Recoiling black holes: prospects for detection and implications of spin alignment

Laura Blecha,[1][*][†] Debora Sijacki,[2] Luke Zoltan Kelley,[3] Paul Torrey,[4] Mark Vogelsberger,[4] Dylan Nelson,[3] Volker Springel,[5,6] Gregory Snyder[7] and Lars Hernquist[3]

[1]*Department of Astronomy, University of Maryland, 1113 PSC, Bldg. 415, College Park, MD 20742, USA*
[2]*Institute of Astronomy and Kavli Institute for Cosmology, Cambridge University, Madingley Road, Cambridge CB3 0HA, UK*
[3]*Harvard–Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA*
[4]*Department of Physics, Kavli Institute for Astrophysics and Space Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA*
[5]*Heidelberg Institute for Theoretical Studies, Schloss-Wolfsbrunnenweg 35, D-69118 Heidelberg, Germany*
[6]*Zentrum für Astronomie der Universität Heidelberg, ARI, Mönchhofstr. 12-14, D-69120 Heidelberg, Germany*
[7]*Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA*

ABSTRACT

Supermassive black hole (BH) mergers produce powerful gravitational wave emission. Asymmetry in this emission imparts a recoil kick to the merged BH, which can eject the BH from its host galaxy altogether. Recoiling BHs could be observed as offset active galactic nuclei (AGN). Several candidates have been identified, but systematic searches have been hampered by large uncertainties regarding their observability. By extracting merging BHs and host galaxy properties from the Illustris cosmological simulations, we have developed a comprehensive model for recoiling AGN. Here, for the first time, we model the effects of BH spin alignment and recoil dynamics based on the gas richness of host galaxies. We predict that if BH spins are not highly aligned, seeing-limited observations could resolve offset AGN, making them promising targets for all-sky surveys. For randomly oriented spins, $\lesssim 10$ spatially offset AGN may be detectable in *Hubble Space Telescope*-Cosmological Evolution Survey, and $> 10^3$ could be found with the Panoramic Survey Telescope & Rapid Response System (Pan-STARRS), the Large Synoptic Survey Telescope (LSST), *Euclid*, and the Wide-Field Infrared Survey Telescope (WFIRST). Nearly a thousand velocity offset AGN are predicted within the Sloan Digital Sky Survey (SDSS) footprint; the rarity of large broad-line offsets among SDSS quasars is likely due in part to selection effects but suggests that spin alignment plays a role in suppressing recoils. None the less, in our most physically motivated model where alignment occurs only in gas-rich mergers, hundreds of offset AGN should be found in all-sky surveys. Our findings strongly motivate a dedicated search for recoiling AGN.

Key words: accretion, accretion discs – black hole physics – gravitational waves – hydrodynamics – galaxies: active – galaxies: interactions.

1 INTRODUCTION

Mergers between galaxies are important drivers of evolution. They grow stellar bulges, evolve galaxies along the Hubble sequence, and can trigger both star formation and accretion on to central supermassive black holes (BHs), which in turn influence the host via feedback processes (e.g. Sanders et al. 1988; Barnes & Hernquist 1991, 1996; Mihos & Hernquist 1996; Wyithe & Loeb 2003; Di Matteo, Springel & Hernquist 2005; Hopkins et al. 2006, 2008a). While the majority of active galactic nuclei (AGN) in the local Universe do not appear to be triggered by mergers (Cisternas et al. 2011; Schawinski et al. 2011; Kocevski et al. 2012), a clear correlation between merging activity and AGN fuelling has been observed (e.g. Koss et al. 2010; Ellison et al. 2011; Satyapal et al. 2014). There is also evidence that the majority of the highest luminosity quasars are triggered by major mergers (e.g. Urrutia, Lacy & Becker 2008; Treister et al. 2012; Hopkins, Kocevski & Bundy 2014).

In addition to fuelling rapid growth, a major merger between comparable-mass galaxies, each containing a central BH, will inevitably lead to the formation of a BH binary. This binary may

[*]Einstein Fellow.
[†] E-mail: lblecha@astro.umd.edu

eventually merge, releasing vast amounts of energy as gravitational waves (GWs). Such events could be detected directly in the coming years via pulsar timing arrays (PTAs; e.g. Sesana, Vecchio & Colacino 2008; Sesana 2013), or with a future space-based GW interferometer such as *eLISA* (e.g. Amaro-Seoane et al. 2012). Additionally, any asymmetry in the merging BH system – unequal masses or spins – will produce asymmetric GW emission and a net momentum flux, causing the merged BH to 'recoil' in the opposite direction at the moment of merger (Peres 1962; Bekenstein 1973). Numerical relativity simulations in recent years have demonstrated that these kicks may be very large, up to 5000 km s^{-1}, implying that some BHs could be ejected from galaxies entirely (Campanelli et al. 2007; Lousto & Zlochower 2011).

These 'superkicks' of more than a thousand km s^{-1} should be rare, but smaller recoils may still displace BHs from galactic nuclei for substantial periods of time. Recoil events may have been more frequent in the early Universe, when galactic escape speeds were lower and merger rates were higher. This places extra constraints on models of BH seed formation that must produce massive quasars by $z \sim 7$ (e.g. Volonteri & Rees 2006; Li et al. 2007; Sijacki, Springel & Haehnelt 2009). At lower redshift, BH ejections must be relatively rare in massive ellipticals, at least, owing to the observed ubiquity of central BHs in such galaxies for which mass measurements have been obtained. The sample of galaxies with direct mass measurements is still modest, however, and the lower mass range is largely unexplored. A large sample of dynamical mass measurements in massive galaxies, such as brightest cluster galaxies (BCGs), could yield strong evidence for a past recoil event if a galaxy with a 'missing' or very undermassive BH were found (Gerosa & Sesana 2015). Even if the BH is not ejected entirely, a marginally bound BH may oscillate in the galactic potential for many Gyr before returning to the centre (e.g. Madau & Quataert 2004; Blecha & Loeb 2008; Gualandris & Merritt 2008; Komossa & Merritt 2008).

These GW recoil events have implications for BH–galaxy co-evolution. They increase the intrinsic scatter to the observed BH–bulge relations (Volonteri 2007), even when the BH is not ejected entirely from the host (Blecha et al. 2011; Sijacki, Springel & Haehnelt 2011). The displacement of the BH from the galactic centre also reduces the amount of AGN feedback imparted to the nuclear region, allowing a denser stellar cusp to form and delaying the galaxy's transition from the blue to the red sequence (Blecha et al. 2011).

A major uncertainty in predicting the effects of GW recoil is that the kick velocity distribution depends sensitively on the pre-merger BH spin vectors, which are not known. In particular, if the BH binary is embedded in a circumbinary gas disc that drives the binary evolution, torques from the disc may efficiently align the BH spins prior to a merger (e.g. Bogdanović, Reynolds & Miller 2007; Dotti et al. 2010; Miller & Krolik 2013). The maximum recoil kick for perfectly aligned spins is <200 km s^{-1}, even for maximally spinning BHs, and kicks $\gtrsim 600$ km s^{-1} do not occur if the spins are aligned to within a few degrees of the orbital angular momentum. However, even in gas-rich systems where a circumnuclear disc forms, spin alignment may not always be efficient (Lodato & Gerosa 2013). Also, if the accretion flow is chaotic, characterized by low-mass gas streams ($\lesssim 1$ per cent M_{BH}) infalling from random directions, the BH spins may not be efficiently aligned prior to merger. In certain cases, circumbinary discs could even produce stable *counter*-alignment if the disc angular momentum is low and the initial BH misalignment is large (Nixon 2012), and retrograde discs may drive the BHs to co-alescence before alignment can occur (Schnittman & Krolik 2015). Such configurations may occur if outer regions of the disc become

unstable to star formation, or if the nuclear gas inflow is intermittent (e.g. King & Pringle 2006; Lodato et al. 2009; King & Nixon 2013). Furthermore, stellar dynamics may be the dominant driver of BH binary inspiral in some systems, especially in triaxial, gas-poor mergers (e.g. Berczik et al. 2006; Khan, Just & Merritt 2011; Holley-Bockelmann & Khan 2015). These stellar interactions typically will not cause any preferential alignment of BH spins (Merritt & Vasiliev 2012). In all cases, even in the final stages of BH inspiral after the BHs have dynamically decoupled from their surroundings, their spins will undergo further evolution via relativistic precession. This precession can substantially alter the final BH spins prior to merger, driving them towards alignment in some cases and increasing misalignment in others, thereby decreasing or increasing the resultant recoil velocity depending on the mass ratio and initial spins of the BH binary (Berti, Kesden & Sperhake 2012; Gerosa et al. 2015; Kesden et al. 2015).

In addition to its possible role in aligning BH spins, gas in merging galaxies may strongly influence the recoil trajectories themselves. The inflow of cold gas to central regions during a gas-rich major merger can dramatically increase the central escape speed of the galaxy for a period of time, and gas drag can also help retain BHs in galactic nuclei (Blecha et al. 2011; Sijacki et al. 2011). Thus, even if spin alignment via gas dynamics is less efficient than numerical results suggest, large BH displacements in gas-rich systems may be rare.

In contrast, recoil events in gas-poor mergers can be much longer lived, particularly if the inner stellar profile has a shallow density core. Cuspy stellar density profiles may be transformed into cores (and cores may be enlarged) by BH binary inspiral (e.g. Quinlan 1996; Milosavljević & Merritt 2001; Yu 2002) and by the impulsive removal of the BH and its bound cusp of stars (Merritt et al. 2004).

A kicked BH that is actively accreting will carry along its inner accretion disc and broad emission line (BL) region, and may also sweep up gas from its surroundings, such that it could be observed as an AGN that is spatially or kinematically offset from its host galaxy[1] (e.g. Madau & Quataert 2004; Loeb 2007; Blecha & Loeb 2008; Volonteri & Madau 2008; Fujita 2009). In the case of kinematic offsets, the line-of-sight (LOS) velocity of the recoiling BH could be detected in the object's spectrum as an offset between the BLs and the narrow emission lines (NLs), which will remain bound to the host galaxy (e.g. Merritt et al. 2006). An offset could also be detectable between the AGN BLs and the rest-frame redshift of the host galaxy itself. Observable offset AGN lifetimes can be up to tens of Myr for both spatial and kinematic offsets, for a wide range in kick speeds (Blecha et al. 2011; Guedes et al. 2011). Prior to the direct detection of GWs, recoiling AGN could offer the best evidence of BH mergers.

In the last few years, a handful of candidate (spatially or velocity) offset AGN have been identified (Komossa, Zhou & Lu 2008; Comerford et al. 2009; Shields et al. 2009b; Batcheldor et al. 2010; Civano et al. 2010; Robinson et al. 2010; Steinhardt et al. 2012; Koss et al. 2014; Lena et al. 2014). None have yet been confirmed (see Komossa 2012, for a review), and the recoil scenario is disfavoured for some (Shields et al. 2009b; Decarli et al. 2014), but two candidates in particular appear promising. CID-42 is a highly disturbed galaxy with an AGN that appears both spatially *and* kinematically offset from the host nucleus, though the possibility that

---

[1] Following convention, we refer to any actively accreting BH as an 'AGN' but note that in the event of a GW recoil, active BHs will generally *not* be coincident with the galactic nucleus.

the galaxy is in a pre-merger phase, with a second, quiescent BH in the nucleus, is difficult to rule out (Comerford et al. 2009; Civano et al. 2010, 2012; Blecha et al. 2013; Novak et al. 2015). SDSS 1133 is a bright point source offset from a nearby dwarf galaxy; it has no evidence for an extended emission component down to a scale of 12 pc (Koss et al. 2014) and is observed over a 65 yr baseline. An alternative scenario in which the object experienced an extreme, 50 yr luminous blue variable-type outburst followed by an unusually long-lived Type IIn supernova (SN) cannot be excluded, but further monitoring of the source should be conclusive in the near future.

These largely serendipitous discoveries motivate a systematic search for offset AGN. To date, a few searches for velocity offset quasars have been carried out (Bonning, Shields & Salviander 2007; Tsalmantza et al. 2011; Eracleous et al. 2012). Bonning et al. (2007) searched for consistent, symmetric velocity shifts in broad Hβ and Mg II lines in SDSS DR5 quasar spectra and found a null result for $\Delta v > 800 \, \text{km s}^{-1}$. Tsalmantza et al. (2011) searched for offset BLs among SDSS DR7 quasars, focusing on binary BH candidates, and identified 32 quasars with significant BL shifts. Eracleous et al. (2012) similarly examined the SDSS DR7 quasars for broad Hβ lines offset by $\Delta v > 1000 \, \text{km s}^{-1}$ and found 88 such objects. In these samples, some objects have BL shifts $>5000 \, \text{km s}^{-1}$, the theoretical maximum recoil velocity, and some are best explained as double-peaked emitters. Offset BLs may also be produced by high-velocity outflows or the orbital motion of a single active BH in a sub-parsec binary pair. These results indicate that recoiling AGN with observable velocity offsets are rare among SDSS quasars, if present at all.

Spatially offset AGN remain largely unexplored. Lena et al. (2014) have conducted the only systematic search for spatial offsets thus far, focusing on very small-scale offsets (<10 pc) in nearby giant core ellipticals. They find displacements between the AGN and the galactic photocentre in 6 of 14 ellipticals, including M87 where a displacement was previously noted by Batcheldor et al. (2010). However, the alignment of the displacement axis with a radio jet in four of the AGN argues against the recoil scenario for these objects. No searches have yet been carried out for AGN with larger spatial offsets.

Designing a targeted search for offset AGN requires better theoretical understanding of when and where significant displacements are most likely to occur. Using semi-analytic models, and assuming randomly oriented BH spins, Volonteri & Madau (2008, hereafter VM08) predict rates of observable spatially offset AGN; their results are of particular relevance to this work and are compared with our findings in Section 4.2. The effect of pre-merger spin alignment remains a major open question, and the influence of gas on recoil dynamics has been studied only for isolated galaxy mergers, rather than in a cosmological framework. The host galaxy properties of offset AGN are also unexplored, with the exception of a recent study (Gerosa & Sesana 2015), which concludes that a larger sample of BH mass measurements in BCGs could find evidence for past superkicks.

Here, we address these open questions by constructing a model for the observability of spatially and velocity offset AGN, utilizing data from the state-of-the-art cosmological hydrodynamic simulation Illustris (e.g. Genel et al. 2014; Vogelsberger et al. 2014a,b; Nelson et al. 2015). The Illustris simulation has been shown to successfully reproduce many observed properties of galaxies and their BHs, including the stellar and BH mass functions, cosmic star formation rate density, galaxy merger rate, baryonic Tully–Fisher relation, and quasar luminosity function

(Genel et al. 2014; Vogelsberger et al. 2014a; Sijacki et al. 2015). We examine the offset lifetimes and space density of observable offset AGN that could be detected in various surveys, and we make the first predictions regarding the host galaxy properties of offset AGN. We consider a range of possible pre-merger BH spin distributions in order to study the effect of spin alignment on offset AGN observability, and we discuss how detections of recoils might be used to constrain the degree of alignment. These predictions will be testable in existing and future surveys with *Hubble Space Telescope* (*HST*), *Chandra*, Pan-STARRS, the James Webb Space Telescope, LSST, *Euclid*, and WFIRST.

The outline of this paper is as follows. The cosmological simulations are described in Section 2.1, BH spin models and recoil velocity distributions are detailed in Section 2.2, and the model for recoiling AGN is outlined in Sections 2.3–2.5. Section 3 describes our results concerning the BH merger rate (Section 3.1), recoil trajectories (Section 3.2), recoiling AGN properties (Section 3.3), spatial and velocity offsets (Section 3.4), offset AGN observability (Section 3.5), and host galaxy properties (Section 3.6). The dependence of our results on model assumptions is discussed in Section 3.7. We discuss the prospects for identification and follow-up of promising recoil candidates in Section 4.1 and compare with previous work in Section 4.2. Our conclusions are summarized in Section 5.

2 METHODOLOGY

Our basic procedure for constructing a recoiling AGN model is as follows: (1) extract the properties of merging BH pairs and their host galaxies from the Illustris cosmological simulations, (2) assign pre-merger spins to the BHs and calculate the resulting recoil kick velocity, according to an assumed spin distribution, (3) construct an analytic potential model for the galaxy merger remnant at the time of BH merger, based on the properties of the progenitor galaxies, (4) integrate the trajectory of the recoiling BH in this potential, including stellar dynamical friction, and (5) calculate the AGN luminosity of the recoiling BH at each timestep and determine its observability as an offset AGN. Here we describe this procedure in detail.

2.1 Cosmological simulations

We use the data from the Illustris cosmological simulation project (e.g. Genel et al. 2014; Vogelsberger et al. 2014a) as the basis for our models. The Illustris simulations were conducted with the moving-mesh hydrodynamics code AREPO (Springel 2010). The highest resolution simulation ('Illustris-1') spans a comoving volume of $(106.5 \, \text{Mpc})^3$ and has 2×1820^3 resolution elements, with a dark matter (DM) mass resolution of $6.26 \times 10^6 \, \text{M}_\odot$ and a typical gas cell mass of $1.26 \times 10^6 \, \text{M}_\odot$. The 'Illustris-2' and 'Illustris-3' simulations have the same volume with 2×910^3 and 2×455^3 resolution elements, respectively. The following cosmological parameters, which are consistent with the *Wilkinson Microwave Anisotropy Probe* 9 measurements, are used in the simulation and assumed throughout this paper: $\Omega_\text{m} = 0.2726$, $\Omega_\Lambda = 0.7274$, $\Omega_\text{b} = 0.0456$, $\sigma_8 = 0.809$, and $H_0 = 100 \, h \, \text{km s}^{-1} \, \text{Mpc}^{-1}$ with $h = 0.704$. Sub-resolution physical models are included for primordial and metal-line cooling, star formation and stellar feedback (Springel & Hernquist 2003), gas recycling and chemical enrichment, and BH seeding, accretion, and feedback. These prescriptions are described in detail in Vogelsberger et al. (2013) and Torrey et al. (2014). Here, we outline the relevant aspects of the prescriptions for BH formation, growth, and feedback in the Illustris simulation. These

models successfully reproduce BH and AGN populations in good agreement with observational constraints, including the BH mass function at $z = 0$ and the quasar luminosity function (Sijacki et al. 2015).

Throughout the simulation, an on-the-fly Friends-of-Friends (FOF) algorithm is used to identify bound subhaloes, and a BH particle with a seed mass of $10^5 \, h^{-1} \, M_\odot$ is placed in each subhalo more massive than $7.1 \times 10^{10} \, M_\odot$ that does not already contain a BH. Gas accretion on to BHs is parametrized in terms of a Bondi-like formula, capped at the Eddington limit. A repositioning scheme is implemented to centre BHs on the gravitational potential minimum of their hosts, which prevents spurious BH wandering from numerical two-body scattering. As a result, BH velocities are ill-defined, and the velocity of the BH relative to the surrounding gas is not taken into account in calculating the accretion rate. Additionally, a pressure criterion is introduced to reduce the BH accretion rate if the gas pressure near the BH is insufficient to compress the gas against thermal feedback, which prevents the formation of unphysically large, hot gas bubbles around accreting BHs.

In addition to gas accretion, BHs are allowed to grow via mergers. Two BHs are assumed to have merged when their separation falls below the gravitational softening length. Owing to the effect of the repositioning scheme on BH velocities, no criterion is imposed on the relative BH velocity in order for them to merge. Furthermore, in some instances, when a satellite halo containing a BH is determined to have merged with the central halo, the repositioning scheme may cause the low-mass satellite BH to merge with the central BH on an artificially short time-scale. We find that this effect is largely confined to BHs with mass $<10^6 \, M_\odot$, and we accordingly exclude from our analysis all mergers in which either BH mass is below this value.

Three distinct modes of BH feedback are included in the simulation. Thermal or 'quasar mode' feedback is introduced by coupling 5 per cent of the AGN luminosity to surrounding gas particles as thermal energy (Springel, Di Matteo & Hernquist 2005). When the accretion rate falls below 0.05 of the Eddington rate, AGN feedback is assumed to be dominated by a radiatively inefficient 'radio mode', in which hot, buoyant bubbles with a radius of 100 kpc are injected into the ambient gas (Sijacki et al. 2007). Additionally, 'radiative' feedback is included by assuming a fixed AGN spectral energy distribution (SED) and modifying the net gas cooling rate in the presence of strong ionizing radiation. This radiative feedback is most relevant for BHs accreting near the Eddington limit.

In addition to this constraint on the BH mass, we also require that each host halo have a minimum of 300 DM and 80 stellar particles, which in Illustris-1 corresponds to masses of $M_{DM} = 2 \times 10^9 \, M_\odot$ and $M_* = 10^8 \, M_\odot$. In practice, the BH mass constraint removes most of the lowest mass dwarf galaxies from our sample, and those remaining are largely satellites that undergo minor mergers with larger galaxies. We find that less than 0.1 per cent of mergers have a total halo mass $<10^{10} \, M_\odot$ and ~ 0.1 per cent have a total stellar mass $<10^9 \, M_\odot$.

The BH merger time-scales for more massive BHs can also be underestimated in the simulation to a lesser degree, owing to the repositioning scheme and the resolution limit, which impose the condition that BH binary inspiral is always rapid on sub-resolution scales. This may be a reasonable assumption in the case of gas-rich major mergers, where circumnuclear gas discs and highly asymmetric stellar potentials can greatly reduce the BH inspiral time-scales (e.g. Escala et al. 2005; Berczik et al. 2006; Dotti et al. 2007; Mayer et al. 2007), but BH binary lifetimes in gas-poor or minor mergers are uncertain and may be substantially longer.

To account for this uncertainty, we impose a delay for *all* BH mergers between the merger time in the simulation and the merger time assumed in our analysis. The time delay is assumed to scale with the BH mass ratio and the host galaxy gas fraction. Specifically, we add a time delay of 0.1 Gyr/q to each merger, where q is the BH mass ratio, motivated by the $1/M_{BH}$ scaling of the Chandrasekhar (1943) dynamical friction time-scale for a given primary galaxy mass. Because gas-driven BH inspiral should be efficient only in the regime where the gas disc is massive compared to the BHs, we add an additional 0.5 Gyr delay if the merged host galaxy has a cold (star-forming) gas fraction of $f_{gas, sf} < 0.1$, where $f_{gas, sf} \equiv M_{gas, sf}/(M_* + M_{gas, sf})$. Our results are not sensitive to the exact choice of these parameters; in practice, the effect of the delay imposed on gas-poor mergers is subdominant to the q-dependent delay. With this prescription, 26 per cent of BH 'mergers' in Illustris are still unmerged at $z = 0$, but only 2.5 per cent of binaries with $\log q > -1.5$ have not coalesced by $z = 0$. This delayed merger time is used as the actual merger time in all subsequent analysis, and only mergers occurring at $z > 0$ are considered.

In addition to the simulation snapshots, data for each BH merger in Illustris, as well as gas accretion data during active BH phases, are output at much higher time resolution.[2] We note that a small fraction of these data files for Illustris-1 in the redshift range 0.15–0.38 were corrupted. This has only a modest effect on the statistics of our results, well within the uncertainties from model assumptions and resolution convergence. Except where otherwise noted, the results below refer to the Illustris-1 simulation. We discuss the convergence of the Illustris-1, 2, and 3 results in Appendix A.

We also note that in some snapshots, a small number of BHs are not associated with any subhalo. Detailed analysis indicates that this can occur if there is a mismatch between the time when a subhalo is determined to have merged with a larger halo or FOF group and the time when its central BH is repositioned on the new potential minimum. Such phases are short-lived and affect <0.7 per cent of BHs at any point in the simulation. We therefore ignore these BHs in our analysis, with negligible impact on our results.

2.2 BH spin models and kick velocities

The recoil velocity of a merged BH depends on the mass ratio and spin vectors of the progenitor BHs. We obtain the BH mass ratio just prior to merger directly from the simulation data. For the spins, we must assume a physically motivated distribution. As discussed above, BH spins will be preferentially aligned with the orbital angular momentum and with each other prior to merger if their inspiral is driven by torques from a circumbinary gas disc. If the BH inspiral is instead driven predominantly by stellar interactions, or if the accretion flow is chaotic, the spin orientations may be closer to random. Coherent gas accretion will also spin up the BHs on average, while repeated mergers at random orientations cause BH spin magnitudes to tend towards an approximately thermal distribution. The efficiency of each of these processes in various merger environments is not known, but by considering a range of models for BH spins, we can explore their relative effects on the resulting recoil kicks and on the observability of offset AGN.

The most optimistic assumption for BH spins – yielding the largest kicks – is that the BHs are always rapidly spinning and that

---

[2] These data will be made publicly available in the coming months at the permanent site www.illustris-project.org/data, where the other simulation data are already available (Nelson et al. 2015).
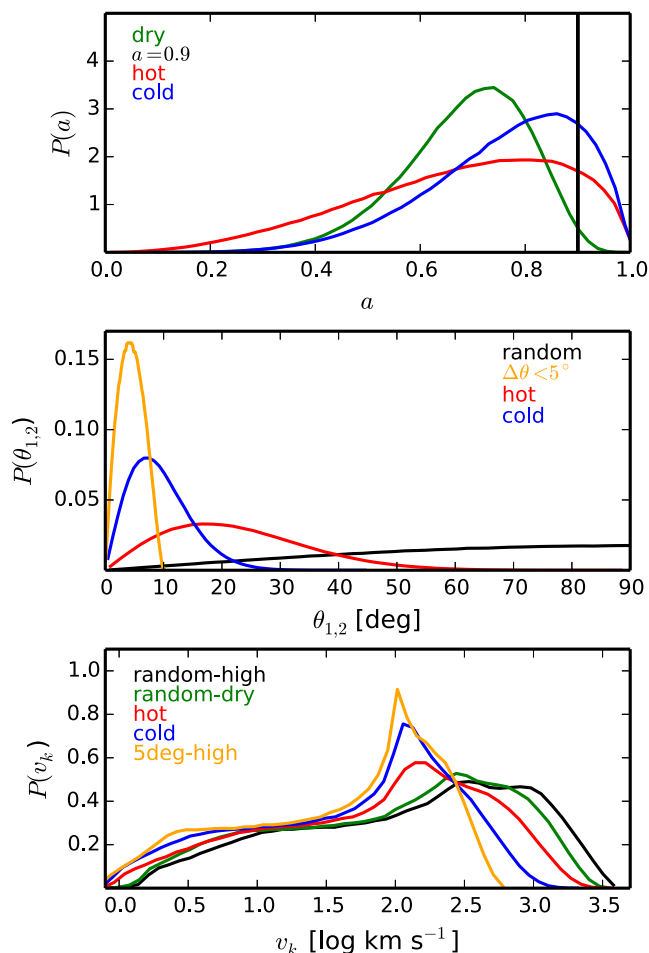
Figure 1. Normalized distributions of spin magnitude, misalignment angle, and recoil kick velocity. Top panel: spin magnitude distributions for the `hot` (red), `cold` (blue), and 'dry merger' (green) models, as defined in the text; for comparison, the 'high' spin value of $a = 0.9$ (black) is also shown. Middle panel: distributions of the angle of misalignment between the BH spin vectors prior to merger, $\theta_{1,2}$, for the models used. The black curve is a distribution uniform in $\cos(\theta_{1,2})$; for clarity, we map $\theta \rightarrow 180° - \theta$ for $\theta_{1,2} > 90°$. The red and blue curves again denote the `hot` and `cold` distributions, respectively. For the orange curve, the misalignment between each BH spin and the orbital angular momentum ($\Delta\theta$) is constrained to be $<5°$; the distribution is uniform in $\cos(\theta)$ over this range. Bottom panel: recoil kick distributions resulting from equation (1) for the spin models used, assuming that the mass ratio q has a log-uniform distribution in the range $-2 < \log(q) < 0$. The `random-high` model uses $a = 0.9$ and random orientations, the `random-dry` model uses the 'dry' spin magnitude distribution and random orientations, the `hot` and `cold` models correspond to partially and nearly aligned spins, and the `5deg` model assumes $a = 0.9$ and $\Delta\theta < 5°$. The high-velocity tail of kicks is strongly suppressed when spin alignment is efficient.

their spins are randomly oriented prior to merger. We refer to this (using a dimensionless spin parameter $a = 0.9$) as the `random-high` model. We also consider a spin distribution resulting from repeated randomly oriented mergers, which could arise from a succession of dry mergers. Lousto et al. (2012) fit this spin magnitude distribution with a beta function that peaks at $a \sim 0.7$ and has a large tail towards low spins (Fig. 1). This is denoted as the `random-dry` model.

At the other extreme, we may assume that BH spins always undergo some amount of alignment prior to merger. We use the `hot`

and `cold` spin evolution models as described in Dotti et al. (2010) and Lousto et al. (2012), which result in partial alignment (within $\sim 30°$) or near alignment ($\lesssim 10°$) of BH spins, respectively. We adopt the analytic approximations to the spin angle and magnitude distributions given in Lousto et al. (2012), which are based on the simulations of Dotti et al. (2010). Finally, we consider the case in which spin alignment is always very efficient, such that spins are aligned to within $5°$ of the binary orbital angular momentum. The misalignment angle of each BH is chosen at random in $\cos(\theta)$ over the interval $(0°, 5°)$, and the spin magnitude is assumed to be high ($a = 0.9$). We refer to this as the `5deg` model.

Of course, it is likely that reality lies somewhere between these extremes, with efficient spin alignment occurring in some BH mergers and not in others. Recent studies of BH spin evolution using semi-analytic models have considered a mixture of coherent and chaotic accretion (Barausse 2012; Sesana et al. 2014). Sesana et al. (2014) find that empirical constraints on BH spins are better fitted by this type of intermediate scenario than if accretion is assumed to be always coherent or always chaotic. As binary BH spin alignment may be driven by circumbinary gas discs, we construct hybrid spin models by utilizing information about the cold gas content of the progenitor host galaxies in Illustris. In general, circumnuclear discs with $M_{disc} \gtrsim M_{BH}$ can significantly influence the inspiralling BHs. We do not attempt to estimate the fraction of gas in the progenitor galaxies that will end up in a circumnuclear disc, but rather assume that gas-rich galaxy mergers are likely to trigger the formation of massive circumnuclear discs. Thus, we define galaxies as 'gas-rich' if they have a cold (star-forming) gas fraction of > 10 per cent, where again the gas fraction $f_{gas, sf}$ is defined relative to the total stellar mass of the progenitor subhaloes. With this definition, the large majority of galaxy mergers are classified as gas-rich – more than 85 per cent of those at $z < 1$ and nearly all of those at higher redshift. We note that the gas-rich fraction could be over-represented in Illustris, as feedback processes do not sufficiently suppress star formation at the high- and low-mass end of the halo mass function (e.g. Genel et al. 2014; Vogelsberger et al. 2014a; Snyder et al. 2015). However, the stellar mass function is otherwise in reasonable agreement with observations, as are the molecular gas fractions and the star formation rate density over cosmic time (Genel et al. 2014; Vogelsberger et al. 2014a). Moreover, we find that a lower gas-rich merger fraction would only *increase* the observability of recoiling AGN. Alternate definitions of gas-rich mergers are discussed in detail in Section 3.7, but our results are not very sensitive to the exact definition used.

We define one hybrid model in which BHs in gas-rich mergers have nearly aligned spins (drawn from the `cold` distribution), and BHs in gas-poor mergers have spins drawn from the `random-high` model. We denote this spin model as `fgas-a`. We also consider a more conservative model in which the same critical $f_{gas, sf}$ determines whether spins are drawn from the `random-dry` model (for gas-poor mergers) or the `5deg` model (for gas-rich mergers); this hybrid model is denoted as `fgas-b`.

Throughout this paper, we refer to these three classes of spin models as 'random', 'hybrid', and 'aligned'. Table 1 summarizes the definitions of each spin model. We finally note that the inclusion of relativistic precession effects on BH spins is beyond the scope of this work, but that recently developed efficient techniques for calculating spin precession create a promising avenue for future studies (Gerosa et al. 2015; Kesden et al. 2015).

Once the pre-merger BH spins have been obtained from the distribution for a given spin model, the recoil kick velocity can be calculated using a fitting formula based on results from

Table 1. BH spin models used to determine the recoil kick distributions. We consider three classes of spin models: *random*, *hybrid*, and *aligned*. Columns: (1) name of model, (2) spin magnitude distribution for gas-rich mergers, (3) distribution of spin misalignment angle (relative to the orbital angular momentum) for gas-rich mergers, (4) spin magnitude distribution for gas-poor mergers, (5) spin misalignment angle distribution for gas-poor mergers, (6) critical fraction of star-forming gas used to define gas-rich versus gas-poor mergers ($f_{gas,sf} \equiv M_{gas,sf}/(M_{gas,sf} + M_*)$). The spin angle and magnitude distributions denoted as 'dry', 'cold', and 'hot' are taken from Lousto et al. (2012) and shown in Fig. 1. Otherwise, the spin angle distributions are randomized in $\cos(\theta)$ over the interval shown.

| Name | Gas-rich | | Gas-poor | | $f_{gas,\,sf}$ |
	$|\boldsymbol{a}_{1,2}|$	$\cos(\theta_{1,2})$	$|\boldsymbol{a}_{1,2}|$	$\cos(\theta_{1,2})$	
Random					
random-high	0.9	$(-1,1)$	0.9	$(-1,1)$	–
random-dry	Dry	$(-1,1)$	Dry	$(-1,1)$	–
Hybrid					
fgas-a	Cold	Cold	0.9	$(-1,1)$	0.1
fgas-b	0.9	$(0.996,1)$	Dry	$(-1,1)$	0.1
Aligned					
hot	Hot	Hot	Hot	Hot	–
cold	Cold	Cold	Cold	Cold	–
5deg	0.9	$(0.996,1)$	0.9	$(0.996,1)$	–

numerical relativity simulations. We use the following formula given by Lousto et al. (2012):

$$\boldsymbol{v}_{recoil} = v_m \hat{\boldsymbol{e}}_{\perp,1} + v_\perp (\cos\xi\,\hat{\boldsymbol{e}}_{\perp,1} + \sin\xi\,\hat{\boldsymbol{e}}_{\perp,2}) + v_\parallel \hat{\boldsymbol{e}}_\parallel, \quad (1)$$

$$v_m = A\eta^2 \sqrt{1-4\eta}\,(1+B\eta), \quad (2)$$

$$v_\perp = \frac{H\eta^2}{(1+q)}(a_{2\parallel} - qa_{1\parallel}), \quad (3)$$

$$v_\parallel = \frac{16\eta^2}{(1+q)} \left[V_{1,1} + V_A \tilde{S}_\parallel + V_B \tilde{S}_\parallel^2 + V_C \tilde{S}_\parallel^3 \right]$$
$$\times |\boldsymbol{a}_{2\perp} - q\boldsymbol{a}_{1\perp}|\cos(\phi_\Delta - \phi_1), \quad (4)$$

where $\eta \equiv q/(1+q)^2$ is the symmetric mass ratio, $\perp$ and $\parallel$ refer to vector components perpendicular and parallel to the orbital angular momentum, respectively, and $\hat{\boldsymbol{e}}_{\perp,1}$ and $\hat{\boldsymbol{e}}_{\perp,2}$ are orthogonal unit vectors in the orbital plane. The vector $\tilde{\boldsymbol{S}} \equiv 2(\boldsymbol{a}_2 + q^2\boldsymbol{a}_1)/(1+q)^2$, and ϕ_Δ is the angle between the in-plane component $\boldsymbol{\Delta}_\perp$ of the vector $\boldsymbol{\Delta} \equiv M^2(\boldsymbol{a}_2 - q\boldsymbol{a}_1)/(1+q)$ and the infall direction at merger. The phase angle ϕ_1 depends on the initial conditions of the binary and is assumed to be random. The best-fitting values of $A = 1.2 \times 10^4$ km s^{-1}, $B = -0.93$, $H = 6.9 \times 10^3$ km s^{-1}, and $\xi = 145°$ are taken from González et al. (2007) and Lousto & Zlochower (2008), and the coefficients $V_{1,1} = 3677.76$, $V_A = 2481.21$, $V_B = 1792.45$, and $V_C = 1506.52$ (all in km s^{-1}) are defined in Lousto et al. (2012).

The kick distributions resulting from each spin model are shown in Fig. 1 assuming a log-uniform mass ratio distribution in the range $-2 < \log(q) < 0$. It is clear that the high-velocity tail is greatly suppressed when spin alignment is efficient.

2.3 Host galaxy model

To model the merged host galaxy, we first match the BHs in each merging pair to their progenitor subhaloes using data from the closest prior simulation snapshot. While the simulations provide de-

tailed information about the density profiles of the progenitor host galaxies of merging BHs, the structure of the merger remnant at the time of the BH merger (which generally occurs between snapshots) is not known. Major mergers in particular should produce remnant galaxies with drastically different structure than their progenitors, and these systems are most likely to produce observable recoils. Thus, we opt to use the mass and stellar half-mass radii of the progenitors to construct an analytic model for the gravitational potential of the merger remnant.

In many cases, the BHs already inhabit a common subhalo in the snapshot prior to their merger, and we use the (DM, stellar, and gas) mass of this common subhalo as the total mass of the merged host galaxy. If the BHs inhabit unique subhaloes, the sums of the DM, stellar, and gas masses of each progenitor are used for the total mass in the merger remnant. We note that Rodriguez-Gomez et al. (2015) have found that the mass ratio of two merging subhaloes can vary dramatically between the time of first infall and the time of halo merger, owing to the difficulty in assigning particles between two nearby haloes. However, the *combined* mass of the merging haloes is generally robust, and this is the quantity used in our calculations. This provides further motivation for our choice to construct analytic merger remnant potential models rather than extracting density profiles directly from the simulation. Again, we consider only BH mergers for which each progenitor has at least 300 DM and 80 stellar particles.

The DM component is modelled as a Hernquist (1990) potential matched to an equivalent NFW profile (Navarro, Frenk & White 1997) with concentration parameter $c_{vir} = 10.5(M_{vir}/10^{12}\,h^{-1}\,\mathrm{M}_\odot)^{-0.11}(1+z)^{-1}$ (following Bullock et al. 2001; Macciò et al. 2007). The stellar component is modelled as a spherical bulge with a softened isothermal density profile, $\rho_* = \sigma_*^2/(2\pi G(r^2 + r_{soft}^2)$, with velocity dispersion $\sigma_*^2 = GM_*/R_{bulge}$ and $r_{soft} = r_{infl}$, where the influence radius r_{infl} of the BH is GM_{BH}/σ_*^2. The profile is truncated at an inner radius R_{ej} within which matter remains bound to the recoiling BH. This truncated, softened profile accounts for the removal of the cusp of stars that is carried along with the BH and prevents unphysically large central densities.

We truncate the stellar bulge at an outer radius R_{bulge} as well; this radius is scaled to the stellar half-mass radii of the progenitor galaxies. Specifically, we take the maximum radius of the stellar bulge in the merger remnant to be a simple average of the progenitor half-mass radii; this choice is based on the assumption that the inner density profile is steep during a major merger, and is also meant to account for the fact that the stellar half-light radii in Illustris are up to a factor of 2 larger than in observed galaxies at $z = 0$, at least at the low-mass end (Snyder et al. 2015). To additionally ensure that the stellar bulges do not have artificially low densities in low-mass galaxies, we compare the calculated stellar velocity dispersion to that inferred from the $M_{BH}-\sigma_*$ relation of McConnell & Ma (2013). If the calculated σ_* derived from M_* and R_{bulge} falls more than 3σ below the empirical relation (assuming 0.4 dex scatter in $\log M_{BH}$), we set σ_* to lie on the relation and calculate R_{bulge} from this instead of using the value inferred from the simulation. In practice, this criterion is imposed on most low-mass bulges ($M_* \lesssim 1$–$2 \times 10^{10}\,\mathrm{M}_\odot$), affecting 14 per cent of all merger remnants. Finally, an absolute maximum bulge radius of 15 kpc is also imposed; this affects about 1 per cent of all merger hosts. The assumption of compact stellar bulges likely overestimates the inner stellar density in some cases, particularly in dry mergers between elliptical galaxies, when the BH inspiral time-scale is very long, or if a disc is able to re-form after the merger. However, we consider this to be a conservative choice, as the steeper inner profile ensures

that we do not overestimate the fraction of recoiling BHs that can escape the bulge.

We also include a gas disc in our galaxy models; hydrodynamic simulations of recoiling BHs in isolated galaxy mergers indicate that the dense, circumnuclear gas discs that form during major gas-rich mergers can strongly suppress recoil trajectories (Blecha et al. 2011; Sijacki et al. 2011). We assume that the cold, star-forming gas in the progenitor haloes has condensed into a circumnuclear gas disc by the time of the merger, and this gas is modelled with a $1/r$ Mestel surface density profile. This is a reasonable approximation to the surface density profiles obtained from zoom-in simulations of gas-rich mergers by Hopkins & Quataert (2010), spanning radii of ~ 0.1 pc to 1 kpc. Motivated by these results, we assume that the inner disc surface density scales with the star-forming gas fraction as $\log \Sigma_{0.1} = 2 \log (f_{\rm gas,\, sf}/0.1) + 12$, where $\Sigma_{0.1}$ is the surface density at 0.1 pc in $M_\odot$ kpc^{-2}. The disc component is ignored in galaxies with $\log f_{\rm gas,\, sf} < -1.5$. Typical disc masses are $\log (M_{\rm disc}/M_\odot)$ ~ 9.5–10.5 at $z < 1$ and $\log (M_{\rm disc}/M_\odot) \sim 10$–11 at $z > 1$. The disc is truncated at an inner radius $R_{\rm ej}$ within which gas is bound to the recoiling BH, and at an outer radius $R_{\rm max} = G M_{\rm disc}/v_0^2$, where the fiducial circular velocity v_0 of the Mestel disc is set by the above scaling of $\Sigma_{0.1}$. An absolute maximum disc radius of 15 kpc is also imposed; this affects about 6 per cent of merger hosts.

2.4 Recoil trajectories

Once the merged galaxy model is assigned, the recoiling BH trajectory is integrated in this potential for each recoil event with $v_{\rm k}/v_{\rm esc} > 0.1$. All spatial and velocity offsets are calculated relative to the centre of the static host potential. We include stellar dynamical friction via the Chandrasekhar formula (Chandrasekhar 1943), taking $M_{\rm BH} + M_{\rm disc,\, ej}$ as the total mass of the recoiling object. Gas dynamical friction is neglected; because the disc is assumed to be thin ($h/r \lesssim 0.1$ for gas-rich systems), only a few per cent of recoiling BHs will be kicked directly into the disc if their orientations are random. If BH spin alignment is efficient and the angular momenta of the accretion disc and the large-scale disc are also aligned, there may be a preference for kicks into the disc plane, but in this case the kicks will be small and difficult to observe. For large kicks, the v^{-2} scaling of dynamical friction ensures that it will be relatively inefficient as the BH leaves the centre of the galaxy. The main uncertainty introduced by neglecting gas dynamical friction is its larger possible effect on recoil events with moderate $v_{\rm k}/v_{\rm esc}$ that return to the centre on short time-scales, in which case the damping of their motion may be underestimated in our model. However, we demonstrate below that these objects do not dominate the population of observable recoils.

The integration is stopped when one of the following occurs: (a) the velocity and galactocentric distance of the BH are less than critical values $v < 0.2\sigma_*$ and $R < R_{\rm infl}$, (b) the AGN luminosity falls below the observable limit and $v_{\rm k} > v_{\rm esc}$, or (c) the integration reaches a Hubble time.

Because the GW recoil kicks are implemented in post-processing, we must explicitly remove from our analysis mergers involving a BH that either (a) was previously ejected from its host galaxy or (b) experienced a previous recoil kick and has not yet returned to the galactic centre. For the fiducial galaxy potential model, this removes 5 per cent of BH mergers that meet all other criteria, if the `random-high` spin model is assumed. 2 per cent of mergers are excluded in the `random-dry` model, and <1 per cent are excluded in all other spin models, where large kicks are less common.

Our final sample therefore includes BH mergers that meet the following criteria: (i) each BH has $M_{\rm BH} \geq 10^6$ $M_\odot$, (ii) each BH is hosted in a progenitor with ≥ 300 DM (80 stellar) particles, (iii) the 'delayed' BH merger time (defined above) occurs at $z > 0$, and (iv) the merger occurs at a sufficiently high redshift to appear in the past light cone of an observer at $z = 0$ situated in the centre of the simulation box. (Because we are concerned only with the statistics of observable recoils over the entire sky, this is relevant only at extremely low redshifts; position in the box is a negligible consideration at cosmological redshifts.) For Illustris-1, this yields a sample of 8993 BH mergers. Finally, we ensure that (v) neither BH is still displaced from a previous recoil event at the time of merger, which primarily affects the random spin models. The final sample sizes for the `random-high` and `random-dry` spin models are 8576 and 8800, respectively.

2.5 AGN luminosities

Blecha & Loeb (2008) and Blecha et al. (2011) have developed an analytic model for the mass and extent of the accretion disc carried along with a recoiling BH that is actively accreting at the time of the kick. The model assumes a thin 'α-disc' that becomes self-gravitating beyond the radius where the Toomre Q parameter (Toomre 1964) equals unity. The radius of the disc bound to the ejected BH is that at which the orbital speed $v_{\rm orb} \approx v_{\rm k}$. Typical ejected disc masses are a few per cent of the BH mass for observable offset AGN (for further details see Fig. 7).

The BH accretion rate at the time of the kick is taken directly from the simulation. The ejected, isolated accretion disc will then diffuse outwards over time, causing the accretion rate to monotonically decline $\propto t^{-19/16}$ (assuming Thomson scattering opacity; Cannizzo, Lee & Goodman 1990; Pringle 1991). Following Blecha et al. (2011), we use this to calculate the accretion rate at each integration timestep of the recoil trajectory. The accretion rate is converted into a bolometric AGN luminosity $L_{\rm bol} = \epsilon \dot{M} c^2$, where the radiative efficiency ϵ is assumed to be 0.1 at high Eddington ratios. At low Eddington ratios, the accretion flow is assumed to become radiatively inefficient. Following Narayan & McClintock (2008), ϵ is multiplied by an accretion rate-dependent factor $\dot{M}/(f_{\rm riaf} \dot{M}_{\rm Edd})$ when the Eddington ratio falls below a critical value $f_{\rm riaf}$. We adopt $f_{\rm riaf} = 0.05$, which is the critical Eddington ratio for the transition to 'radio mode' feedback in Illustris. Also, because very low Eddington ratios are unrealistic in the context of our thin α-disc accretion model, a minimum Eddington ratio must be imposed to prevent unfeasibly long offset AGN lifetimes. We conservatively adopt a minimum observable Eddington ratio of 10^{-2}. While lower Eddington ratios may be detectable for some massive BHs, this choice is motivated by the distribution of Eddington ratios in observed quasars (Shen & Kelly 2012). Section 3.7 examines how our results are affected if a lower minimum value of $f_{\rm Edd}$ is assumed.

Our procedure allows the observable offset AGN lifetime to be calculated for arbitrary telescope resolution and sensitivity. For each recoil event, a random viewing angle is assigned, and projected separations ($\Delta R_{\rm proj}$) of twice the angular resolution are defined as resolvable spatial offsets. For most of our analysis, the minimum resolvable LOS velocity offset is assumed to be $\Delta v_{\rm LOS} > 600$ km s^{-1}. However, in some spectra, especially those with very broad lines (full width at half-maximum, FWHM $\gg 1000$ km s^{-1}), offsets this small may not be measurable. Thus, we also consider a more conservative minimum $\Delta v_{\rm LOS} = 1000$ km s^{-1}.

The minimum observable $L_{\rm bol}$ for each event is determined by first defining a desired flux sensitivity and band, and then applying an

inverse bolometric correction to the corresponding luminosity at the redshift of the recoil event. We use the luminosity-dependent optical and hard X-ray bolometric corrections from Hopkins, Richards & Hernquist (2007). For the near-IR (NIR), we assume a constant bolometric correction of 11.0, derived from the mean SED of Richards et al. (2006) at 2 μm. AGN bolometric corrections are not well constrained, particularly at optical wavelengths where obscuration and host galaxy contamination have large effects. Even at X-ray wavelengths, the fraction of Compton-thick sources is uncertain, and bolometric corrections have been found to vary with Eddington ratio (Vasudevan & Fabian 2007; Lusso et al. 2012). We neglect obscuration effects here, as these should be less significant for *offset* AGN (though the early post-recoil stage of velocity offset AGN may be a notable exception, as discussed in Section 3.5).

Despite the inherent uncertainties, our results depend only weakly on the choice of bolometric correction. This is primarily because at low to moderate redshift, the minimum $f_{Edd} = 10^{-2}$ criterion is typically a more stringent constraint on the AGN lifetime than the absolute flux limit derived from the (reverse) bolometric correction. The consequences of this are discussed in detail in Section 3.3.

3 RESULTS

3.1 Merging BH population

The differential all-sky merger rate of BHs in Illustris at $z = 0$, $d^2N/d\log(z)\, dt$, is shown in Fig. 2. Two BHs are allowed to merge
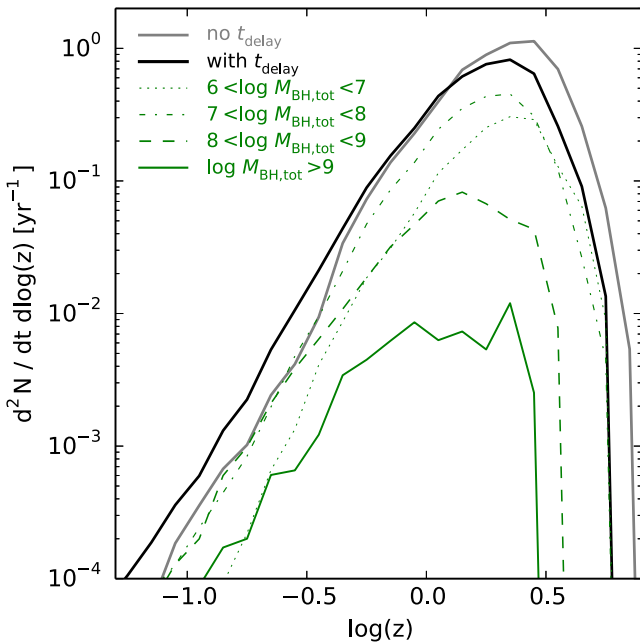

Figure 2. The BH merger rate in Illustris at $z = 0$, per unit redshift (in logarithmic bins). The thick grey curve denotes the 'raw' merger rate for all BH mergers that meet the selection criteria defined above, except that the actual BH merger redshift from the simulation is assumed instead of the 'delayed' redshift. The thick black curve shows the merger rate with a delay time-scale added to each merger that depends on BH mass ratio and host gas fraction, as described in the text. The thin green lines denote the merger rate for different remnant BH mass bins, as indicated in the plot legend. The raw merger rate peaks at $z \sim 2$, in good agreement with semi-analytic models, and the delayed merger rate peaks at $z \sim 1.5$. The cumulative merger rate over cosmic time also agrees well with semi-analytic models, but more low-redshift mergers occur in Illustris.

in the simulation when their separation is less than a smoothing length. In cosmological simulations, this spatial scale (~ 0.1–1 kpc in Illustris) is far larger than the regime in which BH inspiral is dominated by GW emission. Thus, as described above, we add a delay time-scale $t_{delay} = 0.1\,\text{Gyr}/q + t_{gas}$ to the time of each BH merger, where $t_{gas} = 0.5$ Gyr if the fraction of mass in cold (star-forming) gas $f_{gas,\,sf} < 0.1$ and zero otherwise. More detailed modelling of binary BH inspiral on sub-resolution scales is beyond the scope of this work, and will be examined in an upcoming study (Kelley et al., in preparation).

The green lines in Fig. 2 show the merger rate in different mass bins, for the remnant BH mass $M_{tot} = M_1 + M_2$. Over a large redshift range $0.5 < z < 3$, the merger rate is dominated by BHs in the mass range 10^7–10^8 M$_\odot$. Lower mass BHs dominate at higher redshifts, and higher mass BHs dominate at lower redshifts, consistent with expectations from hierarchical growth. Low-mass BH mergers ($\lesssim 10^6$ M$_\odot$) are of interest as GW sources for a future space-based interferometer such as *eLISA*, but these are excluded from our analysis owing to resolution limits, as discussed above.

The cumulative merger rate of this BH sample at $z = 0$, integrated over cosmic time, is 0.42 yr^{-1}, and the merger rate for BHs with total mass 10^7–10^9 M$_\odot$ is 0.27 yr^{-1}. This is in reasonable agreement with results from semi-analytic modelling (e.g. Sesana et al. 2004; Barausse 2012), especially considering the very different BH seeding and growth prescriptions in the semi-analytic models versus Illustris. The BH merger rate also peaks at similar redshifts ($z \gtrsim 2$) in both Illustris and the semi-analytic models, though after the merger delay time-scale is implemented, the peak shifts to slightly lower redshift ($z \sim 1.5$) in our sample. However, there are more high-redshift mergers in the semi-analytic models, owing to the BH seed formation prescriptions, which form seeds at higher redshifts than in Illustris. The first BHs in Illustris form at $z \sim 10$–11, and the first mergers occur at $z < 7$. The merger rate at low redshift ($z < 1$) is correspondingly higher in Illustris than in the models of Sesana et al. (2004, for BH binaries with total mass $>10^6$ M$_\odot$). In the 'heavy seed' model of Barausse (2012), the cumulative merger rate (for $M_{tot} > 10^6$ M$_\odot$) is about a factor of 3 higher than in Illustris, such that the merger rate per unit redshift is slightly higher than the Illustris rate at all redshifts. Further comparison with comparable cuts on host halo masses improves the agreement with both the Barausse (2012) and Sesana et al. (2004) models, though some differences in the mass and redshift distribution of the merger rates remain, as does the overall trend towards more high-redshift mergers in the semi-analytic models (Barausse & Sesana, private communication). Empirical constraints cannot yet distinguish between these differences in the details of BH formation, growth, and merger histories. These distinctions between cosmological simulations and semi-analytic models are important to bear in mind, and they contribute to the differences in offset AGN rates predicted here and in VM08 (Section 4.2). However, the level of agreement of the overall merger rates in Illustris versus semi-analytic models is encouraging. We finally note that the *galaxy* merger rates in Illustris are in reasonable agreement with observations (Rodriguez-Gomez et al. 2015), though they disagree qualitatively with rates derived from semi-analytic models based on cosmological *N*-body simulations (Guo & White 2008).

The most massive BH mergers ($M_{tot} \gtrsim 10^9$ M$_\odot$) will create luminous GW signals that could be detected with PTAs; we find a cumulative merger rate of 0.006 yr^{-1} for these massive mergers. This is consistent with predictions from semi-analytic models that individual GW events resolvable with PTAs should be rare (e.g. Sesana, Vecchio & Volonteri 2009). However, the stochastic
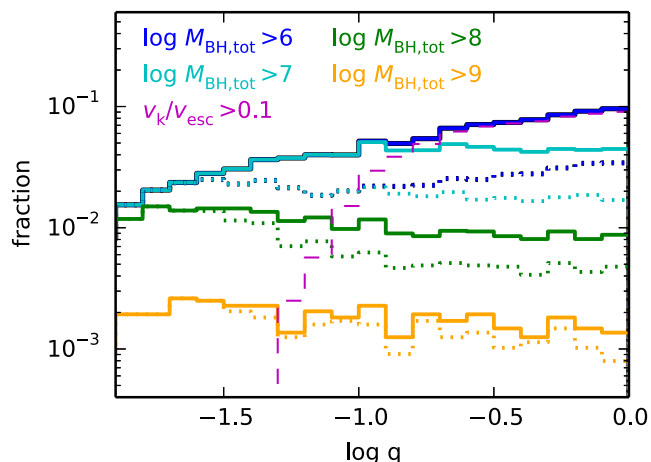
Figure 3. The merging BH mass ratio distribution in Illustris, separated into total BH mass bins. As indicated in the plot legend, the blue curve shows the distribution for all BH masses (above the minimum 2×10^6 M$_\odot$), and the cyan, green, and orange curves correspond to remnant BH masses of $\log (M_{BH}/M_\odot) > 7$, 8, and 9, respectively. For most BH masses, the q distribution is nearly flat in log space over the range $-2 < \log q < 0$. The dotted lines show the distributions for only BH mergers occurring at $z < 1$. The thin magenta line denotes mergers that yield a recoil velocity >0.1 v_{esc}; no mergers with $\log q \lesssim -1.3$ produce significant recoils.

background from unresolved massive BH mergers may be detectable with PTAs in the coming years. The predicted GW signal from BH mergers in Illustris will be explored in detail in Kelley et al. (in preparation).

Fig. 3 shows the distribution of merging BH mass ratios q, separated by total BH binary mass. For a wide range of BH masses, the $\log q$ distribution is nearly flat over the range $-2 < \log q < 0$. Low mass ratios are favoured slightly for total BH masses $>10^8$ M$_\odot$; these massive BHs are fewer in number, so they more often undergo unequal-mass mergers. The trend reverses when mergers with total mass $M_{BH} = 10^7$–10^8 M$_\odot$ are included, and when all BH masses $>10^6$ M$_\odot$ are considered, the distribution clearly rises towards $\log q = 0$. Recall that a lower mass limit of 10^6 M$_\odot$ has been imposed for individual merging BHs, such that BH binaries with total mass $<10^7$ M$_\odot$ cannot have $\log q < -1$.

The dotted lines indicate the q distribution for only mergers occurring at $z < 1$. We see that major mergers are more common at high redshift. This is partly a result of hierarchical growth; minor mergers become more frequent as larger galaxies (and BHs) form. It is also a product of the long dynamical friction time-scales for minor mergers, reflected in the q-dependent delay time-scales we have imposed on the simulation merger time. The limited box size of Illustris, $(106.5 \text{ Mpc})^3$, could further contribute to the limited number of major mergers between low-redshift, massive BHs. For mergers at $z < 1$, the distribution is nearly flat or declining with $\log q$ for all BH masses.

Also shown in this figure is the subset of BH mergers that produce a recoil kick with $v_k/v_{esc} > 0.1$ (in the `random-dry` spin model), which is the minimum kick velocity for which recoil trajectories are calculated. Only mergers with $\log q \gtrsim -1.5$ produce such recoils, with most requiring $\log q > -1$.

We note that the merging BH mass ratio distribution in Illustris is qualitatively different from that in the models of VM08, which exhibits a sharp cutoff at $q \gtrsim 0.3$. Because the recoil kick distribution depends sensitively on BH mass ratios, this is an important distinction. While there are uncertainties in the BH merger time-scales

in Illustris, as discussed above, they are largest for *minor* mergers. A more detailed comparison of our results with those of VM08 is given in Section 4.2. We additionally note that recent numerical simulations of BH binaries in circumnuclear gas discs suggest that accretion on to the lower mass BH may be more rapid, driving the final mass ratios closer to unity than the progenitor distribution (Roedig et al. 2012; Farris et al. 2014).

3.2 Recoiling BH dynamics

Fig. 4 shows the distribution of kick velocities scaled to the host escape speed, for a random spin model (`random-dry`), a hybrid model (`fgas-b`), and an aligned model (`5deg`). For the random spin models, the distribution peaks at $v_k/v_{esc} \sim 0.3$, unlike the hybrid and aligned models, for which the distribution is heavily skewed towards low v_k/v_{esc}. In the `5deg` model, recoiling BHs almost never escape the galaxy entirely. (No recoils have $v_k/v_{esc} > 1$ in the `5deg` model, but this is limited by mass resolution and by statistics. The largest kicks produced in the `5deg` model could escape the lowest mass galaxies in our sample, and would easily escape dwarf galaxies unresolved in the simulation.) Escaping BHs are also very rare in the `cold` spin model, comprising <0.05 per cent of all kicks. The escape fractions in the hybrid spin models are 0.4 and 0.2 per cent for `fgas-a` and `fgas-b`, respectively, comparable to the `hot` spin model (0.6 per cent escaping). In contrast, the random spin models eject a full 6 per cent (`random-dry`) and 12 per cent (`random-high`) of all merged BHs.

The coloured lines in Fig. 4 show the distribution of v_k/v_{esc} separated by total merging BH mass. Because BH mass correlates with galaxy mass and v_k depends only on the BH mass *ratio*, large v_k/v_{esc} events are rarer for high-mass BHs. Despite this, in the random *and* the hybrid spin models, even the most massive BHs ($>10^9$ M$_\odot$) occasionally escape their hosts. Thus, if superkicks do occur, even massive hosts may be left without a central BH.

For bound recoiling BHs ($v_k/v_{esc} < 1$), trajectories can be characterized by the maximum apocentric distance from the host (R_{max}) and the time spent off-centre before returning to the galactic nucleus (t_{return}). These quantities are plotted versus v_k/v_{esc} for selected spin models in Fig. 5. Naturally, both R_{max} and t_{return} increase with v_k/v_{esc}, but the scatter is large, owing to the large dynamic range in halo, stellar, and gas model parameters. There is little to no correlation below $v_k/v_{esc} \sim 0.5$, where recoil trajectories are strongly suppressed by the dense stellar and gas components. These low-velocity recoiling BHs are generally confined to the central kpc of the galaxy, with return time $<10^7$ yr, and some are displaced by less than a parsec. In contrast, large kicks with $v_k/v_{esc} \gtrsim 0.8$ can displace the BH for >1 Gyr, with $R_{max} \gtrsim 100$ kpc.

Pre-merger spin alignment greatly suppresses high-velocity recoils. The hybrid `fgas-b` spin model produces a much smaller fraction of recoils with $v_k/v_{esc} > 0.6$ than the random spin model, and the `5deg` model yields no kicks above $v_k/v_{esc} = 0.8$. In the latter model, with extremely efficient spin alignment, recoiling BHs are nearly always confined to the central kpc of the galaxy (though again, this does not apply to dwarf galaxies that are below the resolution limit of the simulation).

3.3 Merging and recoiling AGN properties

The AGN Eddington ratios ($f_{Edd} \equiv L_{AGN}/L_{Edd}$) in Illustris evolve strongly towards lower values with cosmic time (Sijacki et al. 2015, fig. 9). At $z = 4$, when BHs have low masses and their hosts are generally gas-rich, the f_{Edd} distribution peaks near unity. At $z = 0$,
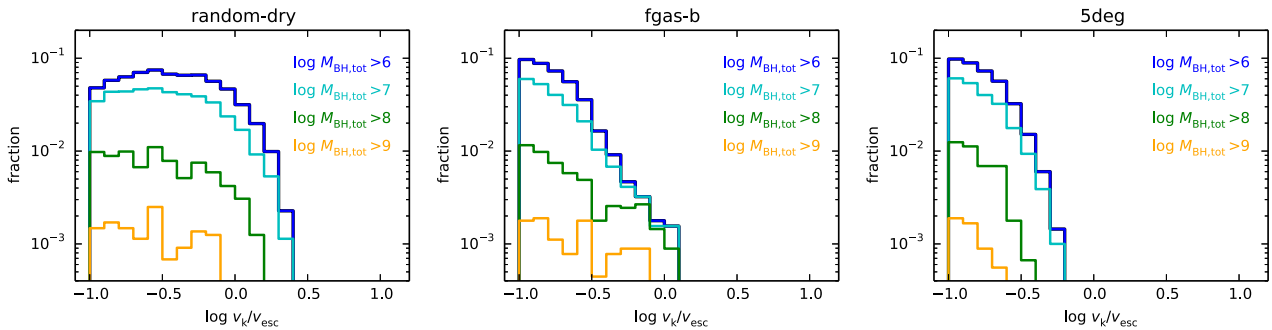
Figure 4. The distribution of kick velocities scaled to the host escape speed, v_k/v_{esc}, is shown for a random spin model (`random-dry`, left-hand panel), a hybrid model (`fgas-b`, middle panel), and an aligned model (`5deg`, right-hand panel). The distributions are separated by total merging BH mass using the same colour scheme as in Fig. 3, and as indicated in the plot legend. The random model has a large tail of escaping BHs, the hybrid model has only a few, and the aligned model has none.
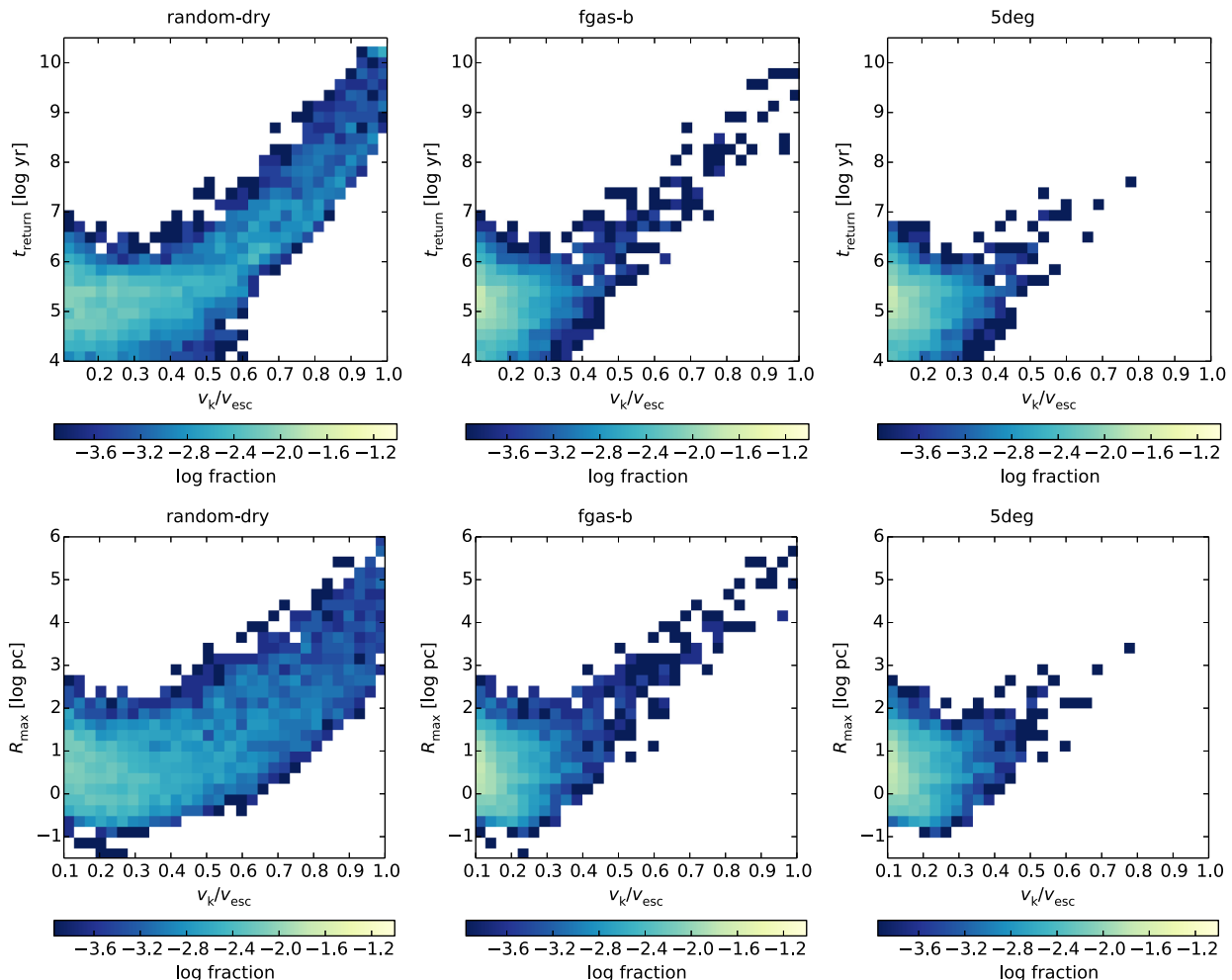

Figure 5. The characteristics of bound recoil trajectories ($v_k/v_{esc} < 1$) are shown for three different spin models: `random-dry` (top panels), `fgas-b` (middle panels), and `5deg` (bottom panels). In each case, the left-hand panel shows the distribution of return times t_{return} for the recoiling BH to settle back to the galactic centre, as a function of v_k/v_{esc}. The right-hand panel shows the maximum galactocentric distance of the BH orbit, R_{max}, versus v_k/v_{esc}. There is little dependence of either quantity on v_k/v_{esc} for values $\lesssim 0.5$, but marginally bound recoils can wander in the halo for many Gyr. Such events are much rarer in the hybrid model than the random model, and almost never occur in the aligned spin model.

Eddington-limited AGN are very rare, and the median f_{Edd} for all BHs is 3.7×10^{-4}.

The Eddington ratios for *merging* BHs are quite different. We find that the f_{Edd} distribution is heavily skewed towards unity for merging BHs at all redshifts. Fig. 6 (top panel) shows the f_{Edd} distribution in several redshift bins, for all merging BHs that meet our minimum mass criteria. At low redshifts, a tail extends to very low Eddington ratios ($f_{Edd} < 10^{-5}$), but more than half of all merging BHs have accretion rates capped at the Eddington limit at the time of merger. The weak evolution of f_{Edd} with redshift relative to the
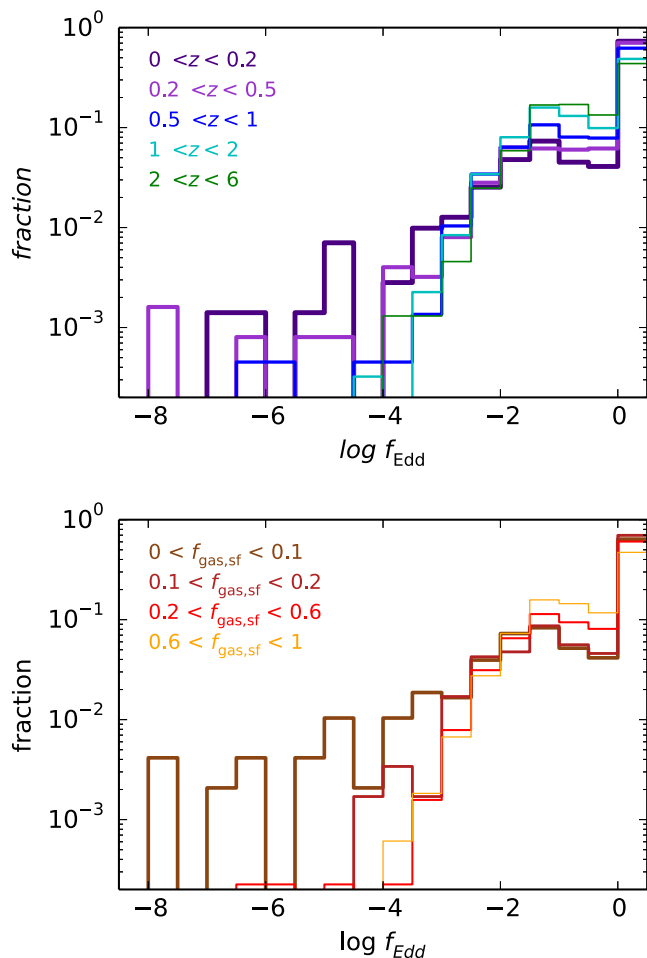
Figure 7. The mass distribution of the accretion disc bound to the recoiling BH, scaled to BH mass. The thick black line shows the distribution for all mergers, while the thin lines denote the distributions for only recoil events that produce observable spatially offset (solid blue line) and velocity offset (dashed red) AGN. For the offset AGN, the `random-dry` spin model and *HST*-COSMOS sensitivity (AB($F814W$) = 27.2 mag for point sources; Koekemoer et al. 2007), resolution (0.1 arcsec), and $\Delta v_{\rm LOS} > 600$ km s^{-1} are assumed, and a minimum offset lifetime of 10^5 yr is imposed. Typical disc masses are a few per cent of the BH mass.

Figure 6. Top panel: the distribution of Eddington ratios ($f_{\rm Edd} \equiv L_{\rm AGN}/L_{\rm Edd}$) is shown for merging BHs in Illustris, separated by redshift bin according to the plot legend (thicker lines correspond to lower z). Bottom panel: same distribution, but separated by the mass fraction (relative to stars) of cold, star-forming gas in the host (thicker lines correspond to lower $f_{\rm gas, sf}$). Unlike the $f_{\rm Edd}$ distribution for *all* BHs (cf. Sijacki et al. 2015), the distribution for merging BHs heavily favours values near unity at all redshifts and for a wide range in gas fractions.

overall BH population suggests that merger dynamics, rather than cosmic epoch, determine the fuelling rate of these AGN.

The bottom panel of Fig. 6 shows the $f_{\rm Edd}$ distribution separated by the star-forming gas fraction in the host, illustrating that higher-$f_{\rm gas}$ galaxies have higher Eddington ratios on average, but even mergers with $f_{\rm gas, sf} < 0.1$ have an $f_{\rm Edd}$ distribution skewed towards unity. This indicates that the cold gas in these galaxies is efficiently funnelled to the central regions during major mergers. The ability of major mergers to trigger AGN fuelling is well studied in simulations of isolated galaxy mergers (e.g. Wyithe & Loeb 2003; Di Matteo et al. 2005; Hopkins et al. 2006, 2008a), and a clear correlation between merging activity and AGN has been observed in real systems (e.g. Sanders et al. 1988; Urrutia et al. 2008; Koss et al. 2010; Ellison et al. 2011; Treister et al. 2012; Satyapal et al. 2014). Our results demonstrate that this effect is self-consistently produced in the Illustris cosmological simulations, and that a strong correlation exists during the late stages of the galaxy merger. In the cases where a significant delay occurs between the BH merger time in the simulation and the actual BH merger on subgrid scales, the accretion rate may be lower at the time of the merger and recoil, but
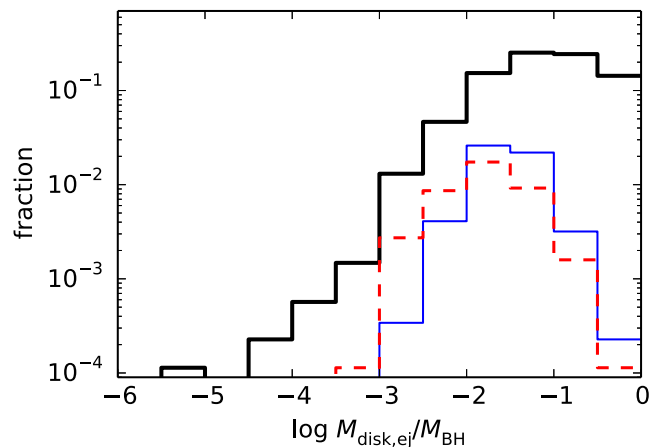
the longest delays occur for minor mergers that are less likely to produce observable recoils. Once the BH is ejected from the galaxy, however, its accretion rate declines with time, so the recoiling AGN have $L \sim L_{\rm Edd}$ only briefly.

Note that even though the recoiling BH accretion rate in our models declines monotonically after the recoil kick, the total AGN lifetime is generally *longer* than if a constant accretion rate (and corresponding AGN lifetime $M_{\rm disc}/\dot{M}_{\rm BH}$) is assumed. For the minimum observable value of $f_{\rm Edd} = 10^{-2}$ assumed in our models, the AGN lifetime may be nearly 10 times longer than a constant $M_{\rm disc}/\dot{M}_{\rm BH}$ lifetime.

As described in Section 2.3, recoiling BHs should carry along material that orbits the BH with $v_{\rm orb} \gtrsim v_{\rm k}$. The amount of mass bound to the recoiling BH, along with the accretion rate at the time of the kick, is used to determine its AGN lifetime. We calculate the mass and radius of this bound material assuming that it is composed of an α-disc that becomes self-gravitating beyond the radius where the Toomre $Q = 1$. The resulting ejected disc masses for recoils with $v_{\rm k}/v_{\rm esc} > 0.1$ are shown in Fig. 7, for the `random-dry` spin model. The overall distribution of disc masses peaks at $\sim 0.1 M_{\rm BH}$ and has a tail extending to $< 10^{-5} M_{\rm BH}$. Among observable offset AGN, very few have $M_{\rm disc} \sim M_{\rm BH}$, because these correspond to low-velocity recoils; this is true regardless of the degree of spin alignment. The disc masses for observable velocity offset AGN are slightly smaller than for spatially offset AGN, because the minimum velocity offset requirement means that the former have higher kick speeds on average. In the random spin models, most offset AGN have $M_{\rm disc}/M_{\rm BH} = 1$–10 per cent. In the hybrid and aligned models (with lower kick velocities), typical disc masses are slightly higher, ~ 3–30 per cent $M_{\rm BH}$.

Fig. 8 shows the relationship between $v_{\rm k}/v_{\rm esc}$ and the spatially or velocity offset AGN lifetime. Here and throughout our analysis, a random viewing angle is assigned to each recoil event, and the offset lifetime is calculated as the total time for which the recoiling AGN is detectable and has a resolvable projected spatial offset or a resolvable (LOS velocity offset. The `random-dry` and
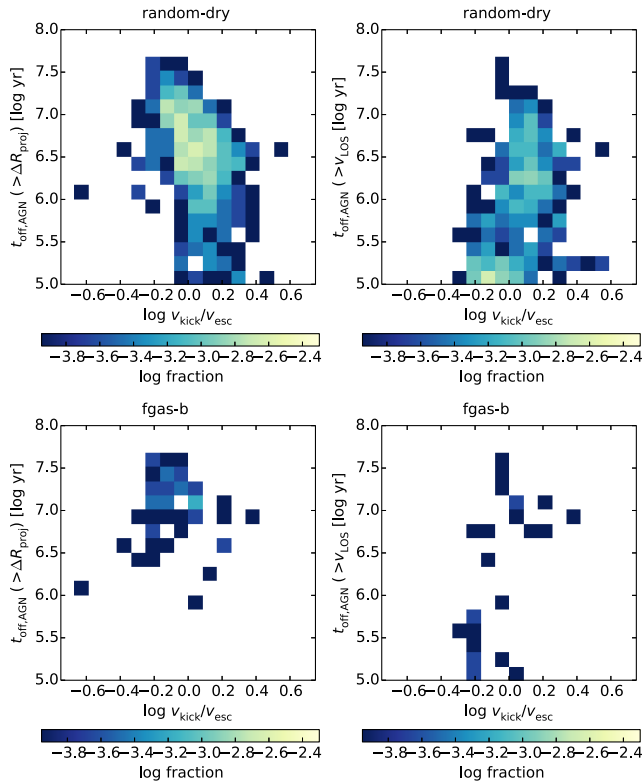
Figure 8. The distributions of observable offset AGN lifetime versus v_k/v_{esc} are shown for resolvable projected spatial offsets (left-hand panels) and for LOS velocity offsets (right-hand panels), weighted by the total number of recoil events. *HST*-COSMOS sensitivity and resolution, and a minimum $\Delta v_{LOS} > 600$ km s^{-1}, are assumed. The top panels correspond to the `random-dry` spin model, which has 8800 total recoil events, 492 of which yield an observable spatially offset AGN, and 413 of which produce an observable velocity offset AGN. The bottom panels correspond to the `fgas-b` model, which has 8989 total events, 54 (30) of which produce observable spatial (velocity) offsets. In all cases, the longest lifetimes occur for kick speeds near the host escape speed.

`fgas-b` spin models are shown, assuming *HST*-COSMOS sensitivity (AB(*F814W*) = 27.2 mag for point sources; Koekemoer et al. 2007), resolution (0.1 arcsec), and $\Delta v_{LOS} > 600$ km s^{-1}.

The spatially offset AGN lifetimes are longest for recoils at or somewhat below the escape speed ($v_k/v_{esc} \sim 0.6$–1). This is because marginally bound BHs experience little deceleration and spend most of their time at large apocentres, well separated from the host centroid. Escaping BHs also produce large offsets, of course, but this is countered by the fact that the accretion disc mass carried with the BH decreases $\propto v_k^2$, such that AGN lifetimes are vanishingly small for $v_k \gg v_{esc}$. Such extreme kicks are also rare. At high redshift, escaping BHs do begin to dominate the population of observable offset AGN, owing to both angular resolution limits and the suppression of low-velocity recoils in gas-rich mergers. The fact that offset AGN lifetimes are maximized for $v_k \lesssim v_{esc}$ has important consequences for the observability of recoils, foremost that spatial resolution should not be a strongly limiting factor in systematic searches for recoiling AGN.

Similarly, velocity offset AGN have the longest lifetimes for recoils near the escape speed (Fig. 8, right-hand panels). On average, velocity offset AGN have shorter lifetimes than spatially offset AGN, because many recoil events have $v_k > 600$ km s^{-1} initially

but are quickly decelerated in the host potential. This is especially true at high redshift, where gas-rich mergers tend to suppress long-lived recoils.

These trends are qualitatively similar for models that include pre-merger BH spin alignment, except that there are fewer offset AGN overall, and the high-v_k/v_{esc} tail of events is mostly truncated (Fig. 8, bottom panels). Recoil events in gas-rich mergers that produce short-lived velocity offset AGN in the `random-dry` model are mostly suppressed in the hybrid model.

3.4 Spatial and velocity offsets

Fig. 9 shows time-weighted distributions of projected spatial (ΔR_{proj}) and velocity (Δv_{LOS}) offsets for *all* recoiling AGN in the simulation (i.e. not only those with resolvable offsets are shown). The panels compare the distributions for the `random-dry`, `fgas-b`, and `5deg` spin models. For the `random-dry` model, the distribution is dominated by recoiling AGN with velocities above a few hundred km s^{-1} and spatial separations $\gtrsim 0.1$ arcsec (physical separations $\gtrsim 1$ kpc are common). This follows from the result discussed above: recoils with $v_k \lesssim v_{esc}$ have the longest offset lifetimes. The largest offsets (>1000 km s^{-1} and >10 kpc) result primarily from BHs that escape their host galaxies, and the angular offset distribution has a tail that extends beyond 10 arcsec. This demonstrates that seeing-limited surveys could resolve spatially offset AGN.

The right-hand panel of Fig. 9 demonstrates that if spins are always nearly aligned, spatial offsets >0.1 arcsec are rare, and resolvable velocity offsets (>600 km s^{-1}) never occur in the `5deg` spin model. The confirmation of a single recoiling AGN with a resolvable velocity offset would therefore be a strong indication of spin misalignment $\gtrsim 5°$. We also see that the ΔR_{proj}, Δv_{LOS} distribution is no longer skewed towards the largest offsets in this case, owing to the lack of escaping BHs and short return times in this spin model.

For the `fgas-b` hybrid spin model (middle panels of Fig. 9), the distributions of ΔR_{proj} and Δv_{LOS} offsets are bimodal. Offsets greater than 0.5 arcsec and 600 km s^{-1} are produced almost exclusively in gas-poor mergers, for which BH spins are assumed to be random. Small offsets are produced mainly by mergers with $f_{gas, sf} > 0.1$, for which BH spins are nearly aligned. This creates distinct signatures in the host galaxy population relative to the non-hybrid spin models, as we show in Section 3.6.

In Fig. 10, the time-weighted distributions of bolometric AGN flux and Eddington ratio are shown versus ΔR_{proj} and Δv_{LOS}, for the `random-dry` spin model. As in Fig. 9, all recoiling AGN are shown, not just those with resolvable offsets. We see that for angular separations $\gtrsim 1$ arcsec, recoiling AGN are generally above the flux limit. This reflects the fact that the largest angular separations occur at low redshift, where the minimum AGN luminosity is generally determined by the criterion $f_{Edd} > 10^{-2}$ instead of by the flux limit. Even in the case of velocity offset AGN, where there is no dependence on angular scale, offset AGN lifetimes at low z are usually limited by the f_{Edd} criterion rather than the flux limit, at least for *HST* sensitivity. The bottom panels in Fig. 10 confirm that the population of recoiling AGN is dominated by low Eddington ratios (<0.1), particularly for large ΔR_{proj} and Δv_{LOS} offsets. (Lower velocity recoils typically return to the galactic centre before their luminosity can decline substantially.)

The same distributions are shown for the `5deg` spin model in Fig. 11. Owing to the short return times of recoils in this model, the flux limit is rarely reached, and the f_{Edd} distribution is heavily skewed towards high values.
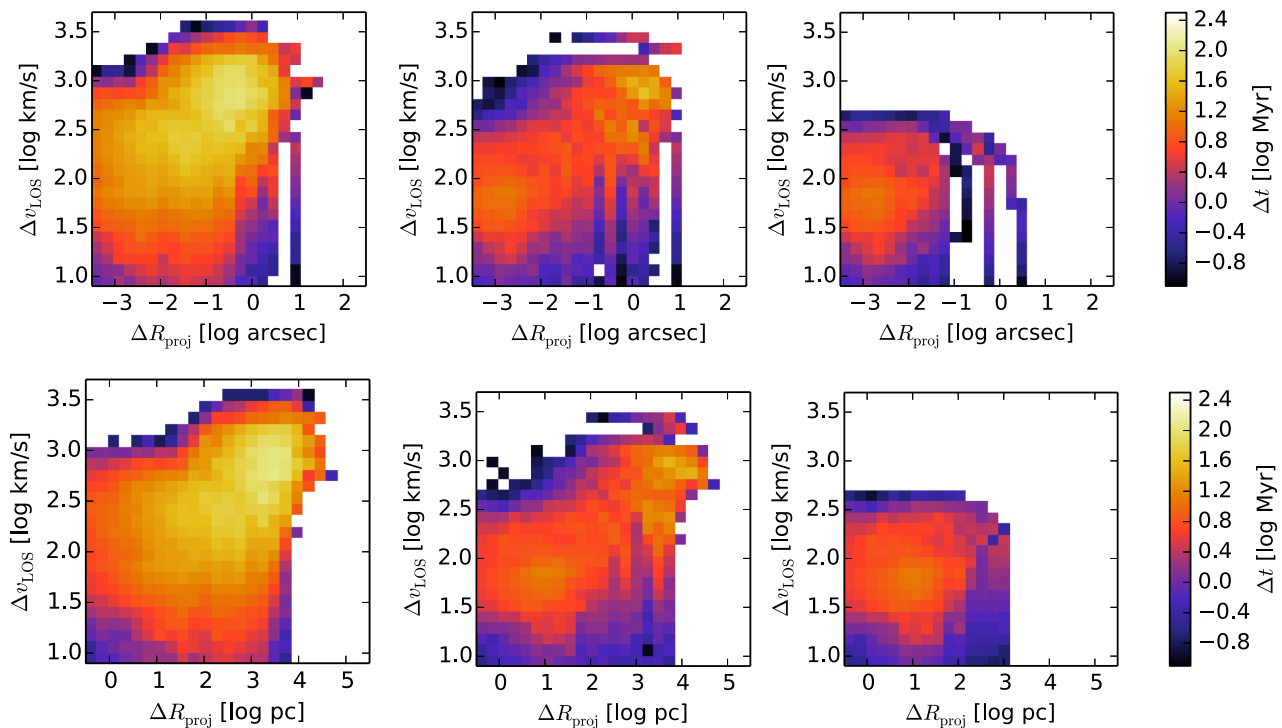
Figure 9. The time-weighted distributions of projected spatial offset versus LOS velocity offset are shown for all recoiling AGN. The colour scale indicates the amount of time a BH spends at a given (ΔR_{proj}, Δv_{LOS}) while active as an AGN and offset from the host nucleus, integrated over all merger events over cosmic time, *regardless* of whether the recoil event produces a resolvable offset. The sensitivity of *HST*-COSMOS is assumed. The top panels show the distribution of angular offsets, and the bottom panels show the physical spatial offsets. The left-hand panels correspond to the `random-dry` spin model, middle panels show the `fgas-b` model, and right-hand panels show the `5deg` model. Large offsets >1 kpc and >1000 km s^{-1} are *favoured* in the random spin model but do not occur at all if BH spins are nearly aligned.

One important caveat to these findings is that our models include only accretion on to recoiling BHs from the disc carried along with the BH at the time of the kick. As shown in Blecha et al. (2011), recoiling BHs on bound orbits may encounter fresh fuel on subsequent passages through the galactic centre, substantially extending the AGN lifetime in some cases. This may create a population of offset AGN with smaller spatial offsets, such that the actual distribution is less biased towards large separations than our results suggest. We also assume the recoils always occur in a dense stellar cusp that forms during the galaxy merger (e.g. Mihos & Hernquist 1994; Hopkins et al. 2008b, 2009). If a recoil event instead occurs in a cored elliptical (perhaps following a dry merger, where the BH inspiral time may be long), the BH may have a long phase of small-scale oscillations about the galactic centre (Gualandris & Merritt 2008; Lena et al. 2014). However, any new gas supply encountered by the recoiling BH on small scales would only *add* to the offset lifetimes predicted by our models.

Offset AGN that exhibit *both* velocity and spatial offsets, such as CID-42 (Civano et al. 2010, 2012), will be especially compelling recoil candidates. We find that a *majority* of velocity offset AGN (weighted by offset lifetime) will have projected spatial offsets resolvable with high-resolution imaging (e.g. *HST*). This fraction increases with increasing spin alignment, from $\gtrsim 50$ per cent in the random spin models to ~90 per cent in the `cold` model. Thus, high-resolution imaging may be a fruitful means of follow-up for any velocity offset recoil candidates. Even with SDSS resolution, ~4–20 per cent of velocity offset AGN may have resolvable spatial offsets; these primarily occur at low redshift ($z < 0.5$).

In contrast, only ~10–20 per cent of spatially offset AGN should have discernible velocity offsets $\Delta v_{LOS} > 600$ km s^{-1} (<15 per cent in the aligned spin models). With a more stringent velocity offset criterion ($\Delta v_{LOS} > 1000$ km s^{-1}), <10 per cent of spatially offset AGN exhibit simultaneous velocity offsets, and no spatial/velocity offset overlap occurs in the aligned spin models. However, we demonstrate below that even this small fraction of spatially selected offset AGN exhibiting velocity shifts could yield a population of strong recoil candidates.

One difficulty with confirming the nature of objects such as CID-42 is that the offset AGN is superimposed on the host galaxy bulge, such that it cannot be easily distinguished from a member of an inspiralling BH pair. CID-42 has no evidence for a second AGN in the host nucleus, but excluding the presence of a *quiescent* BH is more difficult. The strongest offset AGN candidates will likely be those that have not only a resolvable spatial offset (and possibly a simultaneous velocity offset), but are also well separated from the host galaxy's stellar light.

We find that a non-negligible (time-weighted) fraction of spatially offset AGN resolvable with *HST* also have separations $\Delta R_{proj} > R_{bulge}$, where R_{bulge} is the radius of the stellar bulge component in our models. Specifically, ~20–30 per cent of spatially offset AGN have $\Delta R_{proj} > R_{bulge}$ in the random and hybrid spin models, and a similar fraction of AGN with simultaneous spatial and velocity offsets are also found outside the stellar bulge. Separations $> R_{bulge}$ are rare in the aligned spin models. For offset AGN resolvable with SDSS, ~40–50 per cent of spatially offset AGN may have $\Delta R_{proj} > R_{bulge}$ (in the random and hybrid spin models), and ~20–40 per cent of simultaneously spatially and velocity offset
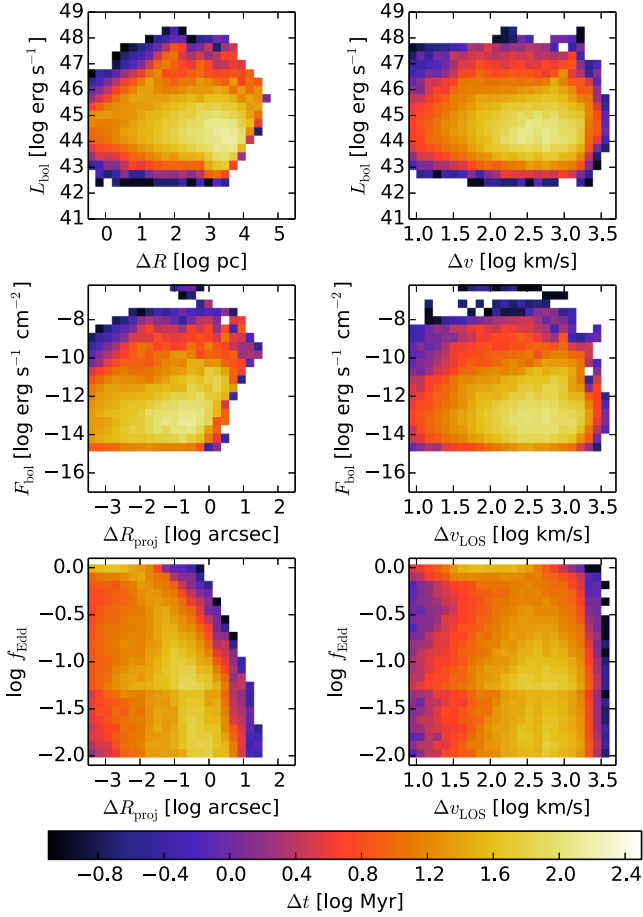
Figure 10. Similar to Fig. 9, the time-weighted distributions of projected spatial offsets (left-hand panels) and LOS velocity offsets (right-hand panels) are shown for all recoiling AGN in the `random-dry` spin model (and *HST*-COSMOS sensitivity), for all timesteps during each recoil event and for all redshifts. Here the offsets are plotted versus the bolometric AGN luminosity (top panels), bolometric AGN flux (middle panels), and Eddington ratio L_{bol}/L_{Edd} (bottom panels). The Eddington ratio is constrained to have a maximum of 1 and a minimum of 10^{-2} in our models. The feature apparent at $f_{Edd} = 0.05$ arises from the transition to the radiatively inefficient accretion regime. Most offset AGN have $f_{Edd} < 0.1$, and for spatial offsets $\gtrsim 1$ arcsec, observable offset AGN are rarely flux-limited.

Figure 11. The same distributions as in Fig. 10 are shown, but for the `5deg` spin model. The short return times of recoiling BHs in this model prevent the flux limit from being reached in most cases.

AGN should be found at such separations. We caution that the stellar bulges in our galaxy potential models may be more compact than in reality if there is a long delay between the galaxy coalescence and the BH merger and recoil, so R_{bulge} may be underestimated in some cases. None the less, these results indicate that a subsample of offset AGN displaced beyond the host stellar bulge could be found, and these would be ideal targets for follow-up and confirmation of recoiling BHs.

3.5 Recoiling AGN observability

3.5.1 Redshift distribution and general trends

The predicted number of observable offset AGN per square degree is shown in Fig. 12, as a function of redshift, for various surveys and for three different spin models (`random-dry`, `fgas-b`, and `cold`). In the `random-dry` spin model, 2–3 per deg² spatially offset AGN could be detected in *HST*-COSMOS (Koekemoer et al. 2007) out to $z = 3$, and ~1 per deg² could be found at $z < 1$.
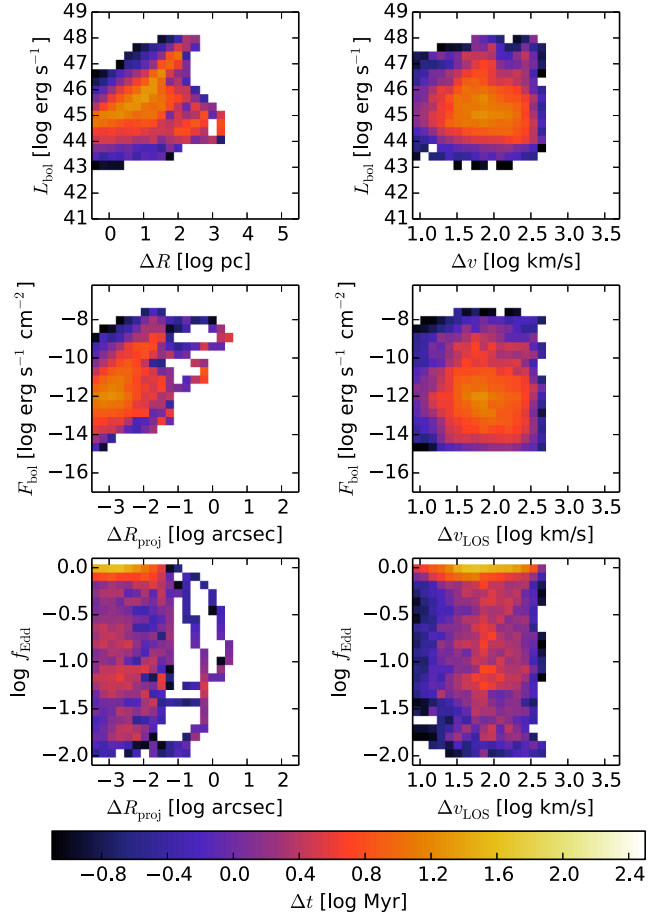
The predicted source counts decrease with survey sensitivity and spatial resolution; SDSS could see about 1 spatially offset AGN per 40 deg², and none should be detectable at $z \gtrsim 0.8$.

HST could detect *velocity* offset AGN in slightly lower numbers; $\gtrsim 1$ per deg² are predicted out to $z = 3$ for a minimum resolvable offset of $\Delta v_{LOS} > 600$ km s⁻¹. In general, for angular resolutions better than ~0.5 arcsec, we find that spatially offset AGN should be found in greater numbers than velocity offset AGN. The reverse is true for seeing-limited surveys with spectroscopic coverage; in the random spin model, SDSS should find ~10 times more velocity offset AGN than spatially offset AGN for a minimum $\Delta v_{LOS} > 600$ km s⁻¹. Even if only offsets $\Delta v > 1000$ km s⁻¹ can be resolved, velocity offsets could still be detected in greater numbers with SDSS.

The predicted number of observable velocity offset AGN (for $\Delta v_{LOS} > 600$ km s⁻¹) differs by at most a factor of a few between *HST* and SDSS, and even less at low redshift, despite the large difference in sensitivity. As discussed above, the imposed minimum AGN Eddington ratio of 10^{-2} generally corresponds to a higher luminosity limit than does the assumed flux limit, for low to moderate redshifts. Thus, our results depend only weakly on the absolute flux limit, which becomes important only at higher z.

This trend is even more prominent in the hybrid spin models, where there is almost no dependence on the assumed flux
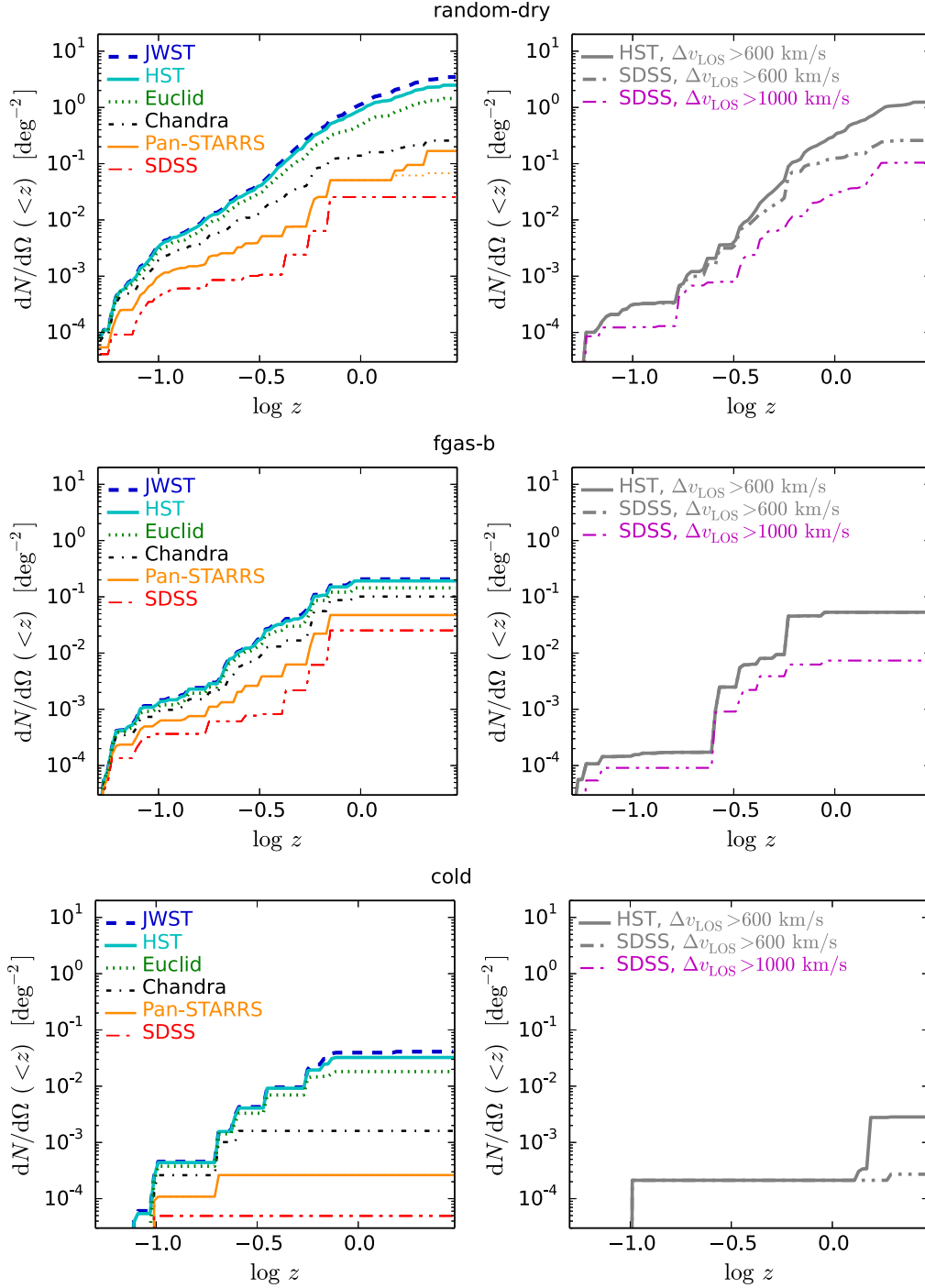
Figure 12. The cumulative number of offset AGN observable per deg^2 is shown as a function of survey volume (in log z), for flux-limited surveys with a given resolution and sensitivity. Top, middle, and bottom panels show results for the `random-dry`, `fgas-b`, and `cold` spin models, respectively. The left-hand panels show the number of observable spatially offset AGN. As indicated on the plot legends, the thick solid blue curve corresponds to the sensitivity of *HST*-COSMOS in the *F*814*W* band, assuming 0.1 arcsec resolution, and the dashed cyan curve corresponds to the target *JWST* sensitivity at 2 μm, assuming 0.07 arcsec resolution. The green dotted line corresponds to the target sensitivity of the *Euclid* wide survey in the visible band, assuming 0.2 arcsec resolution. The black dash–dotted line corresponds to *Chandra* observations at the sensitivity of the C-COSMOS survey at 2–10 keV, for 0.5 arcsec resolution, and the thin solid orange line denotes the PS1-MDS result for the *r* band, assuming a seeing-limited resolution of 1 arcsec. (The dotted orange line, visible only in the random spin model as distinct from the MDS, corresponds to the PS1-3pi survey). Finally, the dot–dot–dashed red line shows results for the SDSS *r* band, for a PSF size of 1.4 arcsec. In each case, the minimum resolvable spatial offset $\Delta R_{\rm proj}$ is assumed to be twice the resolution limit. The right-hand panels show predictions for observable velocity offset AGN. Only results for *HST* (solid line) and SDSS (dot–dot–dashed lines) are shown. Grey lines correspond to a minimum resolvable LOS offset of $\Delta v_{\rm LOS} > 600$ km s^{-1}. For SDSS, we also show results for $\Delta v_{\rm LOS} > 1000$ km s^{-1} (magenta line). The random models predict up to a few spatially offset AGN per deg^2 detectable with space-based instruments, and up to one velocity offset AGN per 10 deg^2 with $\Delta v_{\rm LOS} > 1000$ km s^{-1} observable in the SDSS. Even in the hybrid models, where spins are nearly aligned in gas-rich mergers, up to a few spatially offset AGN per 10 deg^2 could be found.

limit. Relative to the random spin models, the hybrid models produce fewer superkicks, and those that do occur are limited to dry mergers, which are found predominantly at low redshifts (virtually none occur at $z > 1$). Thus, the AGN lifetime is almost always set by the minimum f_{Edd}, and increasing the angular resolution only moderately increases the number of observable spatial offsets for these low-redshift sources (∼factor of 10–12 higher number counts for a factor of 20 increase in spatial resolution). Note that the SDSS predictions for spatially offset AGN are similar between the random and hybrid spin models; most recoil events in the random spin model that produce a spatial offset observable with SDSS also produce an observable offset in the hybrid spin model. This reflects the fact that SDSS is most sensitive to luminous AGN at low to moderate redshift, which are more likely to occur in dry mergers that can produce superkicks in the hybrid model.

In the aligned (`cold`) spin model, fewer observable recoils are predicted than in the random or hybrid models, as expected, and there is a steeper dependence on angular resolution. Only one spatially offset AGN per 30 deg^2 should be detectable with *HST*, and velocity offset AGN should occur at a rate of <1 per 100 deg^2.

We note that, despite the prevalence of major mergers among BHs with $M_{BH} < 10^7$ M$_\odot$ (Fig. 3), these low-mass BHs do not dominate the number counts of observable offset AGN in any of our models. In fact, for the hybrid spin models, massive BHs ($M_{BH} > 10^8$ M$_\odot$) dominate the sample of offset AGN, because these BHs are preferentially hosted in gas-poor galaxies where superkicks can occur. Additionally, for seeing-limited surveys, spatially offset AGN tend to have high BH masses regardless of spin model. The offset AGN lifetime depends on BH mass, such that low-mass BHs will often exhaust their fuel supply before reaching a resolvable separation from the host galaxy.

This trend occurs even though massive BHs are generally hosted in massive galaxies, which have high central escape speeds. As a result, only recoiling AGN with $v_k/v_{esc} \gtrsim 0.8$ produce spatial offsets that are resolvable with seeing-limited surveys. One caveat to this is that our sample of dwarf galaxies is incomplete owing to mass resolution limits. A more complete sample of dwarf galaxies could produce a population of lower mass AGN with resolvable offsets.

As mentioned above, our accretion disc model generally results in longer recoiling AGN lifetimes than if a constant accretion rate is assumed. This is especially important for spatially offset AGN, for which offset AGN lifetimes are more often limited by declining luminosity than by a return to the galactic centre. (Velocity offset AGN have shorter lifetimes on average, owing to deceleration by the host galaxy potential.) In most cases, our models predict lower rates of observable spatially offset AGN by up to a factor of a few if a constant $M_{disc}/\dot{M}_{BH}$ lifetime is assumed. The difference is largest for the aligned spin models and seeing-limited resolution, where in the constant-$\dot{M}$ model, few BHs achieve a resolvable separation during their active lifetime. In contrast, velocity offset AGN lifetimes are often limited by deceleration in the host potential, rather than by the luminosity limit, and the predicted numbers of observable velocity offset AGN can be a factor of a few lower *or* higher when a constant $\dot{M}_{BH}$ is assumed.

3.5.2 Imaging surveys

Fig. 13 shows the cumulative number of spatially offset AGN per square degree at $z < 3$ observable in selected current and future surveys, for all spin models considered in this study. We first examine the random spin models: `random-high` and `random-dry` (left-hand panels). Up to 7 per deg^2 spatially offset AGN may be detectable with *JWST*. In the hard X-ray band, up to ∼1 per deg^2 offset AGN could be detected at the sensitivity of the *Chandra Deep Field*-North (CDF-N; Alexander et al. 2003) survey, and ∼0.3 per deg^2 at the limit of the *Chandra*-COSMOS Legacy (C-COSMOS; Civano et al. 2013) survey. Seeing-limited surveys (assuming 1 arcsec effective resolution) may resolve up to 1 spatially offset AGN per 4 deg^2, and even SDSS, with an effective PSF size of 1.4 arcsec, could find up to 1 per 14 deg^2.

The bottom panels of Fig. 13 show the total predicted number of observable spatially offset AGN in a given survey, versus survey area. Results for spatial offsets are given for *HST*-COSMOS (and a hypothetical *JWST* survey with the same area), a *Euclid* wide survey (Laureijs et al. 2011), a 2000 deg^2 WFIRST weak-lensing survey (Spergel et al. 2013), CDF-N, C-COSMOS, PS1-MDS, PS1-3pi, SDSS DR7, and the LSST Main Survey. The random spin models predict that five to nine spatially offset AGN may be found in *HST*-COSMOS, while none are likely to be identified with *Chandra* in the CDF-N or C-COSMOS surveys. In contrast, future wide-area, high-resolution surveys could detect spatially offset AGN in large numbers: WFIRST could detect several thousand such objects, and *Euclid* could detect tens of thousands.[3]

An important prediction of our models is that, because offset AGN spend most of their time at large separations, seeing-limited surveys should be able to resolve spatially offset AGN. This suggests that offset AGN may be promising targets for ground-based, large-area surveys. For random spins, ∼15–20 offset AGN may be found in PS1-MDS, and hundreds of spatially offset AGN could be found in SDSS. The predictions are even more optimistic for LSST and PS1-3pi, which could detect up to several thousand spatially offset AGN.

Allowing for some amount of pre-merger BH spin alignment reduces the number of observable offset AGN, but does not eliminate the possibility of detecting recoils (Fig. 13, middle and right-hand panels). For the hybrid spin models, less than one offset AGN would be expected in *HST*-COSMOS. However, two to three may be detectable in PS1-MDS, and large populations of spatially offset AGN are still detectable in SDSS (up to ∼200 in DR7), LSST (up to ∼700 in the Main Survey), and WFIRST (> 300). The hybrid spin models also predict that the PS1-3pi and *Euclid* wide surveys could each detect in excess of 10^3 spatially offset AGN.

The predictions for the three aligned spin models (right-hand panels of Fig. 13) differ by several orders of magnitude. For high-resolution observations (*HST*, *JWST*, *Euclid*, and WFIRST), the `hot` spin model produces comparable numbers of spatially offset AGN as the hybrid models. A few tens of offset AGN could be found with LSST and PS1-3pi, and $>10^3$ could be detected with *Euclid*. In the `cold` spin model, a handful might be detectable with LSST and PS1-3pi, but wide-area surveys with space-based instruments (i.e. WFIRST and *Euclid*) would be required to find a population of offset AGN. In the `5deg` model, the maximum displacements achieved barely exceed the angular resolution of SDSS, and offset AGN are unlikely to ever be detected – even *Euclid* is predicted to find at most a few objects over the entire sky.

---

[3] The predictions for *Euclid* assume a point-source sensitivity of 27.3 AB mag in the visible band; the predicted numbers are up to a factor of 2 lower for an assumed sensitivity of 24 AB mag in the *J* band.
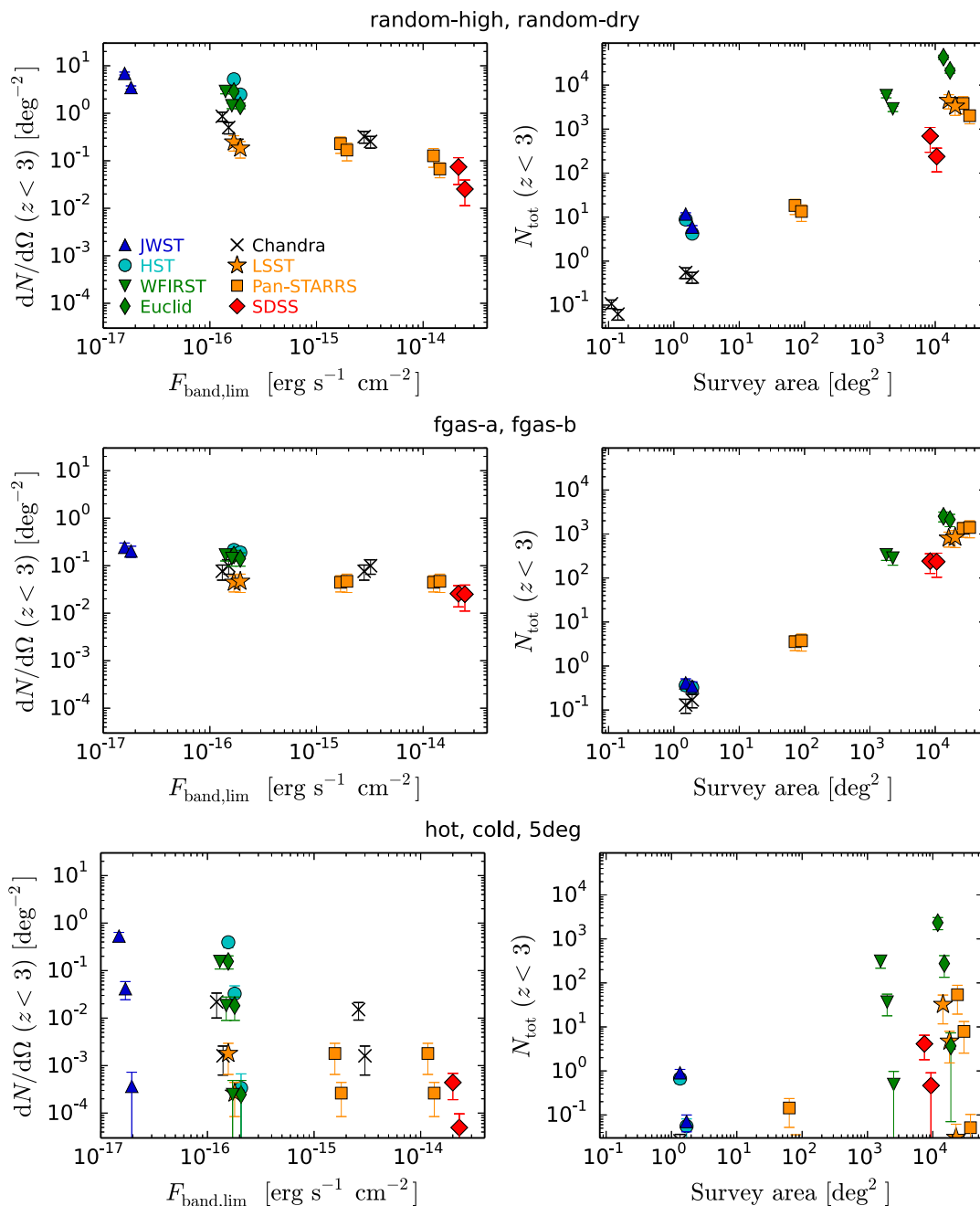
Figure 13. The cumulative number of observable spatially offset AGN at $z < 3$ is shown for various surveys. The top panels show the random spin models (`random-high` and `random-dry`, from left to right; a small offset along the x-axis is added for clarity). The middle panels show the hybrid spin models (`fgas-a` and `fgas-b`) in the same manner. The bottom panels show the aligned spin models (`hot`, `cold`, and `5deg`). Left column: for each spin model and observational survey, the cumulative number of observable recoiling AGN with resolvable (projected) spatial offsets *per square degree*, out to $z = 3$, is shown versus the limiting flux in the relevant band. For the optical and NIR bands, the quantity plotted is νF_ν for the effective wavelength of the band, while the X-ray fluxes correspond to $F_{2-10\,\mathrm{keV}}$. Points correspond to various observational surveys, as indicated in the plot legend and as follows. Blue triangles: *JWST* at 2 μm, 0.07 arcsec resolution. Cyan circles: *HST*-COSMOS in the *F*814*W* band, 0.1 arcsec resolution. Green inverted triangles: WFIRST high-latitude survey (*J* band, 0.2 arcsec resolution). Green thin diamonds: *Euclid* wide survey (visible band, 0.2 arcsec resolution). Black X's: C-COSMOS and CDF-N (2–10 keV, 0.5 arcsec resolution). Orange points: a seeing-limited resolution of 1 arcsec is assumed for LSST (Main Survey *r*-band; orange stars) and Pan-STARRS (*r*-band, for PS1-MDS and PS1-3pi; orange squares). Red diamonds: SDSS DR7, assuming a PSF size of 1.4 arcsec for the *r* band. In each case, the minimum resolvable spatial offset ΔR_{proj} is assumed to be twice the resolution limit. Error bars are derived from time-weighted Poisson statistics for the cumulative number counts. Right column: the total number of spatially offset AGN that may be detectable in a given survey is shown, versus survey area. In the same manner as above, the points correspond to *HST*-COSMOS and a hypothetical *JWST* survey of the same area (1.7 deg²), a 15 000 deg² *Euclid* wide survey, a 2000 deg² WFIRST high-latitude survey, the 448 arcmin² CDF-N and 1.7 deg² *Chandra*-COSMOS surveys, the 80 deg² PS1-MDS, SDSS DR7, the 18 000 deg² LSST Main Survey, and the 30 000 deg² PS1-3pi survey. The random spin models predict that $>10^3$ offset AGN could be found with Pan-STARRS, LSST, and WFIRST, and $>10^4$ with *Euclid*. The hybrid models predict $\sim 10^2$–10^3 offset AGN to be detectable in these surveys. In the most extreme aligned model (`5deg`), however, recoiling AGN may never be detected.

3.5.3 Spectroscopic surveys

Because SDSS also has spectroscopic coverage, we can estimate the number of detectable velocity offset AGN as well. The random spin models produce $\sim$0.1 per deg^2 velocity offset AGN with $\Delta v_{\rm LOS} > 1000$ km s^{-1} (Fig. 12, left-hand panel), implying that nearly 10^3 such objects should exist in the DR7 footprint.

This result is of particular interest, because to date, the SDSS spectroscopic quasar data base is the only data set on which large-scale systematic searches for offset AGN have been carried out. Bonning et al. (2007) searched for quasar spectra at $0.1 < z < 0.81$ in which the broad Hβ and Mg II lines have symmetric profiles and consistent velocity shifts relative to the NLs. They found a null result for $\Delta v > 800$ km s^{-1}. Using different selection criteria, Tsalmantza et al. (2011) searched for shifted BLs in quasar spectra at $0.1 < z < 1.5$ and identified 32 candidates of interest. Eracleous et al. (2012) searched for BL offsets of $\Delta v > 1000$ km s^{-1} in Hβ for quasars at $z < 0.7$ and found 88 such velocity offset quasars. However, these BL offsets could also arise from high-velocity outflows, double-peaked emitters, or binary BH motion, and none have been confirmed to date as recoiling BHs.

For our random spin models, only $\sim$10–30 per cent of all velocity offset AGN occur at $z < 0.7$, such that up to several hundred offset AGN are predicted in the relevant redshift range. This is still substantially more than are present in the DR7 quasar catalogue, but we stress that these observational results cannot be compared at face value to our models. There are several important distinctions between the offset AGN in our sample and the selection criteria for SDSS quasars (Schneider et al. 2010), which will cause some offset AGN to be excluded from the latter. The initial selection of SDSS quasar candidates is based on their colours, which means that low-luminosity AGN whose flux is dominated by the host galaxy are excluded.

The SDSS quasars are also required to have a measurable BL with FWHM > 1000 km s^{-1}, such that Type II (narrow-line) AGN are excluded as well. However, a BL is also required for detecting velocity offset AGN, meaning that some fraction of velocity offset AGN should be undetectable in any case.

We have also neglected the effects of obscuration in our models. Obscuring dust in the nuclear region is a very common feature of AGN, required to explain the optical, mid-IR, and X-ray spectral features of many objects, and may often have a toroidal geometry (e.g. Urry & Padovani 1995; Ueda et al. 2003; Stern et al. 2005). While recent work has focused primarily on the role of dusty tori in obscuring low-velocity recoil events (Raffai, Haiman & Frei 2016), we find that dust obscuration is unlikely to be a major factor in the observability of spatially offset AGN, at least for kick distributions that assume some BH spin misalignment. This is because little of the obscuring material beyond the BL region should remain bound to the BH, and Fig. 9 illustrates that offset AGN are most often found at separations larger than the scale of the torus (see also Komossa & Merritt 2008). In contrast, *velocity* offsets will occur essentially at the moment of recoil, and thus may be obscured from the observer for the early phase of their lifetime. If obscuring material extends beyond the dusty AGN torus to galactic scales, owing perhaps to a coincident nuclear starburst, some recoiling AGN could be obscured for a significant portion of their offset lifetimes. Recall, however, that a majority of velocity offset AGN should also have spatial offsets detectable with *HST*, and up to 20 per cent may have spatial offsets detectable with SDSS. Thus, obscuration may have at most a moderate effect on the observability of velocity offset AGN.

The SDSS DR7 quasars are additionally required to be fainter than $i \approx 15$ mag (to avoid saturation of the spectrograph) and more luminous than $M_i = -22$. These magnitude limits, at least, have only a modest effect on the offset AGN population at $z < 0.7$, if random spins are assumed. If we approximate these cuts for our sample [using the same optical bolometric corrections from Hopkins et al. (2007) that are used throughout the text], the predicted number of velocity offset AGN ($\Delta v > 1000$) is reduced by about 30 per cent for the `random-dry` model and by only 5 per cent in the `random-high` model.

An obvious alternative explanation for the observed rarity of large BL offsets in SDSS is that pre-merger spin alignment plays at least some role in reducing the number of superkicks with $v_{\rm k} > 1000$ km s^{-1}. The hybrid spin models predict that $\sim$70 (`fgas-b` model) to 220 (`fgas-a`) velocity offset AGN with $\Delta v > 1000$ km s^{-1} may exist in the SDSS footprint, most of which are found at $z < 0.7$. With the estimated magnitude limits described above, these numbers are reduced to $\sim$30 and 200, respectively. Thus, even if spins are preferentially aligned only in gas-rich mergers, our results imply that a population of velocity offset AGN may be missed by the SDSS DR7 selection criteria.

In the `hot` aligned spin model, up to 10 velocity offset AGN (with $\Delta v_{\rm LOS} > 1000$ km s^{-1}) could be found with SDSS, with only $\sim$2 at $z < 0.7$. None are predicted to be observable in the `cold` or `5deg` models. Thus, follow-up studies of candidate recoils in SDSS could place indirect constraints on pre-merger BH spins.

3.6 Host galaxy properties

To characterize the host galaxies of offset AGN, we first examine their stellar mass distribution and its evolution with redshift (Fig. 14). The stellar mass distribution for all BH merger hosts peaks at $\sim$10^{10} M$_\odot$ at the highest redshifts, and evolves via hierarchical growth towards larger masses at lower z. The distribution is limited at both the low- *and* high-mass end by the resolution of the simulation; specifically, the minimum BH mass and stellar/halo particle cuts imposed on each progenitor galaxy exclude many dwarf–dwarf mergers and satellite–halo mergers.

For the random spin models, and assuming *HST*-COSMOS sensitivity and resolution, the offset AGN hosts have a broad mass distribution that peaks at $\log (M_*/{\rm M_\odot}) \sim 10$–11 (Fig. 14, top panel). At $z < 1$, offset AGN favour lower mass hosts relative to the overall distribution. Nominally, low-mass galaxies should more easily produce offset AGN owing to their low central escape speeds, and indeed, for the random spin models, nearly 30 per cent of BH mergers occurring in galaxies with $\log (M_*/{\rm M_\odot}) < 10$ at low redshifts give rise to an observable spatially offset AGN.

At $z > 1$, BH merger hosts have lower M_* on average, but the spatially offset AGN hosts do not. At most a few per cent of low-mass galaxy merger remnants ($\log (M_*/{\rm M_\odot}) < 10$) produce observable spatial offsets. In part, this is because low-mass galaxies generally host smaller BHs with lower AGN luminosities. These low-mass bulges are also increasingly gas-rich at high z, making it difficult for BHs to escape the nucleus.

Fig. 15 shows the distribution of M_* versus the star-forming gas fraction ($f_{\rm gas, sf}$) for all BH merger hosts and for spatially offset AGN hosts. The merging galaxy population shows a strong 'quenching' trend evolving from low-mass, gas-rich galaxies at high z to higher-mass, gas-poor galaxies at $z \sim 0$. Comparing with the hosts of spatially offset AGN, we see that they follow the same general trend, but that galaxies with low to moderate gas fractions are favoured; extremely gas rich systems rarely produce offset AGN. This is
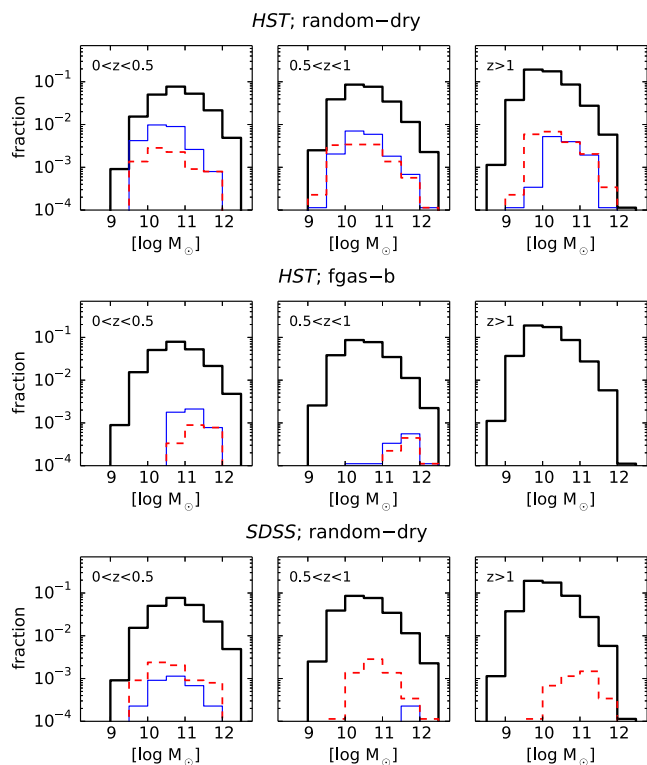
Figure 14. Top row: the distributions of host stellar masses are shown for all merging BHs in Illustris (thick black line), in three redshift bins. The host stellar mass is defined as the combined mass of the progenitor galaxies in the simulation. The thin solid blue line denotes the subset of galaxy merger remnants that host an observable spatially offset AGN, for the `random-dry` spin model and assuming *HST*-COSMOS sensitivity and resolution. The thin red dashed line similarly denotes the hosts of observable velocity offset AGN ($\Delta v_{\rm LOS} > 600$ km s^{-1}). In both cases, only offset AGN with a minimum offset lifetime $>10^5$ yr are shown. Middle row: same as the top row, but for the `fgas-b` spin model. Bottom row: same as the top row, but assuming SDSS sensitivity and resolution. The host mass distribution is broad but peaks at log $(M_*/{\rm M_\odot}) = 10$–11 for the random and aligned spin models. In the hybrid spin models, higher mass hosts are preferred despite their higher escape speeds, owing to their low gas content, and observable recoils are very rare at $z > 1$.
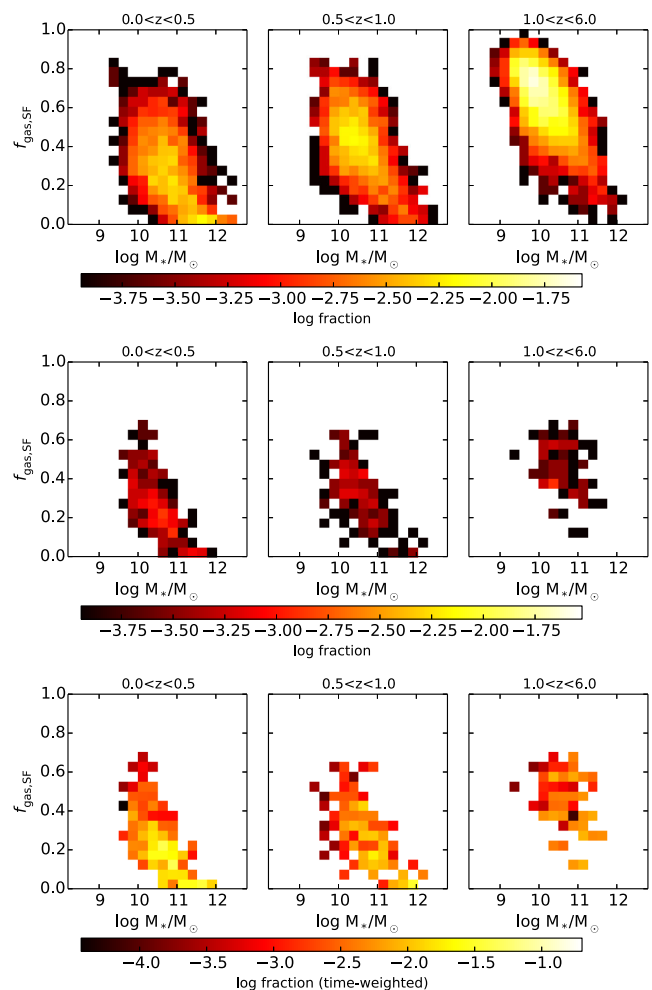

Figure 15. The distributions of stellar mass (M_*) versus the fraction of cold, star-forming gas ($f_{\rm gas,\,sf}$, relative to M_*) are plotted for different redshift bins. The top row shows the distributions for all merging BHs in our Illustris sample (8800 mergers), while the middle and bottom rows show only the 492 mergers that produce observable spatially offset recoiling AGN in the `random-dry` model, assuming *HST*-COSMOS resolution and sensitivity. In the middle row, the distributions are weighted by number, and in the bottom row, they are weighted by time. The overall population of merger hosts evolves from low masses and high gas content to higher mass, quenched galaxies. In contrast, the hosts of observable offset AGN favour lower gas fractions at all redshifts.

even clearer in the bottom panel of Fig. 15, which shows the same distribution weighted by offset AGN lifetime. The longest-lived offset AGN occur preferentially in galaxies with low $f_{\rm gas,\,sf}$.

The M_*–$f_{\rm gas,\,sf}$ distributions for velocity offset AGN hosts are similar to those for spatially offset AGN. Velocity offset AGN occur more often at $z > 1$ in low-mass, gas-rich galaxies than do spatial offsets, but these generally have short lifetimes, <1 Myr. The suppression of recoils in high-z, gas-rich galaxies is another reason that our results depend only weakly on the assumed flux limit, because it is primarily at high redshifts where the minimum observable flux is more constraining than the minimum $f_{\rm Edd} = 10^{-2}$.

In the aligned spin models, offset AGN are much rarer overall, and almost none occur in the `5deg` model. Recoil velocities are generally too small to produce observable offsets in either gas-rich galaxies or massive, quenched galaxies. As a result, offset AGN in the aligned spin models favour similar host stellar masses as the random models (log $(M_*/{\rm M_\odot}) \sim 10$–11). No spatially offset AGN are found at $z > 1$ in the `cold` spin model or at $z > 2$ in the `hot` model.

When the degree of BH spin alignment depends on the host gas fraction, as in the hybrid spin models, recoils are preferentially suppressed in gas-rich hosts by spin alignment in addition to the aforementioned dynamical effects. As a result, offset AGN are preferentially found in higher mass hosts, which have lower gas content on average. This is consistent with the above observation that massive BHs dominate the offset AGN population in the hybrid models. In the `fgas-b` model (Fig. 14, middle panels), the extreme alignment of the `5deg` spin model for gas-rich hosts eliminates virtually all observable offsets in hosts with log $(M_*/{\rm M_\odot}) < 10.5$. At $z > 1$, there are no merger remnants with $f_{\rm gas,\,sf} < 0.1$, and no observable offset AGN. The effect is less severe in the `fgas-a` model, where spins are less highly aligned, but higher-mass, low-$f_{\rm gas}$ galaxies are still favoured as offset AGN hosts.

To summarize, we find that observable offset AGN favour host masses of log $(M_*/{\rm M_\odot}) = 10$–11 in the random spin models and

that higher mass hosts are favoured in the hybrid spin models. Offset AGN are more likely to be found in hosts with low to moderate cold gas fractions, especially in the hybrid spin models. These trends indicate that the host properties of offset AGN could provide indirect information about BH spin alignment. However, because the distributions of M_* and $f_{\mathrm{gas, sf}}$ are broad, a statistical sample of offset AGN would be required to distinguish between spin models. Alternatively, finding even one example of an offset BH in a very massive galaxy would strongly argue for spin misalignment in at least some cases.

Distinguishing between BH spin models based on host stellar masses would also likely require space-based observations. The bottom panels of Fig. 14 show the M_* distribution for all BH merger hosts and for offset AGN hosts in the random-dry model, as in the top panels, but for SDSS resolution and sensitivity. Relative to the population detectable with *HST*-COSMOS, there are fewer offset AGN overall, and the preference for low-mass hosts largely disappears. Because lower mass BHs (hosted in lower mass galaxies) have shorter recoiling AGN lifetimes, they are more likely to exhaust their fuel supply before reaching an offset resolvable with ground-based instruments. Also, many low-mass BHs are too faint to be detected in SDSS at $z \gtrsim 0.5$.

We again note that the BH merger sample is incomplete in the dwarf mass regime ($\log (M_*/\mathrm{M_\odot}) \lesssim 10$) owing to mass resolution limits. However, the BH occupation fraction in such galaxies is not empirically constrained in any case. At the massive end, the resolution limits on minor mergers are unimportant, because these events do not produce observable recoils.

It is important to bear in mind that these galaxy merger remnant models are constructed using the *progenitor* galaxy properties in the simulation snapshot prior to each BH merger; as such, they do not account for evolution that may occur between the snapshot and the time of merger. In particular, the gas fractions shown in Fig. 15 can be considered 'initial' values, in that some of this gas will form stars or be expelled during the galaxy merger, prior to the BH merger and recoil.

3.7 Dependence on model parameters

3.7.1 Accretion disc parameters

Accretion on to recoiling BHs is assumed to occur via a thin, radiatively efficient 'α'-disc in the inner regions. At low Eddington ratios, theory predicts that accretion will become radiatively inefficient, and our accretion disc model will become unphysical. In order to avoid artificially long offset AGN lifetimes, a minimum Eddington ratio for the recoiling AGN must be imposed. This is especially important at low redshift, where AGN observations are generally not flux-limited. In our fiducial models, we adopt a minimum $f_{\mathrm{Edd}} = 10^{-2}$ for recoiling AGN, motivated by the distribution of Eddington ratios in observed quasars (e.g. Shen & Kelly 2012). This is a conservative choice; massive BHs in particular may be detectable at lower Eddington ratios. Here we explore the consequences of adopting a lower minimum value of $f_{\mathrm{Edd}} = 10^{-3}$.

The statistics of recoil events and properties of their hosts do not change; only the observable lifetime of each recoil event depends on the minimum Eddington ratio. Spatially offset AGN lifetimes can be $\gtrsim 100$ Myr in the $f_{\mathrm{Edd, min}} = 10^{-3}$ model, while the maximum lifetimes for $f_{\mathrm{Edd, min}} = 10^{-2}$ are ~40 Myr. Similarly, velocity offset AGN have a maximum lifetime of ~60 Myr, versus a maximum of ~30 Myr for $f_{\mathrm{Edd, min}} = 10^{-2}$. (Velocity offset AGN lifetimes are

more often limited by deceleration in the host potential than by the AGN luminosity.)

In most cases, setting $f_{\mathrm{Edd, min}} = 10^{-3}$ produces a factor of a few more spatially offset AGN per square degree than in the fiducial model. Because the AGN reach larger separations before falling below the minimum f_{Edd}, the biggest increase in number of observable recoils occurs for seeing-limited observations, where spatial resolution is most often a limiting factor. Particularly in the aligned spin models, where large spatial offsets are rare, the $f_{\mathrm{Edd, min}} = 10^{-3}$ model produces at least 10 times more spatially offset AGN per deg^2. Specifically, the hot spin model predicts 1–2 spatially offset AGN in *HST*-COSMOS and $>10^3$ in the LSST Main Survey, while even the cold spin model predicts nearly 100 spatially offset AGN to be observable with LSST. Velocity offset AGN number counts for $f_{\mathrm{Edd, min}} = 10^{-3}$ are more similar to the fiducial model, increasing by less than a factor of 2. This is because at low redshift, the offset AGN lifetimes are still more often limited by $f_{\mathrm{Edd, min}}$ than by the absolute flux limit, even for $f_{\mathrm{Edd, min}} = 10^{-3}$. The results are more sensitive to the assumed flux limit than in the fiducial models, but only moderately so.

We also consider how our results depend on the choice of $f_{\mathrm{riaf}} = 0.05$, the critical Eddington ratio at which accretion transitions to a radiatively inefficient regime. If this regime is ignored entirely, such that a constant radiative efficiency of 0.1 is used at all times, the maximum spatially offset AGN lifetime is ~90 Myr. Similar to the $f_{\mathrm{Edd, min}} = 10^{-3}$ model, these longer lifetimes increase the number of predicted observable recoils by up to a factor of a few for most models, and by more than an order of magnitude for seeing-limited observations in the aligned spin models.

Note that assumption of a time-dependent accretion rate is especially important in these models, owing to their longer AGN lifetimes. For both the $f_{\mathrm{Edd, min}} = 10^{-3}$ and radiatively efficient models, assuming a constant accretion rate with AGN lifetime $M_{\mathrm{disc}}/\dot{M}_{\mathrm{BH}}$ reduces the predicted number of spatially offset AGN by a factor of a few to a few tens in most cases. Velocity offset AGN lifetimes are also generally shorter by a factor of a few if a constant accretion rate is assumed.

Finally, the α viscosity parameter is also important in determining the accretion disc lifetime. Our fiducial model assumes $\alpha = 0.3$, motivated by models of self-gravitating accretion discs (e.g. Goodman 2003). The nature of the viscous accretion flow is uncertain, however, and α is commonly assumed to be lower. Accordingly, we have also tested a model with $\alpha = 0.03$. Fig. 16 shows the resulting predictions for the number of observable spatially offset AGN per deg^2, in the same manner as Fig. 13. The lower value of α yields longer offset AGN lifetimes and higher average luminosities. The predicted number of observable recoils is higher by a factor of a few to $\gtrsim 10$ than in the fiducial model, a slightly larger enhancement than seen in the $f_{\mathrm{Edd, min}} = 10^{-3}$ or radiatively efficient models. Again, predictions for seeing-limited observations in the aligned spin models are most sensitive to the change in accretion disc parameters.

Overall, we find that our results depend on the accretion model for recoiling AGN, but because the fiducial model is designed to be conservative, variations in the accretion model parameters tend to *increase* the observability of recoils. Moreover, the aligned spin models are the most sensitive to changes in the accretion rate; for the random and hybrid models, the alternate accretion models tested here generally increase the predicted number of observable offset AGN by factors of a few at most.
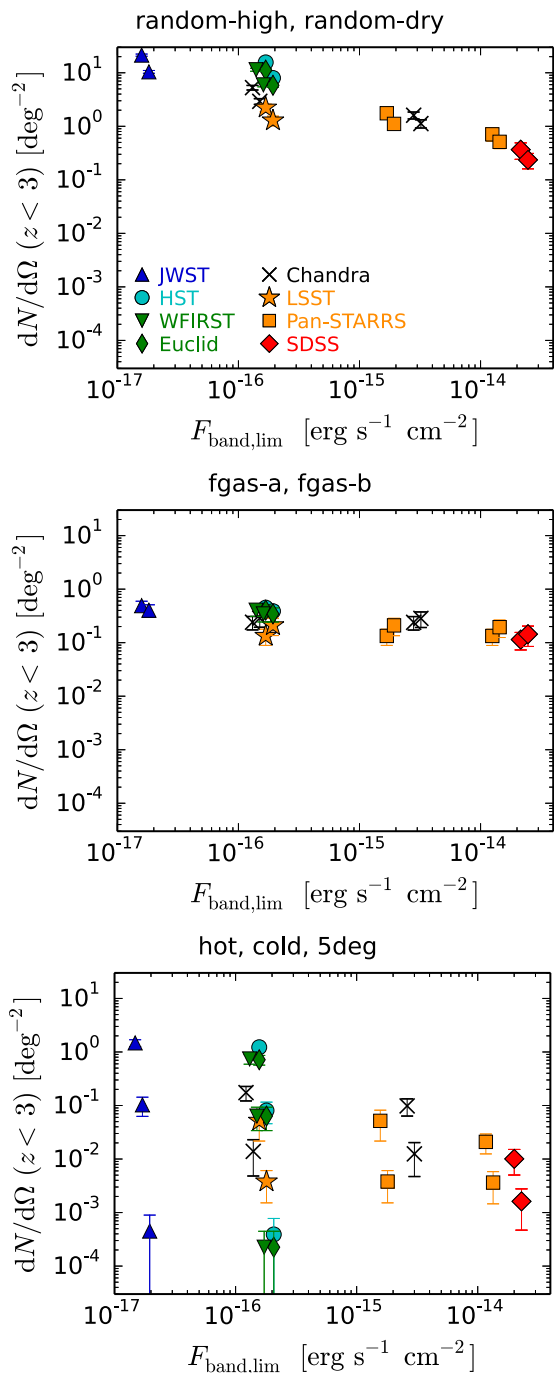
Figure 16. Same as the left column of Fig. 13, but for an α viscosity parameter of 0.03. The predicted rates of observable spatially offset AGN are generally a factor of a few higher than in the fiducial model, but in the aligned spin models the predictions for seeing-limited observations increase by more than an order of magnitude. Similar enhancements relative to the fiducial model are seen for the $f_{\rm Edd,\,min} = 10^{-3}$ and radiatively efficient models.

3.7.2 Circumnuclear gas disc

Our fiducial galaxy models include a compact circumnuclear gas disc with a mass equal to the total mass in cold, star-forming gas in the progenitor galaxies. This is well motivated by the correlation of galaxy mergers with both nuclear starbursts and AGN fuelling,

which is observed empirically and in simulations. Simulations have also shown that the steep central potentials created during gas-rich mergers can greatly suppress recoil trajectories (Blecha et al. 2011; Sijacki et al. 2011). As discussed in Section 3.6, the gas component included in our galaxy models does indeed reduce off-set AGN lifetimes in gas-rich galaxies, such that recoils are more likely to be detected in merger remnants with relatively low gas fractions.

However, the simplified nature of these potential models cannot fully capture the complexity of a galaxy merger remnant. In particular, while recoiling BH orbits will generally be centrophilic, irregularities in the potential will prevent them from returning exactly to the galaxy centre, thus delaying their orbital decay. We have also neglected the consumption of cold gas by star formation in the interval between the simulation snapshot time (when the progenitor galaxy properties are obtained) and the BH merger time, as well as during the recoil event. The amount of cold gas present in the merger remnant model may therefore be systematically overestimated. In order to understand how these assumptions regarding the gas disc model influence our results, we consider the case in which the gas disc is neglected entirely in the galaxy potential model. (The host gas fraction is still used to determine pre-merger BH spin alignment in the hybrid spin models.)

Naturally, galactic escape speeds are lower in the absence of a gas disc component. For the random spin models, the escape fractions are significant: 20 and 30 per cent of recoil events for the `random-dry` and `random-high` models, respectively. Moreover, even when scaling the kick speed to the escape speed, return times (and maximum displacements) are larger for a given $v_{\rm k}/v_{\rm esc}$ in the model without gas. Recoils with $v_{\rm k}/v_{\rm esc} \sim 0.7$ can displace the BH from the galactic centre for up to 1 Gyr, a factor of a few longer than in the fiducial model (Fig. 5). Recoils with low to moderate $v_{\rm k}/v_{\rm esc}$ ($\lesssim 0.6$) are most strongly affected by the shape of the inner density profile. When a dense gas component is included, some BHs with $v_{\rm k}/v_{\rm esc} < 0.6$ have return times $<10^5$ yr (Fig. 5), but in the model without a gas disc, such events are rare even in the aligned spin models.

As a result, offset AGN lifetimes are longer in the model with no gas disc, and a larger fraction of recoil events produce resolvable spatial offsets. Fig. 17 shows the same distributions of offset AGN properties as in Fig. 11, for the `5deg` spin model and *HST* resolution and sensitivity, but for the model with no gas disc. The longer return times for moderate-velocity recoils create a population of flux-limited offset AGN at $\Delta R_{\rm proj} \lesssim 0.5$ arcsec that is largely absent in the fiducial model. These offset AGN would be resolvable with *HST*, *JWST*, *Euclid*, and WFIRST; the no-disc model predicts that up to 14 spatially offset AGN per deg^2 could be found in *HST*-COSMOS, and 7 per deg^2 with *Euclid* and WFIRST. Even the cold spin model predicts that one to two spatially offset AGN could be detected in *HST*-COSMOS. For coarser angular resolutions (>0.5 arcsec), though, the predicted numbers of spatially offset AGN are closer to the fiducial model predictions, because these offsets mostly result from larger kicks that are less strongly affected by the inner potential.

The offset AGN that are observable in the no-disc model but not the fiducial model are almost exclusively hosted in gas-rich mergers ($f_{\rm gas,\,sf} > 0.3$), and most occur at high redshift ($z > 1$). Because gas-rich galaxies have lower masses, on average, low-mass hosts are more heavily favoured in the no-disc model. Also, in contrast to the fiducial model, there is no strong preference in the no-disc model for offset AGN to be hosted in gas-poor galaxies.
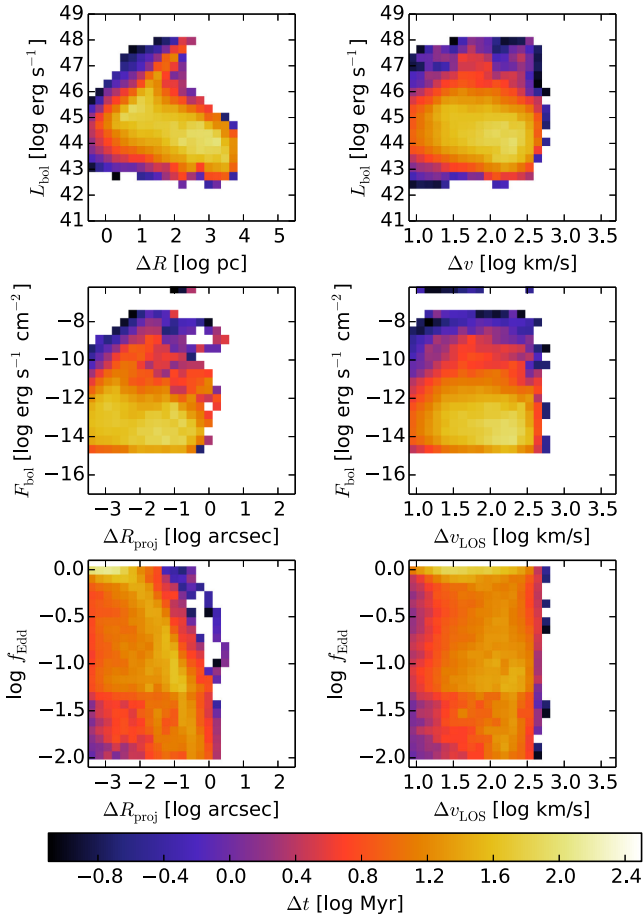
Figure 17. Same as Fig. 11, but for a model in which no gas disc is included in the merger remnant potential. The `5deg` spin model is shown. Recoil events with $v_k/v_{esc} \lesssim 0.6$, which dominate the offset AGN population in this model, are most affected by the absence of a gas disc component.

3.7.3 Gas-rich merger definition

We distinguish between gas-rich and gas-poor galaxy mergers to determine which BH binaries are likely to have an ample reservoir of cold gas capable of aligning the BH spins. Gas-poor mergers are also assumed to require a longer time-scale for the BHs to merge, though this has little effect on our results relative to the mass ratio dependence of our merger time-scale prescription. We have adopted a conservative definition of gas-rich mergers as those with a star-forming gas fraction $f_{gas, sf} > 0.1$. According to this definition, more than 85 per cent of mergers at $z < 1$, and nearly all of those at $z > 1$, are gas-rich. In the hybrid spin models, a less stringent definition of gas-rich mergers would result in fewer aligned spins and a greater number of superkicks.

We test this assertion by considering a model with a critical gas fraction of 0.2, for which $\sim$75 per cent of galaxies at $z < 1$ are defined as gas-rich. For the random and aligned spin models, where the critical gas fraction enters only into the merger delay time-scale, the results do not change significantly. The predicted numbers of observable recoils differ by a few per cent from the fiducial model. For the hybrid spin models, the model with critical $f_{gas, sf} = 0.2$ predicts a factor of 2–3 more observable recoiling AGN (for spatial and velocity offsets) than the fiducial model. The preference for high-mass host galaxies also largely disappears, because greater numbers of low-mass galaxies are defined as gas-poor, and thus can produce high-velocity kicks.

In reality, BH spin alignment should depend on the mass (and angular momentum) of gas in a circumnuclear disc around the BHs, rather than the total gas mass in the galaxy. However, because it is unclear what fraction of the cold gas will condense into a coherent disc during the merger, we have opted to tie spin alignment to the global character of the merger (gas-rich or gas-poor). For completeness, we have also tested a model in which the total star-forming gas mass is compared to the total BH mass, with gas-rich galaxies defined as those with $M_{gas, sf}/M_{BH} > 10$ (such that $M_{disc} \gtrsim M_{BH}$ if $\gtrsim 10$ per cent of the total cold gas mass is condensed into a circumnuclear disc). In this case, >95 per cent of mergers are gas-rich at $z < 1$, indicating that this is an even more conservative definition of gas-rich mergers than our fiducial model.

Again, for random and aligned spins, there is negligible difference between this model and the fiducial case. For hybrid spin models, despite the small fraction of mergers that do not have aligned BH spins, only a factor of $\sim$2 fewer offset AGN are predicted with this definition of gas-rich mergers. We also note that, while offset AGN tend to have high BH masses in the hybrid spin models regardless, this effect is even stronger when spin alignment is tied to the BH mass.

3.7.4 BH merger time-scales

As discussed in Section 2.1, a delay is added to the time of each BH merger in our model, because BH binary evolution cannot be modelled on subgrid scales, and because minor-merger time-scales in particular may be underestimated by the BH prescriptions in Illustris. If we examine the results without this delay time-scale added, we find relatively minor differences; the delay time-scale lowers the predicted numbers of offset AGN by less than a factor of 2 in most cases. In part, this owes to the small fraction of *major* mergers for which the BH binary is still unmerged at $z = 0$ according to this prescription (2.5 per cent of mergers with log $q > -1.5$). Also, the unmerged binaries that are removed from our sample are balanced to some extent by higher redshift binaries that have long delay time-scales, such that the corresponding recoil events occur at lower redshift, where they are more easily observable.

As shown in Fig. 2, the delay time-scale shifts the peak of BH merger activity to slightly lower redshift, from $z \sim 2$ to $z \sim 1.5$. Although this does not produce a large effect on the observability of recoils, these longer merging time-scales may have a more significant impact on GW signals from BH mergers. The inspiral time-scales of BH binaries on subgrid scales, and the resulting GW signal, will be explored in detail in Kelley et al. (in preparation).

4 DISCUSSION

4.1 Detection and confirmation of recoiling agn

Our models predict the numbers of spatially or kinematically offset AGN that are (a) resolvable and (b) above the sensitivity limit for a given observational survey. Once a candidate recoiling AGN is identified via a spatial or kinematic offset, however, extensive multiwavelength follow-up observations may be required to rule out alternate scenarios. Alternative explanations for existing recoil candidates include dual or binary BHs in which only one BH is active (Komossa et al. 2008; Bogdanović, Eracleous & Sigurdsson 2009;

Comerford et al. 2009; Dotti et al. 2009; Batcheldor et al. 2010; Robinson et al. 2010; Tsalmantza et al. 2011; Eracleous et al. 2012; Lena et al. 2014), extreme double-peaked emitters (Shields et al. 2009b; Tsalmantza et al. 2011; Eracleous et al. 2012; Steinhardt et al. 2012), chance superposition (Heckman et al. 2009; Shields, Bonning & Salviander 2009a), or unusual SNe (Koss et al. 2014). Interestingly, Decarli et al. (2014) find that neither the recoil scenario nor other proposed explanations for the velocity offset quasar J0927+2943 are consistent with recent observations.

Type IIn SNe may resemble AGN for months or years after the explosion; they commonly have both BLs and NLs and slow decay times owing to interactions of shocked ejecta with circumstellar material. However, as their spectral features evolve and they eventually fade, such events should be distinguishable from AGN via continued monitoring and multiwavelength observations.

Extreme double-peaked emitters in which only one peak is visible may also resemble recoiling AGN, in that they can produce systemic velocity shifts in BLs (Chen & Halpern 1989; Eracleous et al. 1995). Different BLs often have very different velocity shifts in such cases, while to first order, recoiling AGN should exhibit a consistent shift in all broad lines (owing to the bulk motion of the AGN). Similarly, the orbital motion of sub-parsec binary BHs in which only one BH is active may give rise to velocity-shifted BLs. In this case, the velocity shift will evolve on the orbital time-scale, while recoiling AGN should exhibit a constant velocity shift.

One of the more challenging alternative scenarios to rule out for candidate recoils is that of a well-separated (∼kpc scale) BH *pair* in which only one BH is actively accreting. This is the case for CID-42, which has two distinct nuclei separated by 2.5 kpc, one of which is a confirmed AGN with a BL offset of 1300 km s^{-1} (Civano et al. 2010, 2012; Novak et al. 2015). Note that the magnitude of this spatial and velocity offset is quite consistent with the predictions of our models, if spins are not highly aligned (Fig. 9). The offset BL and lack of evidence for an AGN in the second nucleus support the recoil scenario, and indirect arguments suggest that the host galaxy properties are more consistent with a post-merger phase, but it is difficult to rule out the possibility that a *quiescent* BH is harboured in the second nucleus (Blecha et al. 2013).

An additional challenge in confirming spatially offset AGN is that recoil events occur, by definition, in recently merged galaxies. The host may therefore be highly disturbed, such that identifying the galaxy centroid is nontrivial. In the cases where the BH binary inspiral time may be long, for example in gas-poor mergers, the host may have a more relaxed morphology by the time the BH merger and recoil occurs. However, we have demonstrated that if spins are not highly aligned, offsets larger than a kpc should be common. In such cases, it is not necessary to pinpoint the host centroid to great accuracy in order to determine that the AGN is offset. This is crucial for any programme to search for offset AGN in ground-based, all-sky surveys, where galactic morphological features often will not be resolved in great detail. Similarly, because observable recoiling AGN also tend to have large velocity offsets (≳500–1000 km s^{-1}), any ambiguity in the rest-frame redshift of the disturbed host (owing to residual motion of the NL gas, for example) is unlikely to preclude the detection of such offsets.

As noted earlier, our models do not account for the possibility that a recoiling BH may encounter fresh fuel as it settles back to the galactic centre, which would produce recoiling AGN with much smaller spatial offsets. This could be especially relevant for recoils in dry merger remnants, where the BH may undergo long-lived, small-amplitude oscillations in a core potential (Gualandris & Merritt 2008; Lena et al. 2014). Regardless, our results sug-

gest that if BH spins are not highly aligned, a population of spatially offset recoiling AGN may be resolvable with seeing-limited observations.

The most unambiguous recoiling AGN candidates will likely be those with both spatial and velocity offsets, as in CID-42, or with a projected spatial offset large enough that the AGN is well separated from the stellar light of the host galaxy. Our results indicate that a majority of velocity offset AGN should also have spatial offsets resolvable with *HST*, and 5–20 per cent of spatially offset AGN resolvable with ground-based observations should also have LOS velocity offsets >600 km s^{-1}. Additionally, up to 50 per cent of spatially offset AGN resolvable with the ground-based imaging (and up to 25 per cent of those resolvable with *HST*) should appear at radii beyond the host stellar bulge. Thus, the possibility exists of finding hundreds of such objects with Pan-STARRS and LSST, if spins are not strongly aligned.

Recoiling AGN with large spatial offsets also tend to have low Eddington ratios, as their accretion rate declines over time. We assume that accretion transitions to a radiatively inefficient regime below $f_{Edd} = 0.05$; this regime may often be associated with radio emission in AGN. Thus, high radio-to-optical flux ratios in offset AGN, or even small-scale radio jets, could be another good indication of a past recoil event.

We note that the hybrid and aligned spin models predict few recoils to be detectable at $z > 1$. For the random spin models, this is not the case; velocity offset AGN in particular may be found in equal or greater numbers at $z > 1$ as at lower redshift (Fig. 12). This raises the interesting possibility of detecting recoiling AGN near the cosmic peak of quasar and galaxy merging activity, and such detections would place additional indirect constraints on pre-merger BH spins. Evidence of past BH mergers at high redshift would also place stronger limits on BH merging time-scales than could be obtained in the local Universe. However, if such events are observable, they would be more challenging to confirm with follow-up observations than their low-redshift counterparts.

Our models provide only analytic approximations to the properties and structure of the host galaxies of recoiling AGN, which should be highly dynamic during major mergers. The total stellar mass and star-forming gas content are readily obtained from the progenitors, however, and we can identify trends in these properties of offset AGN hosts. Specifically, although the host stellar mass distribution is broad, stellar masses of log ($M_*/M_\odot$) ∼ 10–11 are favoured in the random and aligned spin models. In the hybrid spin models (for which spin alignment occurs only in gas-rich mergers), offset AGN occur preferentially in higher mass hosts, log ($M_*/M_\odot$) ≳11. This is a testable prediction, but a statistical sample of offset AGN would be required to differentiate between these distinct but overlapping distributions.

We also find that merger remnants with low to moderate star-forming gas fractions are preferred as offset AGN hosts, as recoil trajectories tend to be quickly suppressed in gas-rich mergers. This is especially true of the hybrid spin models, where almost no observable recoils occur in gas-rich merger remnants. We suggest that a systematic search for offset AGN might focus initially on bulge-dominated or disturbed galaxies with low to moderate star formation rates.

Again, we caution that our sample is incomplete at the low-mass end owing to resolution limits, such that offset AGN in low-mass or dwarf hosts are largely excluded from our analysis. This is an important caveat; although the merger rate and BH occupation fraction of dwarf galaxies are poorly constrained, the low escape speeds of low-mass galaxies should result in resolvable offsets in

many if not most recoil events, even in galaxies that are gas-rich. The recoil candidate SDSS 1133 is one possible example of this (Koss et al. 2014). Low-mass merger remnants may therefore host a population of spatially offset AGN in addition to those predicted by our models.

As a final caveat, our results assume that recoiling AGN generally resemble 'normal', stationary AGN, in that they retain an accretion disc and broad line region and have an SED that is well modelled with standard AGN templates. This seems a reasonable first-order assumption, given that the region producing most of the AGN emission will typically have orbital speeds $v_{orb} \gg v_k$ and will feel the recoil event as a negligible perturbation. None the less, the possibility that recoiling AGN appear qualitatively different from stationary AGN cannot be discounted, and until recoil events can be unambiguously confirmed, this remains an uncertainty for any search for such objects.

The detection of even one recoiling AGN with a large velocity offset ($\gtrsim 1000$ km s^{-1}) or spatial offset ($\gtrsim 10$ kpc), or alternatively the discovery of a massive galaxy with no central BH, would strongly indicate that binary BH spins are not efficiently aligned in all cases. Our results suggest that up to hundreds or thousands of spatially offset AGN may be detectable with Pan-STARRS and LSST. While these numbers are dependent on model assumptions, as discussed in Section 3.7, we have attempted to be conservative in constructing the fiducial models. If such a population of recoiling AGN is identified, then further distinctions between BH spin and recoil velocity distributions may be possible. If no convincing recoil candidates are found in large-area surveys, the simplest explanation may be that spin alignment is effective for most binary BHs, or that spin magnitudes are low. Constraints on the efficiency of spin alignment could also provide indirect information about the nature of the accretion flow, specifically whether it is likely to be coherent on time-scales required to align the BH spins, or if it is more often characterized by chaotic, randomly oriented streams.

4.2 Comparison with previous work

Our predicted rates of observable offset AGN are substantially higher than those of VM08, the only previous study of recoiling AGN in a cosmological framework. Here, we identify some key differences that contribute to this discrepancy.

VM08 used semi-analytic merger tree models to determine the BH merger rate and host properties, similarly integrating recoil trajectories for each event in a galaxy potential model. Our predictions for the number of observable spatially offset AGN per deg^2 in the `random-high` spin model (Fig. 13) can be compared with their results for *HST*, *JWST*, and *Chandra*. For our fiducial models, these numbers are comparable to the predictions of the VM08 'halo-only' models, and they are a factor of > 10–100 above their 'halo+bulge' predictions. Given that our fiducial model includes not only a halo and compact stellar bulge but also a dense gas disc, the recoiling BH return times are much shorter than in a halo-only model. Thus, our model predictions are in fact significantly higher than those of VM08.

Neglecting the gas component in our models increases the discrepancy with the VM08 results for a halo+bulge potential, up to a factor of several hundred for *HST* and *JWST*. If we include only a DM halo potential, ~ 15–25 times more spatially offset AGN are predicted than in our fiducial model.

The recoil kick velocity distribution is sensitive to the BH mass ratio distribution, but the kick formula of Baker et al. (2008) used

in VM08 depends even more sensitively on q than is predicted in more recent work (e.g. van Meter et al. 2010; Lousto et al. 2012). Specifically, the Baker et al. (2008) formula scales with η^3, where $\eta \equiv q/(1 + q)^2$ is the symmetric mass ratio, while the more recent formulae include an η^2 term. As a result, the Baker et al. (2008) formula produces fewer high-velocity recoils, and the steeper dependence on q amplifies discrepancies in the underlying distributions. Using this formula in our models lowers our predictions by a factor of ~ 2.

The ejected accretion disc model of Loeb (2007) used in VM08 has a steep dependence on BH mass ($\propto M_{BH}^{2.2}$). Thus, the disc must be capped at $M_{disc} = M_{BH}$ in most cases and depends sensitively on the underlying BH mass distribution. Our fiducial models instead assume the disc model of Blecha & Loeb (2008), which accounts for disc self-gravity and yields typical disc masses of a few per cent of the BH mass. Furthermore, the accretion rate in our models is assumed to decline over time as the disc diffuses outwards. This increases the predicted number of spatially offset AGN by a factor of a few relative to a constant-$\dot{M}$ model.

If we attempt to compare as closely as possible with the results of VM08, using the same potential models, recoil kick formula, and ejected accretion disc model, we still predict spatially offset AGN to be observable at rates that are tens to hundreds of times higher than in any of their models. The discrepancy is larger when comparing to the halo+bulge models than the halo-only case, presumably because the bulge mass is directly scaled to the BH mass and thus is sensitive to differences in the underlying distribution. Owing to the significant differences in numerical techniques and prescriptions for BH seeding, growth, and feedback between the semi-analytic models and cosmological simulations, it is not surprising that the resulting BH merger and recoil properties are qualitatively different. The offset AGN population has a non-trivial dependence on the details of each method, but none the less, a few important differences can be pinpointed.

First, as noted in Section 3.1, the BH mass ratio distribution in Illustris is qualitatively different from that shown in VM08, the latter having a sharp deficit of mergers with $q > 0.3$. It is suggested that this deficit owes to the minimum mass threshold ($M_{tot} > 10^5$ M$_\odot$) imposed, indicating that the BH mass distribution differs qualitatively from that in Illustris (at least for merging BHs), with lower mass BHs favoured in the latter. VM08 also show for comparison a q distribution derived from Monte Carlo sampling of their $z = 0$ BH mass function, which is in much better qualitative agreement with the Illustris result. We can also compare our results with the semi-analytic model results of Barausse (2012); they find fewer $q > 0.3$ BH mergers than in Illustris as well, but the discrepancy is much smaller. If we consider only recoil events with $q < 0.3$ in our models, the rates of observable recoils are reduced by a factor of ~ 5–6.

While it is possible that $q \sim 1$ mergers are over-represented in Illustris, owing to the uncertainties in merger time-scales discussed in Section 2.1, the inspiral times are most likely to be underestimated for *minor* mergers. We have compensated for this by imposing a minimum BH mass of 10^6 M$_\odot$ and a merger delay time-scale that scales with q.

If the merging BH mass distribution is also different in VM08 than in Illustris – specifically if lower masses are favoured in the former, this would contribute to the discrepant results as well. The mass distribution is not explicitly given in VM08, but the reasonable agreement of the integrated merger rate for $M_{BH} > 10^{6-7}$ M$_\odot$ between Illustris and the other semi-analytic models (Sesana et al. 2004; Barausse 2012) suggests that the merging BH mass

distribution in Illustris does not differ too greatly from those studies, at least at the high-mass end and integrated over cosmic time.

We again note that any differences in the mass and mass ratio distributions between Illustris and the models of VM08 are amplified by the sensitive dependence of the Baker et al. (2008) kick formula and the Loeb (2007) disc model to these parameters. However, these factors alone do not appear to account for all of the difference between our comparison model and the results of VM08.

The cumulative BH merger rate for $M_{BH} > 10^6$ M$_\odot$ agrees well between Illustris and semi-analytic models, and the rate peaks at similar redshift (Sesana et al. 2004; Barausse 2012). However, as noted in Section 3.1, there are fewer high-redshift mergers and more low-redshift mergers in Illustris, owing to the different BH seed formation, growth, and merger prescriptions. Differences in the redshift distribution of mergers are likely a significant contributor to the higher rates of observable recoils predicted in this work. Empirical constraints on the BH merger rate are insufficient to distinguish between these models, but the agreement of the Illustris galaxy merger rate with observations is encouraging (Rodriguez-Gomez et al. 2015), as is the agreement with the observed BH mass function and quasar luminosity function (Sijacki et al. 2015).

5 SUMMARY

We have developed models for AGN that are spatially or kinematically offset from their host galaxies owing to GW recoil kicks following a major merger. Using data from the Illustris cosmological simulations, we determine the characteristics of recoiling AGN and their host galaxies over cosmic time. Because recoil events may be preferentially suppressed in gas-rich mergers, the use of hydrodynamic simulations is critical for determining which recoiling AGN may be observable. Each BH merger in our sample is assigned (a) a recoil kick velocity based on the BH mass ratio and assumed spin distribution, (b) an accretion disc that remains bound to the recoiling BH, and (c) a merged host galaxy model based on the progenitor galaxy properties. The recoil trajectory is integrated in this galaxy potential, and the AGN luminosity is calculated at each timestep.

This study is the first to consider hydrodynamic effects on recoiling BHs in a cosmological framework, and the first to examine the effect of pre-merger BH spin alignment on the observability of recoils. Mergers between BHs with aligned or nearly aligned spins produce dramatically lower recoil velocities than do randomly oriented spins, and misaligned spins are required to produce superkicks that can eject BHs from galaxies entirely. BH spin alignment is predicted to occur via interaction with a circumbinary gas disc, though the degree of alignment is not well constrained and likely depends on the properties of the gas disc. We therefore consider three categories of BH spin models: those with randomly oriented spins, those in which BHs always undergo some amount of alignment, and 'hybrid' models in which spin alignment occurs only in gas-rich mergers.

Our main results can be summarized as follows.

(i) Recoiling AGN do *not* spend most of their time at small spatial and velocity offsets, unless pre-merger BH spins are highly aligned. The longest-lived offset AGN are produced by recoils near the host escape speed. This occurs because marginally bound and escaping BHs experience the least deceleration as they leave the host nucleus, and because marginally bound recoiling BHs spend most of their time at large apocentres. Recoil events with $v_k \gg v_{esc}$

are intrinsically rare and have short AGN lifetimes because little gas remains bound to the recoiling BH.

(ii) Spatially offset AGN are most commonly found at projected separations >0.1 arcsec from the host galaxy, with a tail extending beyond 10 arcsec (if BH spins are not strongly aligned). This indicates that even seeing-limited observations can resolve spatially offset AGN, making them promising targets for current and future wide-area surveys such as SDSS, Pan-STARRS, LSST, *Euclid*, and WFIRST. These spatial offsets often correspond to physical projected separations >1 kpc, such that they may be detectable even if the host galaxy is disturbed and its centroid ill-defined.

(iii) If BH spins are randomly oriented, our models predict that nearly 10 spatially offset AGN may be detectable in *HST*-COSMOS. The Pan-STARRS1 Medium Deep Survey could find ∼15–20 offset AGN, and hundreds could be found in SDSS DR7. The all-sky surveys LSST and PS1-3pi could detect up to several thousand spatially offset AGN, as could a wide-area WFIRST survey. An all-sky survey with Euclid could detect up to several tens of thousands of such objects.

(iv) The random spin models also predict hundreds of velocity offset AGN with $\Delta v_{LOS} > 1000$ km s^{-1} in the SDSS DR7 footprint. At face value, this is in tension with the findings of Bonning et al. (2007), Tsalmantza et al. (2011), and Eracleous et al. (2012) that large BL offsets are rare among SDSS quasars. Many of the offset AGN in our sample may be low-luminosity or obscured AGN excluded from the SDSS quasar catalogue, but this may also be a preliminary indication that pre-merger spin alignment is effective in at least some mergers, or that BH spins are low.

(v) A population of recoiling AGN may be detected in current and future surveys even if some pre-merger BH spin alignment occurs. If spin alignment is *always* very efficient (within ∼ 5° of the orbital angular momentum), recoils are unlikely to ever be detected. But if only moderate spin alignment occurs, up to a few tens of spatially offset AGN could be detected with LSST or Pan-STARRS, and if spins are only aligned in gas-rich mergers, hundreds of recoiling AGN may be observable in such surveys. In the latter case, over 2000 recoiling AGN could be detected with Euclid.

(vi) For comparable assumptions regarding the recoil velocity distribution and host galaxy properties, we find a much higher incidence of observable recoiling AGN than a previous study (VM08). We argue that this discrepancy owes primarily to systematic differences in the underlying redshift, mass and mass ratio distributions of merging BHs in Illustris versus the semi-analytic models, as well as the use of more physically motivated models for recoiling AGN accretion discs and an updated recoil kick formula in this work.

(vii) A *majority* of recoiling AGN with resolvable LOS velocity offsets may have simultaneous spatial offsets resolvable with *HST*, and ∼4–20 per cent of velocity offset AGN detectable with SDSS may also have spatial offsets resolvable with SDSS. This indicates that high-resolution imaging could be key for confirming any velocity offset recoil candidates.

(viii) A smaller but non-negligible fraction of spatially offset AGN (<10 per cent) should simultaneously have LOS velocity offsets $\Delta v > 1000$ km s^{-1}. We also find that up to 50 per cent of spatially offset AGN resolvable with seeing-limited observations (and up to 25 per cent of those resolvable with *HST*) should be at projected separations larger than the stellar bulge radius, such that they could be distinguished from inspiralling dual BHs. AGN exhibiting simultaneous spatial and velocity offsets, or that are well separated from the host galaxy stellar light, or both, would be among the strongest candidate recoils.

(ix) The preferred stellar masses of offset AGN hosts differ depending on the degree of pre-merger BH spin alignment. If the spins are always drawn from the same distribution (random or aligned), the host stellar mass distribution peaks around log $(M_*/M_\odot) =$ 10–11. For the hybrid spin models, where the degree of BH spin alignment depends on the host gas content, *high-mass* hosts are preferred despite their larger escape speeds, because they have lower gas fractions on average and are more likely to yield superkicks from misaligned BH spins. These models also yield very few observable recoils at $z > 1$, owing to the gas-rich nature of high-redshift galaxies.

(x) In all cases, offset AGN are most likely to be found in hosts with low to moderate gas fractions, owing to the dynamical suppression of recoil trajectories by dense circumnuclear gas in gas-rich systems. The longest-lived offset AGN occur preferentially in galaxies with low gas fractions.

Our quantitative predictions for the observability of recoiling AGN have some dependence on model assumptions, including the host galaxy potential model and accretion disc parameters. As discussed in Section 3.7, variation with these parameters is often substantially larger than the statistical errors shown in Fig. 13. Because the fiducial model is conservative by design, reasonable variations in model parameters tend to *increase* the predicted number of observable offset AGN. However, the results also depend on the underlying distributions of BH mass, accretion rate, and merger rate in Illustris. Owing to uncertainties in the subgrid BH prescriptions, we cannot exclude the possibility of systematic bias in the merging BH population, but the success of Illustris in reproducing key observables of the BH population (Sijacki et al. 2015) and the reasonable level of convergence of our results (Appendix A) are encouraging.

These results strongly motivate a systematic search for offset AGN in current and future large-area surveys. Discoveries of GW recoils would confirm this prediction of strong-field gravity, and they would provide robust evidence of recent BH mergers, thereby constraining the time-scales for BH binary inspiral. We demonstrate that observations of recoiling AGN could also place indirect constraints on the efficiency of pre-merger spin alignment, which in turn could provide new evidence regarding the nature of the BH accretion flow.

REFERENCES

Alexander D. M. et al., 2003, AJ, 126, 539
Amaro-Seoane P. et al., 2012, Class. Quantum Grav., 29, 124016
Baker J. G., Boggs W. D., Centrella J., Kelly B. J., McWilliams S. T., Miller M. C., van Meter J. R., 2008, ApJ, 682, L29
Barausse E., 2012, MNRAS, 423, 2533
Barnes J. E., Hernquist L. E., 1991, ApJ, 370, L65
Barnes J. E., Hernquist L., 1996, ApJ, 471, 115
Batcheldor D., Robinson A., Axon D. J., Perlman E. S., Merritt D., 2010, ApJ, 717, L6
Bekenstein J. D., 1973, ApJ, 183, 657
Berczik P., Merritt D., Spurzem R., Bischof H.-P., 2006, ApJ, 642, L21
Berti E., Kesden M., Sperhake U., 2012, Phys. Rev. D, 85, 124049
Blecha L., Loeb A., 2008, MNRAS, 390, 1311
Blecha L., Cox T. J., Loeb A., Hernquist L., 2011, MNRAS, 412, 2154
Blecha L., Civano F., Elvis M., Loeb A., 2013, MNRAS, 428, 1341
Bogdanović T., Reynolds C. S., Miller M. C., 2007, ApJ, 661, L147
Bogdanović T., Eracleous M., Sigurdsson S., 2009, ApJ, 697, 288
Bonning E. W., Shields G. A., Salviander S., 2007, ApJ, 666, L13
Bullock J. S., Kolatt T. S., Sigad Y., Somerville R. S., Kravtsov A. V., Klypin A. A., Primack J. R., Dekel A., 2001, MNRAS, 321, 559
Campanelli M., Lousto C. O., Zlochower Y., Merritt D., 2007, Phys. Rev. Lett., 98, 231102
Cannizzo J. K., Lee H. M., Goodman J., 1990, ApJ, 351, 38
Chandrasekhar S., 1943, ApJ, 97, 255
Chen K., Halpern J. P., 1989, ApJ, 344, 115
Cisternas M. et al., 2011, ApJ, 726, 57
Civano F. et al., 2010, ApJ, 717, 209
Civano F. et al., 2012, ApJ, 752, 49
Civano F. M. et al., 2013, in American Astronomical Society Meeting Abstracts, Vol. 221, p. 352.25
Comerford J. M., Griffith R. L., Gerke B. F., Cooper M. C., Newman J. A., Davis M., Stern D., 2009, ApJ, 702, L82
Decarli R., Dotti M., Mazzucchelli C., Montuori C., Volonteri M., 2014, MNRAS, 445, 1558
Di Matteo T., Springel V., Hernquist L., 2005, Nature, 433, 604
Dotti M., Colpi M., Haardt F., Mayer L., 2007, MNRAS, 379, 956
Dotti M., Montuori C., Decarli R., Volonteri M., Colpi M., Haardt F., 2009, MNRAS, 398, L73
Dotti M., Volonteri M., Perego A., Colpi M., Ruszkowski M., Haardt F., 2010, MNRAS, 402, 682
Ellison S. L., Patton D. R., Mendel J. T., Scudder J. M., 2011, MNRAS, 418, 2043
Eracleous M., Livio M., Halpern J. P., Storchi-Bergmann T., 1995, ApJ, 438, 610
Eracleous M., Boroson T. A., Halpern J. P., Liu J., 2012, ApJS, 201, 23
Escala A., Larson R. B., Coppi P. S., Mardones D., 2005, ApJ, 630, 152
Farris B. D., Duffell P., MacFadyen A. I., Haiman Z., 2014, ApJ, 783, 134
Fujita Y., 2009, ApJ, 691, 1050
Genel S. et al., 2014, MNRAS, 445, 175
Gerosa D., Sesana A., 2015, MNRAS, 446, 38
Gerosa D., Kesden M., Sperhake U., Berti E., O'Shaughnessy R., 2015, Phys Rev D, 92, 064016
González J. A., Sperhake U., Brügmann B., Hannam M., Husa S., 2007, Phys. Rev. Lett., 98, 091101
Goodman J., 2003, MNRAS, 339, 937
Gualandris A., Merritt D., 2008, ApJ, 678, 780
Guedes J., Madau P., Mayer L., Callegari S., 2011, ApJ, 729, 125
Guo Q., White S. D. M., 2008, MNRAS, 384, 2
Heckman T. M., Krolik J. H., Moran S. M., Schnittman J., Gezari S., 2009, ApJ, 695, 363
Hernquist L., 1990, ApJ, 356, 359
Holley-Bockelmann K., Khan F. M., 2015, ApJ, 810, 139
Hopkins P. F., Quataert E., 2010, MNRAS, 407, 1529
Hopkins P. F., Hernquist L., Cox T. J., Di Matteo T., Robertson B., Springel V., 2006, ApJS, 163, 1

Hopkins P. F., Richards G. T., Hernquist L., 2007, ApJ, 654, 731
Hopkins P. F., Cox T. J., Kereš D., Hernquist L., 2008a, ApJS, 175, 390
Hopkins P. F., Hernquist L., Cox T. J., Dutta S. N., Rothberg B., 2008b, ApJ, 679, 156
Hopkins P. F., Cox T. J., Dutta S. N., Hernquist L., Kormendy J., Lauer T. R., 2009, ApJS, 181, 135
Hopkins P. F., Kocevski D. D., Bundy K., 2014, MNRAS, 445, 823
Kesden M., Gerosa D., O'Shaughnessy R., Berti E., Sperhake U., 2015, Phys. Rev. Lett., 114, 081103
Khan F. M., Just A., Merritt D., 2011, ApJ, 732, 89
King A., Nixon C., 2013, Class. Quantum Grav., 30, 244006
King A. R., Pringle J. E., 2006, MNRAS, 373, L90
Kocevski D. D. et al., 2012, ApJ, 744, 148
Koekemoer A. M. et al., 2007, ApJS, 172, 196
Komossa S., 2012, Adv. Astron., 2012, 364973
Komossa S., Merritt D., 2008, ApJ, 689, L89
Komossa S., Zhou H., Lu H., 2008, ApJ, 678, L81
Koss M., Mushotzky R., Veilleux S., Winter L., 2010, ApJ, 716, L125
Koss M. et al., 2014, MNRAS, 445, 515
Laureijs R. et al., 2011, preprint (arXiv:1110.3193)
Lena D., Robinson A., Marconi A., Axon D. J., Capetti A., Merritt D., Batcheldor D., 2014, ApJ, 795, 146
Li Y. et al., 2007, ApJ, 665, 187
Lodato G., Gerosa D., 2013, MNRAS, 429, L30
Lodato G., Nayakshin S., King A. R., Pringle J. E., 2009, MNRAS, 398, 1392
Loeb A., 2007, Phys. Rev. Lett., 99, 041103
Lousto C. O., Zlochower Y., 2008, Phys. Rev. D, 77, 044028
Lousto C. O., Zlochower Y., 2011, Phys. Rev. Lett., 107, 231102
Lousto C. O., Zlochower Y., Dotti M., Volonteri M., 2012, Phys. Rev. D, 85, 084015
Lusso E. et al., 2012, MNRAS, 425, 623
Macciò A. V., Dutton A. A., van den Bosch F. C., Moore B., Potter D., Stadel J., 2007, MNRAS, 378, 55
McConnell N. J., Ma C.-P., 2013, ApJ, 764, 184
Madau P., Quataert E., 2004, ApJ, 606, L17
Mayer L., Kazantzidis S., Madau P., Colpi M., Quinn T., Wadsley J., 2007, Science, 316, 1874
Merritt D., Vasiliev E., 2012, Phys. Rev. D, 86, 102002
Merritt D., Milosavljević M., Favata M., Hughes S. A., Holz D. E., 2004, ApJ, 607, L9
Merritt D., Storchi-Bergmann T., Robinson A., Batcheldor D., Axon D., Cid Fernandes R., 2006, MNRAS, 367, 1746
Mihos J. C., Hernquist L., 1994, ApJ, 437, L47
Mihos J. C., Hernquist L., 1996, ApJ, 464, 641
Miller M. C., Krolik J. H., 2013, ApJ, 774, 43
Milosavljević M., Merritt D., 2001, ApJ, 563, 34
Narayan R., McClintock J. E., 2008, New Astron. Rev., 51, 733
Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493
Nelson D. et al., 2015, Astron. Comput., 13, 12
Nixon C. J., 2012, MNRAS, 423, 2597
Novak M. et al., 2015, MNRAS, 447, 1282
Peres A., 1962, Phys. Rev., 128, 2471
Pringle J. E., 1991, MNRAS, 248, 754
Quinlan G. D., 1996, New Astron., 1, 35
Raffai P., Haiman Z., Frei Z., 2016, MNRAS, 455, 484

Richards G. T. et al., 2006, ApJS, 166, 470
Robinson A., Young S., Axon D. J., Kharb P., Smith J. E., 2010, ApJ, 717, L122
Rodriguez-Gomez V. et al., 2015, MNRAS, 449, 49
Roedig C., Sesana A., Dotti M., Cuadra J., Amaro-Seoane P., Haardt F., 2012, A&A, 545, A127
Sanders D. B., Soifer B. T., Elias J. H., Madore B. F., Matthews K., Neugebauer G., Scoville N. Z., 1988, ApJ, 325, 74
Satyapal S., Ellison S. L., McAlpine W., Hickox R. C., Patton D. R., Mendel J. T., 2014, MNRAS, 441, 1297
Schawinski K., Treister E., Urry C. M., Cardamone C. N., Simmons B., Yi S. K., 2011, ApJ, 727, L31
Schneider D. P. et al., 2010, AJ, 139, 2360
Schnittman J. D., Krolik J. H., 2015, ApJ, 806, 88
Sesana A., 2013, MNRAS, 433, L1
Sesana A., Haardt F., Madau P., Volonteri M., 2004, ApJ, 611, 623
Sesana A., Vecchio A., Colacino C. N., 2008, MNRAS, 390, 192
Sesana A., Vecchio A., Volonteri M., 2009, MNRAS, 394, 2255
Sesana A., Barausse E., Dotti M., Rossi E. M., 2014, ApJ, 794, 104
Shen Y., Kelly B. C., 2012, ApJ, 746, 169
Shields G. A., Bonning E. W., Salviander S., 2009a, ApJ, 696, 1367
Shields G. A. et al., 2009b, ApJ, 707, 936
Sijacki D., Springel V., Di Matteo T., Hernquist L., 2007, MNRAS, 380, 877
Sijacki D., Springel V., Haehnelt M. G., 2009, MNRAS, 400, 100
Sijacki D., Springel V., Haehnelt M. G., 2011, MNRAS, 414, 3656
Sijacki D., Vogelsberger M., Genel S., Springel V., Torrey P., Snyder G. F., Nelson D., Hernquist L., 2015, MNRAS, 452, 575
Snyder G. F. et al., 2015, MNRAS, 454, 1886
Spergel D. et al., 2013, preprint (arXiv:1305.5422)
Springel V., 2010, MNRAS, 401, 791
Springel V., Hernquist L., 2003, MNRAS, 339, 289
Springel V., Di Matteo T., Hernquist L., 2005, MNRAS, 361, 776
Steinhardt C. L. et al., 2012, ApJ, 759, 24
Stern D. et al., 2005, ApJ, 631, 163
Toomre A., 1964, ApJ, 139, 1217
Torrey P., Vogelsberger M., Genel S., Sijacki D., Springel V., Hernquist L., 2014, MNRAS, 438, 1985
Treister E., Schawinski K., Urry C. M., Simmons B. D., 2012, ApJ, 758, L39
Tsalmantza P., Decarli R., Dotti M., Hogg D. W., 2011, ApJ, 738, 20
Ueda Y., Akiyama M., Ohta K., Miyaji T., 2003, ApJ, 598, 886
Urrutia T., Lacy M., Becker R. H., 2008, ApJ, 674, 80
Urry C. M., Padovani P., 1995, PASP, 107, 803
van Meter J. R., Miller M. C., Baker J. G., Boggs W. D., Kelly B. J., 2010, ApJ, 719, 1427
Vasudevan R. V., Fabian A. C., 2007, MNRAS, 381, 1235
Vogelsberger M., Genel S., Sijacki D., Torrey P., Springel V., Hernquist L., 2013, MNRAS, 436, 3031
Vogelsberger M. et al., 2014a, MNRAS, 444, 1518
Vogelsberger M. et al., 2014b, Nature, 509, 177
Volonteri M., 2007, ApJ, 663, L5
Volonteri M., Madau P., 2008, ApJ, 687, L57 (VM08)
Volonteri M., Rees M. J., 2006, ApJ, 650, 669
Wyithe J. S. B., Loeb A., 2003, ApJ, 595, 614
Yu Q., 2002, MNRAS, 331, 935

APPENDIX A: RESOLUTION CONVERGENCE

The results presented in Section 3 are based on the highest resolution Illustris simulation (Illustris-1). Here we discuss the level of convergence between the three Illustris simulations (Illustris-1, Illustris-2, and Illustris-3), which differ in mass resolution. The BH merger rate for the three simulations is shown in Fig. A1 (top panel), for all BHs and divided by total BH mass bins. There is a trend towards convergence for high-mass mergers ($M_{tot} > 10^8$ $M_\odot$) and at low redshift ($z \lesssim 0.3$), but there is a clear dependence of merger rate on mass resolution, with higher merger rates in the higher resolution simulations.

As described in Section 2.3, we consider only mergers for which each progenitor has $M_{BH} > 10^6$ $M_\odot$ and is hosted in a galaxy with at least 80 stellar and 300 DM particles. Thus, it is natural to expect a lower merger rate in the lower resolution simulations, where low-mass galaxies are excluded from our analysis. The bottom panel of Fig. A1 compares the merger rate for the three Illustris simulations with a consistent *mass* cut applied to the BH host galaxies in each case, rather than a minimum particle number. Here the Illustris-3 curves are the same as in the top panel, corresponding to a minimum stellar (halo) mass of 6.4×10^9 $M_\odot$ (1.2×10^{11} $M_\odot$), and the Illustris-1 and 2 curves assume the same minimum mass. The convergence improves substantially in this case. This demonstrates that our fiducial sample from Illustris-1 has a higher BH merger rate primarily because more low-mass galaxies can be resolved.

Fig. A2 illustrates the level of resolution convergence of the predicted number of observable spatially offset AGN in our fiducial models. Again, a constant minimum progenitor halo mass is used for the comparison. In general, the convergence is quite reasonable; in most cases the differences between the simulations are within the Poisson errors. Convergence is poorest for the aligned spin models (hot and cold) in seeing-limited surveys, where recoils are unlikely to be observed in any case. The higher values for Illustris-3 in these models reflect its higher BH merger rate when a consistent minimum mass cut is applied in all simulations (Fig. A1, bottom panel).
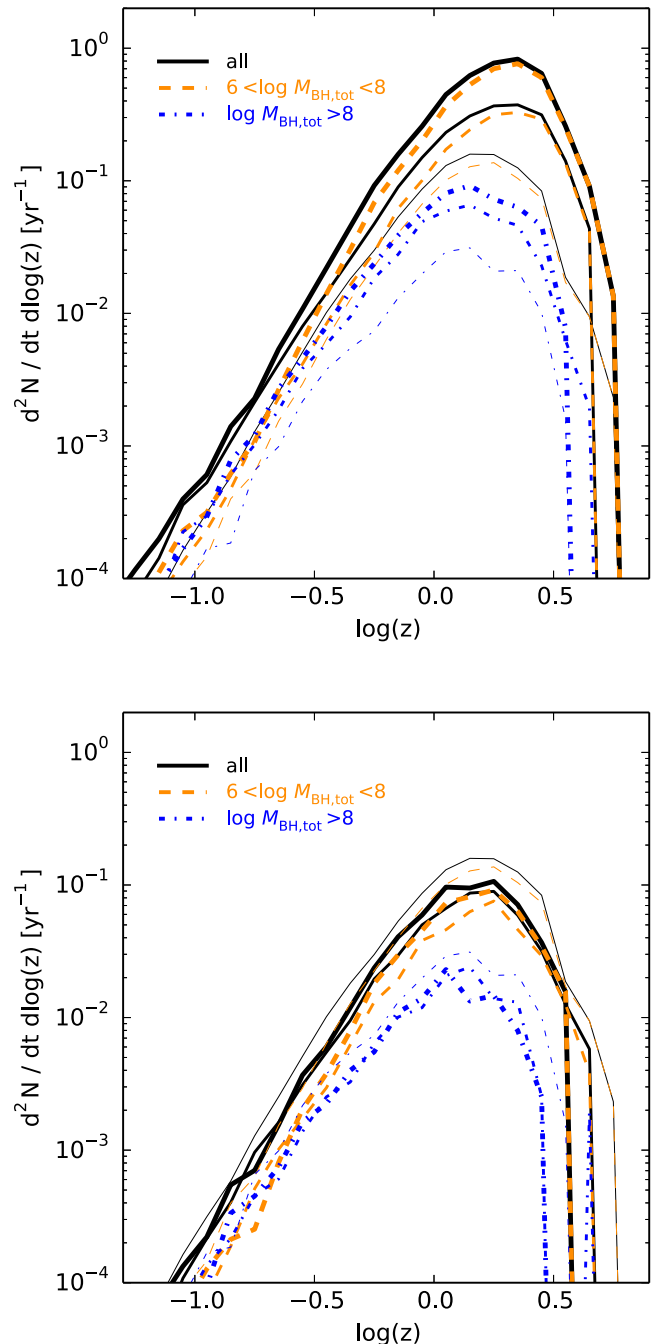

Figure A1. Top panel: the BH merger rate in Illustris is shown in the same manner as in Fig. 2, but here the resolution convergence between Illustris-1, 2, and 3 is shown. Only mergers for which each progenitor galaxy has at least 300 DM and 80 stellar particles are included. As before, the solid black lines denote the total merger rate for our BH sample with a delay time-scale added to each merger, as described in the text. The orange dashed and blue dot–dashed lines show the merger rate separated by total BH mass: $6 < \log M_{BH} < 8$ and $\log M_{BH} > 8$, respectively. In all cases, thicker lines denote higher resolution simulations. Bottom panel: same as the top panel, but with a constant minimum *mass* imposed on progenitor galaxies, rather than a minimum particle number. A minimum stellar (halo) mass of 6.4×10^9 $M_\odot$ (1.2×10^{11} $M_\odot$) is assumed, corresponding to 300 DM (80 stellar) particles in Illustris-3, such that this curve is identical between the two panels. The convergence improves substantially in this case.
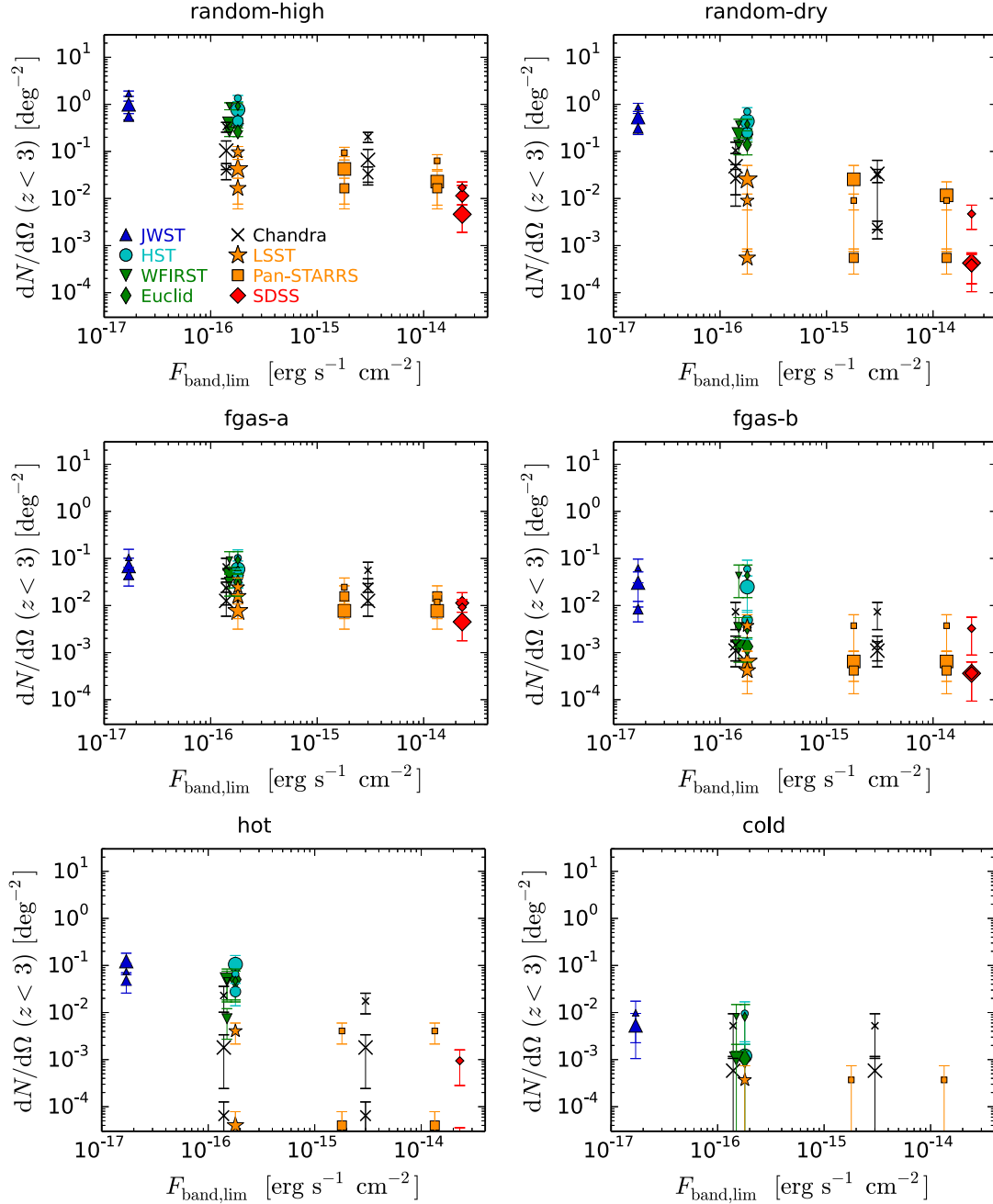
Figure A2. As in the top panels of Fig. 13, each panel shows the cumulative number of observable spatially offset AGN per deg^2, out to $z = 3$, for various surveys as a function of limiting flux in the corresponding band. Different panels correspond to different spin models, as indicated on the plot labels. The `5deg` spin model is not shown, as the number counts are nearly all below the plot range. The points are in the same style as Fig. 13, except that here results are compared between Illustris-1, 2, and 3. Point size denotes the simulation resolution; the largest points correspond to Illustris-1. As in the bottom panel of Fig. A1, constant minimum DM and stellar mass are assumed for the progenitor haloes of merging BHs in each simulation, rather than a constant minimum particle number. As a result, the predictions here for Illustris-1 are lower than in Fig. 13, where lower mass haloes are included.

This paper has been typeset from a TeX/LaTeX file prepared by the author.