densities mapped from the gas velocity fields with very long baseline interferometry (VLBI) are shown as solid lines. For constraints involving stars (gas fractions, SFR), the VLBI+AO constraints are shown as diamonds, at the minimum resolved radii of the AO observations. The nuclear star formation history is modelled for several cases in Davies et al. (2007); we show their estimated current SFR both at the innermost radii where stellar light is measured and at the outer radii where the integrated light is used to determine the SFH. We also show their estimated maximum SFR of each observed burst from the fitted SFH within the observed radius. In all cases the observations broadly bracket the simulations, albeit with larger uncertainties in $f_{\text{gas}}$ and the SFH.

Figure 7. Azimuthally averaged nuclear disc properties versus radius. Each dotted line is a simulation with different galaxy and inflow properties, but one where some nuclear inflow is excited. We chose a random time near the peak of activity for each simulation to show here, but the results are similar over the entire active phase of each. Here for a set of simulations with identical initial conditions but varied $q_{\text{esc}}$, the parameter describing our effective equation of state and stellar feedback model. Lines range from black ($q_{\text{esc}} = 0$; effective $c_s = 100 \text{km s}^{-1}$) to red ($q_{\text{esc}} = 1$; effective $c_s = 100 \text{km s}^{-1}$), evenly spaced in $\log c_s$. Top left: surface mass density profiles. Points with error bars show constraints from AGN maser discs (NGC 3079; Kontratko, Greenhill & Moran 2005, magenta), (NGC 3393; Kontratko, Greenhill & Moran 2008, dark green), (NGC 1068; Lodato & Bertin 2003, pink), (Circinus; Greenhill et al. 2003, light green). Solid lines show constraints from adaptive optics (AO) measurements of AGN (NGC 1068, 1097, 3227, 3783, 4051, 4151, 6814, 7469, Circinus; in violet, green, cyan, blue, dark green, magenta, red, orange, dark blue, respectively; Davies et al. 2007; Hicks et al. 2009). Top centre: gas fraction $(M_{\text{gas}}(<R))/(M_*(<R) + M_{\text{gas}}(<R))$. Diamonds are the AO systems (same colour styles), at the resolution limits for complementary constraints used to derive $f_{\text{gas}}$. Top right: SFR surface density. For each AO area where a measurement is available, two points are shown. The first is the current SFR (small points), second is the peak SFR (larger points), both estimated from the fits to the SFR history inside the minimum and maximum observed radii in Davies et al. (2007). Bottom left: integrated SFR(<R). Bottom centre: vertical velocity dispersion of the gas. Dotted lines show both the resolved and subgrid assumed dispersions (added in quadrature). Bottom right: instantaneous inflow rate through R, generated by gravitational torques from the eccentric disc structure.

Of course, since these properties all scale with the dynamical properties of the system, they are all mutually correlated. A Kennicutt–Schmidt-type law similar to that observed ($\Sigma_{\text{SF}} \propto \Sigma_{\text{gas}}^{-1}$) for the nuclear-scale observations see Hicks et al. (2009) is effectively built into our simulations by subgrid assumption. We have discussed extensively the gravitational origin of the dispersion ($\sigma$). But both $\sigma$ and $\Sigma$ are related to $V_*$, for obvious dynamical reasons, and increase at smaller radii and/or in more massive/dense systems. And $\Sigma$ is tied to $\Sigma_{\text{SF}}$ via the Kennicutt relation. We therefore predict a relation between $\Sigma_{\text{SF}}$ and $\sigma$ for purely gravitational dynamic reasons. In the past, such a correlation has been interpreted as evidence of stellar feedback driving the observed dispersions – we find this may not be necessary.

6 Both the observed and simulated Kennicutt-type laws appear to have an index closer to $\eta \sim 1.7$ rather than the canonical $\eta \sim 1.4$. In the simulations, this is because we assume a local $\rho_\odot \propto r^{-3/2}$, and for the simple case of a gas disc contracting at constant $b/R$ this predicts $\eta = 1.75$.
6 THE COLUMN DENSITY DISTRIBUTION: TO CLUMP OR NOT TO CLUMP?

Thus far, all of our analysis has concerned global properties of the simulated torii, which we have reason to believe should be robust to the exact microstructure of gas on unresolved scales. However, subresolution structure can be important in calculating the column densities observed towards the BH. We therefore consider this now with two simple subresolution models.

6.1 The no-substructure case: smooth torii

One extreme is trivial: we simply take the gas distribution exactly as is from the simulations, without any assumed subgrid substructure. The column density along a given line of sight at each time can then be simply determined (following Hopkins et al. 2005a). We generate \( \sim 1000 \) radial lines of sight (rays) uniformly spaced in solid angle and with its origin at the BH, and integrate the line-of-sight density until outside the galaxy.

This assumption maximizes obscuration, since locking mass up in subresolution clumps would confine mass to smaller covering fractions (see the discussion from simulations in Hopkins et al. 2005b).

6.2 The clumpy torus

In fact, we know that there must be substructure in the gas, because cooling and star formation occur. Most of the mass in the ISM is probably locked into dense cold clumps. Unfortunately our simulation, limited by the physics included, does not predict the clump properties but only indirectly assumes an effective ISM state. However, with some simple assumptions, we can construct a subresolution estimate of all the relevant column properties, without the introduction of any tunable parameters.

Assume temporarily that most of the mass in the ISM is locked into \( N_{\text{cl}} \) dense clumps, with mean mass \( M_{\text{cl}} \), size \( R_{\text{cl}} \) and mean density \( \rho_{\text{cl}} = M_{\text{cl}} / (4\pi R_{\text{cl}}^3) \). Define the density contrast \( \rho = \rho_{\text{cl}} / \bar{\rho} \), with respect to the volume-average background density \( \bar{\rho} \). We make two assumptions, both just at the order-of-magnitude level: that the clumps are quasi-viral, and that they are in pressure equilibrium with the external medium. The first implies that whatever supports the clump generates an effective pressure \( P_{\text{cl}} = \rho_{\text{cl}} V_{\text{cl}}^2 / 3 \), where \( V_{\text{cl}} = (GM_{\text{cl}} / R_{\text{cl}})^{1/2} \). But this is just \( P_{\text{cl}} = G \Sigma_{\text{cl}} / R_{\text{cl}} \), where \( \Sigma_{\text{cl}} \) is the column density through the clump \( \sim \rho_{\text{cl}} R_{\text{cl}} \). To within a factor of 2 or so, this is even true for clumps in free-fall collapse, so is likely to be robust. We know the external effective (volume-average) pressure of the medium, \( P_{\text{eff}} \) – this is just the volume-average pressure used for all smoothed particle hydrodynamics (SPH) calculations. It is straightforward to then set \( \Sigma_{\text{cl}} \sim P_{\text{eff}} / G \), and obtain

\[
\Sigma_{\text{cl}} = \sqrt{P_{\text{eff}} / G}.
\]  

(19)

Pressure equilibrium is a less certain assumption, but if we were to force a mass–radius or linewidth–radius relation similar to the observed Larson’s laws in GMCs (\( \sigma \propto R^{1/2} \)), we would obtain the same dimensional scalings.7 Assuming that clumps follow the Jeans mass and radius in a self-regulating \( Q = 1 \) disc actually also results in the same dimensional scalings, so it may be robust in a variety of regimes.

The probability of a path-length \( \Delta r \) intersecting a cloud is given by \( p = (N_{\text{cl}} / V_{\text{tot}}) \sigma_{\text{cl}} \Delta r \), where \( V_{\text{tot}} \) is the total volume, and \( \sigma_{\text{cl}} \sim \pi R_{\text{cl}}^2 \) the clump cross-section. But since \( N_{\text{cl}} \sim M_{\text{tot}} / M_{\text{cl}} \), this simply reduces to

\[
p_{\text{cl}} \sim \bar{\rho} \Sigma_{\text{cl}}^{-1} \Delta r = \bar{\rho} \left( P_{\text{eff}} / G \right)^{-1/2} \Delta r.
\]  

(20)

The only two quantities we ultimately care about, the probability of intersecting a clump, and the clump column, have the useful feature that the clump density contrast and number of clumps completely cancel out. Thus, for any system where the mass is concentrated in quasi-viral, pressure-equilibrium clumps, we can determine the column density distribution and probability of sightlines seeing clumps based only on reference to well-determined volume-average gas properties in the simulations (\( \bar{\rho} \) and \( P_{\text{eff}} \)). Of course, the external pressure is set in part by our adjustable \( q_{\text{cos}} \), so it is important to examine the consequences of that choice. Higher order detailed radiative transfer effects will depend on the specific clump sizes and other internal properties, but these are not our focus here. Because of the cancellation of the exact size and density contrast (and correspondingly clump mass), the above relations hold for an arbitrary spectrum of clump masses, sizes and/or densities.

The column density along a given line of sight can then be integrated outward from the BH. For each integration step \( \Delta r \) along the ray (taken to be increments of \( \epsilon h_{\text{cl,avg}} \), where \( \epsilon \approx 0.01 \ll 1 \) and \( h_{\text{cl,avg}} \) is the local smoothing length at each point), we determine the probability \( p_{\text{cl}} \) of intersecting a clump, and probabilistically assign the ray a collision or not. If there is a collision, the integrated column is increased by \( \Sigma_{\text{cl}} \). If not, the column is integrated through the ‘diffuse’ (non-clump) phase of the ISM. The mass fraction in this phase (i.e. mass fraction not in star-forming clumps) is determined implicitly in the \textsc{gadget} code (see Springel & Hernquist 2003), but is always small and should have near-unity volume filling factor.

Whether or not these assumptions are justified in detail, this provides a useful toy model, and we show that it can account for a number of observations. Moreover, on galactic scales, the assumptions above have been borne out by a large number of independent observations (Larson 1981; Scoville et al. 1987; Solomon et al. 1987; Fuller & Myers 1992; Ward-Thompson et al. 1994; Andre, Ward-Thompson & Motte 1996; Blitz & Rosolowsky 2006; Rosolowsky 2007). Of course, such clumps as observed locally could not survive the tidal forces near a supermassive BH. But even on nuclear starburst scales, it appears that the star formation efficiency per clump dynamical time is low, implying they must be quasi-viral and not wildly out of pressure equilibrium (Tan, Krumholz & McKee 2006; Krumholz & Tan 2007). Similar constraints come from clump structure in the narrow-line region (e.g. Crenshaw et al. 2000; Rice et al. 2006). And the fact that similar dimensional scalings arise from Jeans considerations implies they are likely to be generic to within factors of a few. Finally, we note that the dynamic range in column density is so large that violations of the above assumptions would have to be more than order-of-magnitude in order to qualitatively affect our conclusions.

6.3 Column densities: model and observations

Fig. 8 compares the resulting column density distribution for simulations with various \( q_{\text{cos}} \) (each sampled at a random time near their peak of accretion). We compare the distribution from the ‘smooth’ and ‘clumpy’ torus models above, and that observed. Because of the dynamic range in \( N_{\text{H}} \) predicted, we are specifically interested

\[7\] These clouds cannot, however, simply follow an extrapolation of the local GMC scalings. The local GMC size–mass relation implies an approximately constant clump surface density \( \Sigma \sim 10^{22} \text{ cm}^{-2} \). But this is much less than the mean surface density of gas already at these radii, so any substructure must obey a relation at least different in normalization.
in comparison with samples sensitive to Compton-thick populations. We compile the (estimated intrinsic) distribution of column densities determined from the INTEGRAL/the Imager on Board the INTEGRAL Satellite (IBIS) AGN sample of Malizia et al. (2009, 20–40 keV), the predominantly Swift/Burst Alert Telescope (BAT) sample of Treister et al. (2009, ~100 keV), and the nearby OIII sample in Risaliti et al. (1999) (this is a type 2-only sample, so we normalize to their estimated total fraction of type 2 AGN). The latter sample is most complete at the highest columns >10^25 cm^-2; none is sensitive to AGN with N_H > 10^26 cm^-2. At lower column densities, these are consistent with a wide variety of hard X-ray observations from e.g. Chandra and XMM (Ueda et al. 2003; La Franca et al. 2005; Silverman et al. 2005; Hasinger 2008).

Unsurprisingly, the predicted columns in the smooth torus model are uniformly large, in conflict with the observations. This is not a problem of there being ‘too much’ gas – recall that the actual total gas masses and gas densities predicted at these scales agreed well with those in observed AGN (Fig. 7). What this shows is that it is not possible to reconcile the observed central masses, gas densities and/or SFRs of AGN with their obscured fractions, without invoking some small-scale gas clumping. The problem cannot simply be that systems are observed at different states either – as pointed out in Hicks et al. (2009), several observed optically unobscured AGN have instantaneous near line-of-sight volume-averaged gas densities in <1–10 pc that should naively imply columns of N_H ~ 10^{25–26} cm^-2, similar to our predictions here without subresolution clumping. And indeed direct observations on this scale have argued for such clumping (Risaliti, Elvis & Nicastro 2002; Mason et al. 2006; Sánchez et al. 2006; Nenkova et al. 2008b; Ramos Almeida et al. 2009; Höning & Kishimoto 2010; Deo et al. 2011).

The column density distribution predicted by the clumpy torus model, on the other hand, agrees well with that observed. The basic features are easily understood: the small mass fraction in the diffuse ISM phase shifts the main peak in the N_H distribution to lower values. The tail towards larger N_H is caused by obscuration by clumps. The relative ‘flatness’ of the tail is broadly expected for vertical profiles similar to those in Fig. 3.

Although the systems plotted differ in some subtle details, there is little dependence on the parametrization of stellar feedback (our q_los parameter). Why should the column density distribution be so insensitive to stellar feedback? Most important are the factors discussed in Section 4.3, i.e. the contribution of gravitational heating which keeps the discs somewhat puffed up, and means that the gaseous scale height does not scale as strongly with q_los as might otherwise be expected.

There are also two handy ‘conspiracies’, in the clumpy torus scenario, which make the predicted column density distribution primarily a function of global, rather than local parameters. In the (near-polar) regime where p_d < 1, it is quite difficult in any model to obtain a column density radically different from those shown. This is because, even if all the mass is locked in cold clumps, a column of at least ~10^{21–22} cm^-2 will arise just from diffuse, non-star-forming galactic gas on much larger scales (see Hopkins et al. 2005c, 2006a). We do see some systematic difference in the lowest columns seen, because the exact mass in the ‘diffuse phase’ depends on the subgrid model – but for almost any reasonable model this mass is small, so these differences are all in the un-obscured range (and therefore dominated by or comparable to galaxy-scale effects). In the opposite (near-disc plane) regime, where p_d > 1, the total column encountered is ~Σ_d p_d ~ ρ Δρ r – i.e. in the optically thick regime the column density is simply the same as that of the ‘average medium’, independent of the gas properties or phase structure so long as the global dynamical properties predicted are similar (physically, this simply represents where clouds will begin to overlap, thus making a more uniform molecular medium). This is true even if we discard our assumptions of virial and/or pressure equilibrium. It is only in the intermediate column regime (which interpolates broadly between the two, so we do not expect any features or particular sensitivity to appear) where the detailed assumed model of clump properties makes some difference.
These values are consistent with various indirect constraints from attempts to model AGN SEDs (Mason et al. 2006; Shi et al. 2006; Nenkova et al. 2008a,b; Mor et al. 2009; Ramos Almeida et al. 2009; Thompson et al. 2009; Hönig & Kishimoto 2010). Almost universally, these studies have found that a similar clumpy torus is required, with a number of clumps of order several along the edge-on lines of sight, characteristic locations/outer radii of most of the clumps from \( \sim 1–100 \) pc from the BH and (where constrained) radial clump distributions with roughly power-law scaling \( \mathrm{d}N/\mathrm{d}r \sim r^{-\eta} \). We find, for our typical gas surface density profiles \( \Sigma_{\text{gas}} \propto r^{-0.5–1} \), a \( \mathrm{d}N/\mathrm{d}r \sim r^{-0.7–1.2} \) over the dynamic range of interest here.

The number of clumps can be crudely estimated from equation (20). It is straightforward to show that this equation reduces to

\[
\langle N_{\text{clumps}} \rangle = 0.5 \int \rho_{\text{cl}} \, \mathrm{d}r \sim (R h) \left( \frac{Q}{\Omega_1} \right)^{-1/2},
\]

where \( h \approx c_s/\Omega_1 \) is the scale height of the torus and \( Q \) is the usual Toomre \( Q \). For a self-regulating disc, therefore, with \( Q \sim 1 \), we naturally expect \( \langle N_{\text{clumps}} \rangle \sim (h/R)^{-1} \) to be a few. The same scaling pertains if we discard pressure equilibrium and instead assume clumps are characteristically Jeans-scale in a \( Q \sim 1 \) disc (since then the scale of clumps within \( R \) is \( \sim h \)).

The characteristic value of a few clumps is also interesting because it implies that one is almost always in the Poisson regime. This has several implications. First, there should be a large scatter between the column observed and actual viewing angle, consistent with a wide variety of observations (see references above). Secondly, clumping has a number of important radiative transfer effects, which will be discussed in subsequent work. Third, this allows for highly variable obscuration. A clump moving through the line of sight can lead to variation in the column density by several orders of magnitude. The detailed variability will depend on the clump size spectrum and other properties, but the maximal variability timescale should scale as \( \sim R_o/R \, Q(R) \); since most of these clumps are at \( \sim 0.1–1 \) pc, the constraint that clumps not be tidally shredded \( (\rho_{\text{cl}} \gtrsim M_{\text{BH}}/R^3) \), and \( N_{\text{cl}} > 1 \) sets an upper limit to the variability timescale of \( \sim 5–100 \) yr (for \( 0.1–1 \) pc), for a \( 10^8 M_\odot \) BH. For partial obscuration, a more realistic clump density contrast and/or larger clump number, the obscuration could vary on a time-scale 0.01–0.1 times this (i.e. months–year). Such rapid, extreme variability in X-ray obscuration has been seen in several AGN (Risaliti et al. 2002, 2005; Immler et al. 2003; Lamer, Uttley & McHardy 2003; Matt, Guainazzi & Maiolino 2003; Fruscione et al. 2005; Guainazzi et al. 2005).

### 7 The Obscured Fraction and Torus Properties

We now examine how the column density distribution depends on global properties. For the sake of comparison with observations, we parametrize the distribution by means of the ‘obscured fraction’: specifically, the fraction above a given column density \( N_{\text{H}} > 10^{22} \) cm\(^{-2} \) (a value typically adopted in observational studies). Henceforth, we ignore the ‘smooth torus’ model – it does not agree with observations and gives uninteresting (always near-unity) obscured fractions.

Fig. 10 compares the obscured fraction in the ‘clumpy’ model with a number of nuclear properties. For each simulation, we measure the relevant properties at randomly sampled times and viewing angles. We show the obscuration versus total mass inside some small radius (essentially, the BH mass), versus gas mass, versus the nuclear SFR and versus the BH Eddington ratio.

Unsurprisingly, \( f_{\text{obs}} \) increases with the gas mass inside a small radius \( \sim 1 \) pc. Note, however, that the correlation is weak:

---

Fig. 9. Obscuration properties from the clumpy-torus model in Fig. 8 versus viewing angle (\( \cos \theta = 0 \) is edge-on; \( \cos \theta = \pm 1 \) is polar). For \( q_{\cos} \) survey: Top: integrated column. Dense clusters of points at lower \( N_{\text{H}} \) pass through the diffuse ISM only, points scattered to higher \( N_{\text{H}} \) encounter clumps along the line of sight. The appearance of bimodality is somewhat artificial (see Fig. 8; for most simulations, there are not actually two peaks). Each point is a sightline to the BH in a given simulation. Bottom: expectation value for the number of clumps encountered along a given sightline. The clump number density is always Poisson, with \( \sim \) a few clumps along the typical edge-on line of sight, and rapidly declines above a scale height of \( h/R \sim 1/2 \).

---

© 2011 The Authors, MNRAS 420, 320–339
Monthly Notices of the Royal Astronomical Society © 2011 RAS
Figure 10. Obscured fraction (fraction of sightlines that encounter a column density $N_H > 10^{22}$ cm$^{-2}$) for our simulations as a function of various properties. Each simulation is sampled at several random times both before, during and after its period of peak activity, and we show results from all simulations with different initial conditions and $q_{\text{eos}}$ values. Points are coloured by the instantaneous gas fraction inside of 10 pc (from red at $f_{\text{gas}} < 0.05$ to dark blue at $f_{\text{gas}} > 0.8$). Top left: versus $M_{\text{BH}}$ (we add ~0.1 dex scatter in $M_{\text{BH}}$ so the points can be distinguished). For otherwise the same conditions (e.g. $M_{\text{gas}}$), $f_{\text{obscured}}$ declines with $M_{\text{BH}}$ as approximately $\propto M_{\text{BH}}^{-1/2}$. Top right: versus $M_{\text{gas}}$ inside 1 pc, at fixed $M_{\text{BH}}$. Obscured fractions increase weakly ($\propto M_{\text{gas}}^{1/4}$) with gas mass. Bottom left: versus the SFR inside 10 pc. There is a weak increase in $f_{\text{obscured}}$ with the circumnuclear SFR, driven by the dependence on $M_{\text{gas}}$ (slightly weaker $\propto \dot{M}_{\text{gas}}^{0.15-0.2}$, because $M_{\text{gas}}^{0.15-0.2}$ is superlinear in $M_{\text{gas}}$). Bottom right: versus inflow rate into small radii, at fixed $M_{\text{BH}} = 3 \times 10^7 M_\odot$ (if this continued to arbitrarily small radii, this would be proportional to the Eddington ratio $M_{\text{BH}}/\dot{M}_{\text{Edd}} = L/L_{\text{Edd}}$). The dependence here is significant, but very weak ($\propto \dot{M}_{\text{gas}}^{0.1}$).

$f_{\text{obsc}} \propto M_{\text{gas}}^{1/4}$. The midplane columns should increase more rapidly with $M_{\text{gas}}$, but these are already optically thick – the obscured fraction grows slowly with the fraction of sightlines above the disc that (at higher column) become optically thick. More interesting is the correlation this implies – $f_{\text{obsc}}$ also increases with the nuclear SFR. Behaviour along these lines has been observed at a wide variety of scales – type 2 AGN are more likely to be found in more rapidly star-forming hosts, and/or hosts with younger stellar populations (Brotherton et al. 1999; Canalizo & Stockton 2001; Jahnke, Kuhlenbrodt & Wisotzki 2004; Yip et al. 2004; Zakamska et al. 2006; Nandra et al. 2007; Silverman et al. 2008). The observational correlation appears to be particularly strong when the nuclear stellar populations are isolated (Imanishi 2002; Imanishi & Wada 2004; Davies et al. 2007; Shi et al. 2007; Wang et al. 2007). Note that the SFRs inside of ~10 pc can reach large ~10 M$_\odot$ yr$^{-1}$ values; however, as shown in Hopkins & Quataert (2010b, fig. 14), this is correlated with the BH inflow rates on these scales as $M_\ast \sim M_{\text{BH}}$ (again, both tracing the gas mass supply) – for a less extreme quasar the ‘zero-point’ expected inside ~10 pc would be more like ~0.1–1 M$_\odot$ yr$^{-1}$.

Conversely, ULIRGs and mergers with more pronounced star formation in their nuclei are more likely to host obscured Seyferts or quasars, whereas those with slightly older populations are more likely to exhibit type 1 signatures (Sanders 1999; Farrah et al. 2003, 2005; Dasyra et al. 2006; Guyon, Sanders & Stockton 2006; Yuan, Kewley & Sanders 2010).

Most likely, at least some of this trend owes to the role of AGN feedback in clearing away some of the gas and dust (see e.g. Sanders et al. 1988a; Granato et al. 2004; Hopkins et al. 2005a,c; Narayanan et al. 2006; Hopkins 2011), but it can simply arise as we see here from the larger gas and dust supply ‘burying’ the AGN until star formation exhausts much of that material. We stress that this is a true nuclear-scale (<10 pc) correlation here, and the nuclear star formation contributes negligibly (~0.3 per cent) to the total SFR; there is not necessarily any predicted correlation between the AGN obscuration and the total/large-scale galaxy SFR.

Given similar gas properties, the obscured fraction decreases with BH mass. This is expected because the BH gravity provides a stabilizing force that tries to ‘flatten’ the torus (for fixed gas...
The dynamical AGN torus

properties, the disc $h/R \propto M_{\text{BH}}^{1/2}$. If more luminous AGN are, on average, more massive BHs, then this suggests an inverse correlation between quasi-stellar object (QSO) luminosity and obscured fractions. Indeed, the existence of such an apparent correlation is well established (Hill et al. 1996; Simpson et al. 1999; Simpson \\& Rawlings 2000; Willott et al. 2000; Steffen et al. 2003; Ueda et al. 2003; Grimes et al. 2004; Sazonov \\& Revnivtsev 2004; Barger \\& Cowie 2005; Hao et al. 2005; Simpson 2005; Gilli et al. 2007; Hickox et al. 2007; Hasinger 2008). However, it is still unclear precisely how much of this correlation owes to alternative possibilities such as simple dilution by the host galaxy and/or differences in the Eddington ratio distribution and accretion rate (for a detailed discussion, see Hopkins et al. 2009b, and references therein). Moreover, without a full cosmological model to predict e.g. the distribution of active BH masses and Eddington ratios, we cannot forward model the BH luminosity distribution to construct a direct comparison with observations. But the predicted scaling here is not especially strong; it may well be that additional physics is needed to recover the full observed correlation – most commonly, AGN feedback is invoked to ‘blow away’ some of the torus in the most luminous systems (see references above).

8 DISCUSSION

We have studied AGN obscuration in a series of multiscale hydrodynamic simulations that can self-consistently follow gas from $>10\,$kpc galactic scales to $<0.1\,$pc. These simulations include the full self-gravity of stars and gas (along with BHs and dark matter), gas cooling and star formation, along with varied prescriptions for feedback from young stars; these are all critical to the behaviour we see, and have not before been simultaneously modelled on nuclear scales. In these simulations, inflows from large scales, when sufficiently large, lead to a cascade of instabilities on small scales, ultimately yielding large nuclear gas masses and accretion rates on to the AGN. The scenario is qualitatively similar to the ‘bars within bars’ model, but there is a high degree of variability and morphological diversity at each stage (with spirals, bars, clumps, flocculent structures, all present and alternatively powering inflows and outflows) – a more apt description would be ‘stuff within stuff’ (Hopkins \\& Quataert 2010b). Once gas nears the radius of influence of the BH, it generically forms an unstable $m = 1$ mode (a lopsided or eccentric disc) that slowly precesses about the BH (Hopkins \\& Quataert 2010a). The stellar and gas disc precess differently, leading to strong gravitational torques that can drive accretion rates of up to $\sim 10\,M_\odot\,\text{yr}^{-1}$ on to the BH.

In this paper, we show that these nuclear, lopsided discs in fact naturally account for the long-invoked ‘toroidal obscuring region’ used to explain the obscuration of type 2 AGN. Up to now, these models have been essentially phenomenological – we show for the first time the formation of sub- pc scale obscuring structures from galaxy-scale inflows, and in a suite of $\sim 100$ simulations show that they are quite generic and arise ubiquitously with this inflow scenario. We show that the global dynamical properties – gas and stellar densities, density profiles, kinematics, gas fractions and SFRs, agree well with observations of AGN obscuring regions from scales as small as $0.1$ to $\gtrsim 100\,$pc. This implies a fundamentally new paradigm in which to view the obscuring region or ‘torus’. Far from being a passive bystander or simple fuel reservoir for the accretion process, it is itself the driver of that accretion. The torus is the gravitational structure on scales within the BH radius of influence that torques the gas and forces continuous gas inflow on to the BH. The same lopsided modes that drive accretion can also provide the scale height, column density distribution and characteristic gas properties of the structure.

As such, the predicted torii have non-trivial substructure: both small-scale clumping in the gas (discussed below), global $m = 1$ patterns, and warps/twists arising from bending modes at a range of radii. On large scales, the $m = 1$ modes tend to manifest as lopsided/eccentric discs, or one-armed spirals; on small scales, they become more tightly wound spirals. Their typical amplitude in surface density is expected to be at the $\sim 10\%$ level (see Hopkins \\& Quataert 2010b); the amplitude of induced non-circular velocities and corresponding magnitude of ‘offsets’ of the BH from the spatial centre of the galaxy nuclei (in units of the BH radius of influence) are about the same. Maser observations may show indications of asymmetry in the structure around nuclei (Schnirrer et al. 2000; Greenhill et al. 2003; Fruscione et al. 2005; Kondratko et al. 2005, 2006b); it is also possible that measurements of the velocity structure of e.g. molecular emission lines from the torus region may be able to measure such asymmetries in the near future.

We have argued that a large number of obscuration properties traditionally associated with ‘feedback’ processes from AGN and star formation may, in fact, be explained by purely gravitational physics. Even in the absence of feedback, properly including the full self-gravity of gas and stars leads to discs with large $h/R$, sufficient to account for the observed column density distribution. This arises because of a combination of large-scale warps and twists (for example, where the lopsided disc mode meets an outer bar) and bending modes within the disc itself. The latter will, even in the absence of any large-scale twists or warps, tend to pump up $h/R$ wherever the eccentric disc mode is excited until an order unity $h/R \sim |a|$ is reached. Since bending modes are fast modes (pattern speed $\sim \Omega$), this can continuously transfer energy from the orbital motion to vertical motions on a single dynamical time, maintaining vertical scale heights even when the cooling time is arbitrarily short.

These warps and twists also naturally lead to the observed lack of correlation between nuclear-scale disc inclination angles and those of their parent/host galaxies. This will be even more prominent in systems which are driven on large scales by mergers, but can occur even in entirely secularly fuelled AGN. They also account for observed gas velocity dispersions in AGN nuclei, and the correlations between those dispersions and quantities such as the local gas mass, SFR and mass in young stars (all via their inherent dynamical correlations, not via any feedback channel). The efficiency of gravitational torques and induced inflow also naturally leads to convergence in nuclear gas masses and density profiles, leaving relic ‘cusps’ similar to those observed (Hopkins & Quataert 2011a).

These mechanisms can naturally explain observed global quantities such as the gas scale heights, masses and density profiles. However, modelling the actual obscured fraction of AGN requires a more explicit model for the actual substructure on Jeans mass scales and well below, in the ISM surrounding black holes. We show in fact that any model which matches the observed dynamical properties (particularly global gas masses), but assumes ‘smooth’ gas (uniformly distributed, say, out to some scale height corresponding to the average obscured fraction), will simultaneously fail to explain the observed column density distributions. A natural explanation for this discrepancy is that the gas is clumpy on multiple scales, broadening the column density distribution along all sightlines. There must, in fact, be structure on the relevant scales, since we know there is star formation at these radii (so some gas must
reside in dense, tidally bound star-forming clumps). Unfortunately, our present models do not explicitly resolve the necessary physics of star formation and GMC formation/destruction via stellar feedback needed in order to explicitly simulate the structure of the gas down to these scales.

However, we find that we can obtain predicted column density distributions in good agreement with those observed if we assume that whatever subgrid clumps exist obey a couple of basic assumptions: namely that they are (at least at the order-of-magnitude level) near both virial and pressure equilibrium (or, instead of pressure equilibrium, that they are Jeans-scale in a $Q \sim 1$ self-regulating disc). These assumptions are sufficient to (statistically) predict the column density distribution that would be observed, regardless of the actual clump mass spectrum and physical origin (and without any adjustable parameter introduced). The predictions agree quite well with the column density distribution of both unobscured, obscured and Compton-thick AGN. This suggests that these basic properties should still hold for substructure in these regions, and that – if so – the uncertain quantities in our simulations (such as the feedback prescription and star formation recipe), make no dramatic difference in the column density distribution, since its key properties are set by the basic dynamics above. Essentially, if these assumptions hold, the distribution of observed column densities towards AGN is itself a natural consequence of gravitational clumping at the Jeans length/mass in a self-gravitating, globally isothermal AGN (see e.g. Hopkins et al. (2005c, 2006a), Hopkins & Hernquist (2006) and Hayward et al. (2011), and we refer to those papers for more details. Observations have also made it clear that a significant fraction of obscuration must come from host galaxies (especially in starbursts and edge-on discs; see e.g. Rigby et al. 2006; Zakamska et al. 2006; Martínez-Sansigre et al. 2009; Lagos et al. 2011). Since our current simulations are the first to simultaneously resolve both the nuclear scales where very large columns arise, and the galaxy scales where more moderate but potentially more isotropic (or at least differently oriented) columns can arise, it will be interesting to investigate the relative contributions to obscuration from different scales, as a function of evolutionary stage and galaxy/BH properties. Because the AGN spectrum changes as it moves through the inner obscuring regions, this will require a full treatment of radiative transfer, as described above.

ACKNOWLEDGMENTS

We thank Eliot Quataert, Nadia Zakamska, Jenny Greene, Patrik Jonsson and Josh Younger for helpful discussions in the development of this work. Support for PFH was provided by the Miller Institute for Basic Research in Science, University of California Berkeley. DN and LH acknowledge support from the NSF via grant AST-1009452.

REFERENCES

Alexander D. M., Bauer F. E., Chapman S. C., Smail I., Blain A. W., Brandt W. N., Ivison R. J., 2005, ApJ, 632, 736
Alexander D. M. et al., 2008, ApJ, 687, 835
Aller M. C., Richstone D. O., 2007, ApJ, 665, 120
Andre P., Ward-Thompson D., Motte F., 1996, A&A, 314, 625
Antonucci R., 1993, ARA&A, 31, 473
Barger A. J., Cowie L. L., 2005, ApJ, 635, 115
Barnes J. E., Hernquist L. E., 1991, ApJ, 370, L65
Barnes J. E., Hernquist L., 1996, ApJ, 471, 115
Bender R. et al., 2005, ApJ, 631, 280
Blitz L., Rosolowsky E., 2006, ApJ, 650, 933
Bournaud F., Elmegreen B. G., Elmegreen D. M., 2007, ApJ, 670, 237
Braatz J. A., Henkel C., Greenhill L. J., Moran J. M., Wilson A. S., 2004, ApJ, 617, L29
Brotherton M. S. et al., 1999, ApJ, 520, L87
Bryant P. M., Scoville N. Z., 1999, AJ, 117, 2632
