

# The power of infrared AGN selection in mergers: a theoretical study

Laura Blecha,<sup>1</sup>★ Gregory F. Snyder,<sup>2</sup> Shobita Satyapal<sup>3</sup> and Sara L. Ellison<sup>4</sup>

<sup>1</sup>Department of Physics, University of Florida, 2001 Museum Rd., Gainesville, FL 32611, USA

<sup>2</sup>Space Telescope Science Institute, 3700 San Martin Dr, Baltimore, MD 21218, USA

<sup>3</sup>Department of Physics and Astronomy, George Mason University, MS 3F3, 4400 University Drive, Fairfax, VA 22030, USA

<sup>4</sup>Department of Physics and Astronomy, University of Victoria, Finnerty Road, Victoria, BC V8P 1A1, Canada

Accepted 2018 May 12. Received 2018 May 2; in original form 2017 October 31

## ABSTRACT

The role of galaxy mergers in fuelling active galactic nuclei (AGNs) is still debated, owing partly to selection effects inherent to studies of the merger/AGN connection. In particular, luminous AGNs are often obscured in late-stage mergers. Mid-infrared (IR) colour selection of dust-enshrouded AGN with, e.g. the *Wide-field Infrared Survey Explorer* (WISE) has uncovered large new populations of obscured AGN. However, this method is sensitive mainly to AGN that dominate emission from the host. To understand how selection biases affect mid-IR studies of the merger/AGN connection, we simulate the evolution of AGN throughout galaxy mergers. Although mid-IR colours closely trace luminous, obscured AGN, we show that nearly half of merger-triggered AGNs are missed with common mid-IR selection criteria, even in late-stage, gas-rich major mergers. At  $z \lesssim 0.5$ , where merger signatures and dual nuclei can most easily be detected, we find that a more lenient  $W1 - W2 > 0.5$  cut greatly improves completeness without significantly decreasing reliability. Extreme nuclear starbursts are briefly able to mimic this AGN signature, but this is largely irrelevant in mergers, where such starbursts are accompanied by AGN. We propose a two-colour cut that yields high completeness and reliability even in starbursting systems. Further, we show that mid-IR colour selection very effectively identifies *dual* AGN hosts, with the highest fraction at the smallest separations ( $< 3$  kpc). Thus, many merger hosts of mid-IR AGN should contain unresolved dual AGN; these are ideal targets for high-resolution follow-up, particularly with the *James Webb Space Telescope*.

**Key words:** accretion, accretion discs – black hole physics – galaxies: active – galaxies: interactions – infrared: galaxies.

## 1 INTRODUCTION

Mergers between galaxies cause morphological transformations, induce bursts of star formation, and fuel active galactic nuclei (AGNs), and the resulting stellar and AGN feedback regulates further evolution. Galaxy mergers therefore offer a plausible mechanism for coordinated evolution between supermassive black hole (BHs) and their host galaxies, giving rise to the observed correlations between BH mass and stellar bulge properties (e.g. Magorrian et al. 1998; Gebhardt et al. 2000; Gültekin et al. 2009; McConnell & Ma 2013). However, despite a great deal of study, the connection between galaxy mergers and AGN fuelling is still not well understood.

The large majority of AGN hosts do not show signs of ongoing merger activity, suggesting that secular mechanisms are responsible for triggering most AGN and leading some to conclude that there is no convincing evidence for a merger/AGN connection (e.g. Cis-

ternas et al. 2011; Kocevski et al. 2012; Schawinski et al. 2012; Villforth et al. 2014, 2017). In contrast, studies that select samples of paired and merging galaxies have found that when mergers do occur, they are more likely than isolated galaxies to host AGN, and that AGN are most likely to be found in the late stages of the merger (e.g. Liu et al. 2011; Ellison et al. 2011; Silverman et al. 2011; Ellison et al. 2013b; Satyapal et al. 2014; Lackner et al. 2014; Weston et al. 2017; Goulding et al. 2018). The merger/AGN association also depends strongly on the AGN luminosity (e.g. Ellison et al. 2013b; Satyapal et al. 2014; Koss et al. 2012). At the highest quasar luminosities, *most* AGN may be hosted in ongoing mergers (e.g. Urrutia, Lacy & Becker 2008; Fan et al. 2016; Vito et al. 2018).

Crucially, the observed correlation between mergers and AGN fuelling is highly sensitive to obscuration by gas and dust, which can be extreme during merger-induced nuclear starbursts. Indeed, the merger fraction of AGN hosts is much higher for infrared (IR) and hard X-ray-selected AGN, which are less sensitive to dust obscuration, than for optical or soft X-ray-selected AGN (e.g. Sanders et al. 1988a,b; Sanders & Mirabel 1996; Veilleux et al. 2009; Koss et al.

★ E-mail: lblecha@ufl.edu

2010; Satyapal et al. 2014; Kocevski et al. 2015). Low-excitation radio galaxies, in contrast, have no apparent connection to galaxy mergers (Ellison, Patton & Hickox 2015). Ultraluminous infrared galaxies (ULIRGs,  $L_{\text{IR}} > 10^{12} L_{\odot}$ ) emit most of their bolometric luminosity in the IR; in the nearby Universe, nearly all of these are strongly merging systems that contain buried starbursts or AGNs (e.g. Sanders et al. 1988a; Veilleux, Kim & Sanders 2002; Tacconi et al. 2008). The most AGN-dominated ULIRGs are found in the latest stages of merging, with the highest IR luminosities (Veilleux et al. 2009).

The presence of dust in the nuclear region is a nearly ubiquitous feature of AGN. Partial obscuration by a dusty ‘torus’ (or another geometry with covering fraction less than unity) is commonly invoked in AGN unification theories to explain the Type I/Type II dichotomy observed in optical AGN spectra (e.g. Urry & Padovani 1995). Dusty AGNs produce distinctive spectral signatures in the mid-IR. The hard ionizing AGN continuum emission heats the surrounding dust, up to the sublimation limit ( $\sim 1500$  K); the reprocessed emission appears as a red power-law slope in the mid-IR spectral energy distribution (SED; defined as  $\sim 3\text{--}30\ \mu\text{m}$ ). This fact has been exploited to identify AGN based on their characteristic red mid-IR colours (e.g. Laurent et al. 2000; Lacy et al. 2004; Stern et al. 2005; Assef et al. 2010; Donley et al. 2012). Large new populations of AGN, most undetected in optical surveys, have been identified via mid-IR colour selection in wide-field surveys, most recently the *Wide-field Infrared Survey Explorer* (WISE) all-sky survey (e.g. Jarrett et al. 2011; Stern et al. 2012; Mateos et al. 2012, 2013; Assef et al. 2013; Secrest et al. 2015).

The WISE bands (W1, W2, W3, and W4), are centred at 3.4, 4.6, 12, and 22  $\mu\text{m}$ , respectively. Stern et al. (2012) proposed a single-colour cut of  $W1 - W2 > 0.8$  (Vega magnitudes) to identify luminous AGN with high reliability and completeness. Others have proposed two-colour cuts to minimize contamination from star-forming galaxies, especially at  $z \gtrsim 1$  (Jarrett et al. 2011, hereafter J11; Mateos et al. 2012; Assef et al. 2013). Mateos et al. (2012) also demonstrate that the completeness is a strong function of AGN luminosity, with high completeness achieved only for the most luminous AGN ( $L_{2-10\text{ keV}} > 10^{44}\text{ erg s}^{-1}$ ).

Satyapal et al. (2014) used WISE colour selection to identify a new population of low-redshift merger-triggered AGNs. Using samples of merging and isolated selected from the Sloan Digital Sky Survey (Ellison et al. 2011, 2013b), they cross-matched with the WISE data and found a significant excess of AGN activity in advanced mergers, relative to isolated galaxies and to the excess measured for optically selected AGN. The mid-IR-selected AGN excess in mergers persists whether a selection cut of  $W1 - W2 > 0.8$  or a more lenient  $W1 - W2 > 0.5$  cut is used; this sample is at low redshift ( $z < 0.2$ ) and is therefore less sensitive to contamination from star-forming galaxies. Recently, Weston et al. (2017) and Goulding et al. (2018) have confirmed the excess of WISE AGN in merging galaxies, and Donley et al. (2018) have found similar results for AGN selected with the *Spitzer*–Infrared Array Camera (IRAC) (especially compared with 2–10 keV X-ray-selected AGN).

Similar trends are seen for hard X-ray-selected AGN in merging galaxies (e.g. Koss et al. 2010, 2012, 2018; Kocevski et al. 2015; Ricci et al. 2017; Lansbury et al. 2017). Hard X-ray AGN selection is highly accurate and uniquely complete, but only to the comparatively shallow depth of current surveys with, e.g. the *Swift*–Burst Alert Telescope and *NuSTAR*. Mid-IR colour selection yields much larger AGN samples, but it is sensitive only to AGNs that are luminous relative to their host galaxies. Thus, to fully utilize mid-IR selection as a probe of the merger/AGN connection, we must under-

stand how the selection completeness varies throughout the course of a merger and in different merger environments.

With this motivation, we use hydrodynamics and radiative transfer simulations of merging galaxies to model the evolution of nuclear obscuration and mid-IR SEDs through different merger stages. A similar approach to studying mid-IR AGN signatures in mergers has been used by Snyder et al. (2013) and Roebuck et al. (2016), who focus on signatures sensitive to the AGN versus starburst contribution in ULIRGs. Snyder et al. (2013) find that AGN signatures can be suppressed by dust self-absorption in extreme starbursting  $z \sim 2$  galaxies, and they propose a combination of spectral signatures accessible to *James Webb Space Telescope* (JWST) that can constrain the AGN fraction. Roebuck et al. (2016) compare the empirical and simulated AGN fraction in ULIRGs and find that the AGN fraction may be underestimated by empirical classifications in some cases.

In this work, we focus on the selection of merger-triggered AGN with WISE; we wish to quantify the completeness and accuracy of mid-IR AGN selection and how this depends on the properties of the merging system. Our other primary goal is determining the incidence of (possibly unresolved) close dual AGN in mid-IR-selected samples – i.e. simultaneously active BH pairs with  $\lesssim 1\text{--}10$  kpc separations. A small but growing sample of dual AGN has been identified in optically selected surveys (e.g. Comerford et al. 2009; Liu et al. 2010; Comerford et al. 2012; Müller-Sánchez et al. 2015), but hard X-ray-selected samples have revealed a much higher incidence of dual AGN in late-stage mergers (Koss et al. 2012). Follow-up studies of WISE AGN in mergers have also seen a high success rate in finding candidate dual AGN (Satyapal et al. 2017; Ellison et al. 2017). We aim to determine the intrinsic dual AGN fraction in WISE-selected sources from our models, as a function of nuclear separation and AGN luminosity. Because we are interested specifically in mergers, we focus on modelling galaxies representative of relatively low-redshift systems ( $z \lesssim 0.5$ ), similar to the observed samples of Satyapal et al. (2014), Weston et al. (2017), and Goulding et al. (2018). At higher redshifts ( $z \sim 1\text{--}2$ ) more typical of the overall WISE AGN population, merger signatures and dual AGN are increasingly difficult to reliably detect.

We use Vega magnitudes throughout this paper. The hydrodynamics simulations and radiative transfer calculations are described in Sections 2.1 and 2.2, respectively. Our results are presented in Section 3. Merger-driven AGN obscuration is discussed in Section 3.5, and AGN lifetimes as a function of merger phase are presented in Section 3.1. We explore the completeness and reliability of mid-IR AGN selection in Sections 3.3 and 3.2. In Section 3.4, we consider the effectiveness of mid-IR selection of dual AGN. Finally, in Section 4, we discuss the implications of our results and summarize our main conclusions.

## 2 METHODOLOGY

### 2.1 Hydrodynamic galaxy merger simulations

As the basis for modelling AGN in mergers, we conduct high-resolution simulations of merging galaxies with GADGET-3, a smoothed-particle hydrodynamics and  $N$ -body code that conserves energy and entropy and uses sub-resolution physical models for radiative heating and cooling, star formation, supernova feedback, metal enrichment, and a multiphase interstellar medium (ISM; Springel & Hernquist 2003; Springel 2005). BHs are modelled as gravitational ‘sink’ particles that accrete gas via an Eddington-limited, Bondi–Hoyle like prescription. Thermal AGN feedback is included by coupling 5 per cent of the accretion luminosity

**Table 1.** Progenitor galaxy parameters. Column 1: name of the progenitor used throughout the text. Columns 2 and 3: total and stellar mass of the galaxy, respectively. Column 4: initial gas fraction in the galaxy disc (by mass,  $M_{\text{gas}}/M_{*, \text{disc}}$ ). Column 5: bulge-to-total mass ratio.

Name	$M_{\text{tot}}$ ( $10^{11} M_{\odot}$ )	$M_*$ ( $10^{10} M_{\odot}$ )	$f_{\text{gas}}$	$B/T$
A0	14	3.9	0.3	0
B0	14	5.0	0.1	0
C0	14	4.1	0.3	0.1
D0	14	4.2	0.3	0.2
E0	14	5.1	0.1	0.2
A1	6.8	2.0	0.3	0
B1	6.8	2.5	0.1	0
C1	6.8	2.0	0.3	0.1
D1	6.8	2.1	0.3	0.2
E1	6.8	2.6	0.1	0.2
A2	2.7	0.78	0.3	0

( $L_{\text{bol}} = \epsilon_{\text{rad}} \dot{M} c^2$ ) to the surrounding gas as thermal energy, with an accretion-dependent radiative efficiency  $\epsilon_{\text{rad}}$  at low accretion rates (cf. Narayan & McClintock 2008). GADGET-3 has been used for a multitude of studies of merging galaxies, including many studies of BH/galaxy co-evolution (e.g. Di Matteo, Springel & Hernquist 2005; Robertson et al. 2006; Cox et al. 2006; Hopkins et al. 2006, 2008; Blecha et al. 2011; Blecha, Loeb & Narayan 2013b).

As mentioned above, our primary goal is to study mid-IR AGN selection in merging galaxies, as probes of dual AGN and the merger/AGN connection. Identifying merging galaxies or dual AGN empirically requires the ability to resolve dual nuclei and to detect low-surface-brightness signatures of morphological disturbance (e.g. tidal tails). Because such signatures are increasingly difficult to identify at higher redshifts, our simulations are designed to be representative of galaxy mergers at relatively low redshift, rather than very gas rich ( $f_{\text{gas}} \gtrsim 0.5$ ), extreme starbursting systems that may be more typical of some merger-triggered *WISE* AGN at the median redshift of the *WISE* AGN population ( $z \sim 1$ ). This choice is in agreement with several recent empirical studies of IR AGN signatures in merging galaxies, which have similarly focused on the low-redshift regime (Satyapal et al. 2014; Weston et al. 2017; Goulding et al. 2018).

Our primary simulation suite includes eight merger simulations with galaxy mass ratios of  $q = 0.2$ – $1$ ; most of these are major mergers with  $q = 0.5$ , but we include one equal-mass merger and one ‘minor’ merger with  $q = 0.2$ . The galaxies consist of a dark matter halo, a disc of gas and stars, a stellar bulge, and a central BH. We focus on merger progenitors with disc-dominated morphologies initially, choosing the bulge-to-total stellar mass ratio ( $B/T$ ), between 0 (i.e. a pure disc initially) and 0.2. The initial gas fraction in the disc (by mass,  $M_{\text{gas}}/M_{*, \text{disc}}$ ) is set to be 0.3 or 0.1, to represent both gas-rich and relatively gas-poor mergers. The fiducial baryonic gravitational softening length and mass resolution are  $\epsilon_{\text{grav}} = 48$  pc and  $m_b = 2.8 \times 10^5 M_{\odot}$ , respectively. We also run a subset of these simulations with 10 times higher mass resolution and  $10^{1/3}$  times smaller softening length,  $\epsilon_{\text{grav}} = 23$  pc. A comparison between the fiducial- and high-resolution results is presented in Appendix A.

The parameters of the progenitor galaxy models are summarized in Table 1. The merger simulation parameters are given in Table 2, along with some key characteristics of each merger, and the simulations that have also been run at higher resolution are denoted in boldface. Throughout the paper, we refer to the mergers either by a name, composed of the names assigned to each progenitor galaxy

**Table 2.** Merger simulation parameters and key characteristics. Column 1: simulation number. Column 2: simulation name, indicating the combination of merging progenitor galaxies as in the notation of Table 1. Simulations names in boldface have also been run at 10 times higher mass resolution. Column 3: merging galaxy mass ratio. Columns 4 and 5: maximum total SFR and specific SFR ( $\text{SFR}/M_*$ ) achieved during the simulation. Column 6: maximum column density (along sightlines to either BH) achieved during the simulation, calculated as the median over all viewing angles in each snapshot. Column 7: maximum bolometric AGN luminosity achieved during the simulation.

#	Name	$q$	$\text{SFR}_{\text{max}}$ ( $M_{\odot} \text{ yr}^{-1}$ )	$\text{sSFR}_{\text{max}}$ ( $\log \text{ yr}^{-1}$ )	$N_{\text{H, max}}$ ( $\log \text{ cm}^{-2}$ )	$L_{\text{AGN, max}}$ ( $\log \text{ erg s}^{-1}$ )
0	<b>A0A0</b>	1	380	−8.4	24.7	46.0
1	<b>A1A0</b>	0.5	220	−8.5	24.5	46.3
2	C1C0	0.5	14	−9.7	23.6	45.2
3	<b>D1D0</b>	0.5	25	−9.5	24.2	45.6
4	<b>A1E0</b>	0.5	46	−9.2	24.4	45.7
5	A2A0	0.2	15	−9.6	23.8	45.0
6	B1B0	0.5	21	−9.6	24.0	45.0
7	E1E0	0.5	1.6	−11	22.4	43.7

(e.g. A1A0), or by the simulation number (0–7) as listed in Table 2. We also frequently refer to simulations 0–5 as the subset of ‘gas-rich’ mergers, simulations 0–4 as the ‘gas-rich, major’ mergers, and simulations 6–7 as the ‘gas-poor’ mergers.

## 2.2 Dust radiative transfer simulations

We conduct radiative transfer simulations in post-processing with the 3D, polychromatic, Monte Carlo dust radiative transfer code SUNRISE (Jonsson 2006; Jonsson, Groves & Cox 2010). This publicly available code has been used extensively with GADGET-3 to model a wide range of isolated and merging galaxy populations (e.g. Lotz et al. 2011; Narayanan et al. 2010; Hayward et al. 2011, 2013; Snyder et al. 2013; Blecha et al. 2013a; Lanz et al. 2014). Stellar emission is calculated from single stellar population STARBURST99 SEDs (Leitherer et al. 1999) based on the age and metallicity of each stellar particle, assuming a Kroupa initial mass function (Kroupa 2002). Emission from HII regions and photodissociation regions (PDRs) around young stars is calculated using the MAPPINGS III models of Groves et al. (2008), where age, metallicity, and gas pressure are taken from the newly formed stellar particles in the GADGET-3 simulation. A covering fraction of 0.2 is assumed for the PDR models, which include dust re-emission as well as polycyclic aromatic hydrocarbon (PAH) absorption and emission.

An AGN SED is implemented based on the BH accretion rate using the luminosity-dependent templates of Hopkins, Richards & Hernquist (2007). In the mid-IR, the AGN SED is derived from the empirically determined mean quasar SEDs of Richards et al. (2006), such that it implicitly includes reprocessed radiation from a sub-resolution dusty ‘torus’ typical of Type I quasars.

The dust distribution is calculated in SUNRISE by projecting the gas-phase metal density onto a 3D adaptively refined grid, where we assume that 40 percent of the metals are in dust (Dwek 1998) and use the Milky Way  $R_V = 3.1$  dust model of Draine & Li (2007). We also tested the effect of using the Draine & Li (2007) dust model based on the Small Magellanic Cloud bar, which has fewer carbonaceous grains and hence weaker PAH features, but found that this did not significantly change our results.

GADGET-3 uses a multiphase model for the ISM, in which cold gas is assumed to be clumpy, with a negligible volume filling fac-



tor. We make the same assumption about the gas (and hence the dust) in our *SUNRISE* models: attenuation and emission from dust in cold gas clouds is neglected in the radiative transfer calculations, including the IR emission. (This is equivalent to the ‘multiphase-on’ model of Hayward et al. 2011). Because galactic-scale dust primarily emits at far-IR wavelengths, ignoring this cold-phase dust emission should not significantly affect the mid-IR diagnostics considered here. In high-redshift, extremely gas-rich mergers where molecular gas dominates the ISM, the zero-volume-filling assumption for cold-phase gas would break down, and a more appropriate choice would be to assume the gas and dust is distributed uniformly across the grid cells (cf. Hayward et al. 2011; Snyder et al. 2013). The galaxies considered in this work are instead designed to be low-redshift analogues with a maximum initial gas fraction of 0.3, which are dominated by hot-phase gas throughout the merger. The most gas rich of these do briefly become ULIRGs with  $L_{\text{IR}} \sim 10^{12} L_{\odot}$  during the peak of merging activity, with starburst and AGN activity driven by an increase in the cold gas fraction in the galactic nuclei.

Using similar *SUNRISE* simulations, Snyder et al. (2013) demonstrate that even the mid-IR signatures of AGN can be obscured by dust self-absorption in the nuclei of extreme starbursting, merging systems, but they find significant suppression of AGN signatures only for merger events featuring highly gas-rich  $z \sim 2$  analogues that reach ‘hyper-LIRG’ luminosities ( $L_{\text{IR}} \sim 10^{13} L_{\odot}$ ). For the less-extreme, less-obscured systems considered here, the multiphase ISM treatment that assumes negligible absorption and emission by dust in cold clumps is considered to be the most appropriate choice. None the less, it is important to bear in mind that even mid-IR selection can miss the most heavily buried, Compton-thick AGN during peak obscuration in mergers, and that even in less-extreme mergers, select sightlines that intersect with cold gas clumps could obscure the AGN signatures in the mid-IR.

One should also keep in mind that these dust and ISM models are approximations of physics on sub-grid scales, an unavoidable consequence of the finite resolution of simulations. Dust grain distributions and gas clumpiness in the ISM of real galaxies will be spatially dependent, evolving, and correlated with processes such as starbursts and AGN heating. The inherent uncertainty in sub-resolution dust and ISM models is an important caveat to any such study. The model parameters considered here are physically motivated and empirically based, and as noted above, we focus on a regime in which the primary concern of this study – mid-IR colour selection of AGN – should not be subject to dramatic effects from these model uncertainties.

Once the source SEDs are determined for stellar particles, star-forming regions, and AGN, and the dust distribution is determined from the gas grid, *SUNRISE* performs Monte Carlo radiative transfer through the grid, computing energy absorption (including dust self-absorption) and thermal re-emission to produce spatially resolved ultraviolet-to-IR SEDs. To ensure convergence of the dust temperature and IR emission calculation, the emergent IR luminosity in each grid cell is required to be converged to within a factor of 0.2 (such that the *integrated* IR SED is converged to a much higher degree). We find that adjusting this tolerance parameter between 0.1 and 0.4 has a negligible effect ( $< 3$  per cent) on the emergent IR SED.

For each merger simulation, we run *SUNRISE* on snapshots at 100 Myr intervals in the early stage of the merger and at 10 Myr intervals in the later stage of the merger. Images and resolved spectra are produced for seven isotropically distributed viewing angles, with a spatial resolution of 500 pc or 167 pc in the early- or late-

merger stages, respectively. (A  $3 \times$  larger field of view is used in the early merger phase, when the galaxies are at larger projected separations.) Unless otherwise specified, all results are presented for rest-frame SEDs, but we have also calculated broad-band mid-IR magnitudes for SEDs redshifted up to  $z = 1$  for the fiducial simulation suite, and up to  $z = 4$  for select simulation snapshots. As discussed below, we find negligible difference in our results for  $z \lesssim 0.5$ . Finally, in order to quantify the contribution of the AGN to the emergent SED, we also re-run each *SUNRISE* simulation with the AGN luminosity artificially set to zero. We refer to these as the ‘AGNx0’ simulations. While this does not remove the effect of AGN feedback on the surrounding gas (because the *GADGET-3* simulation is the same for the fiducial and AGNx0 *SUNRISE* simulations), it allows us to explicitly separate the contribution of stellar and AGN emission to the emergent SED. Among other things, this is essential for measuring the contribution of star-forming regions to the mid-IR colours.

### 3 RESULTS

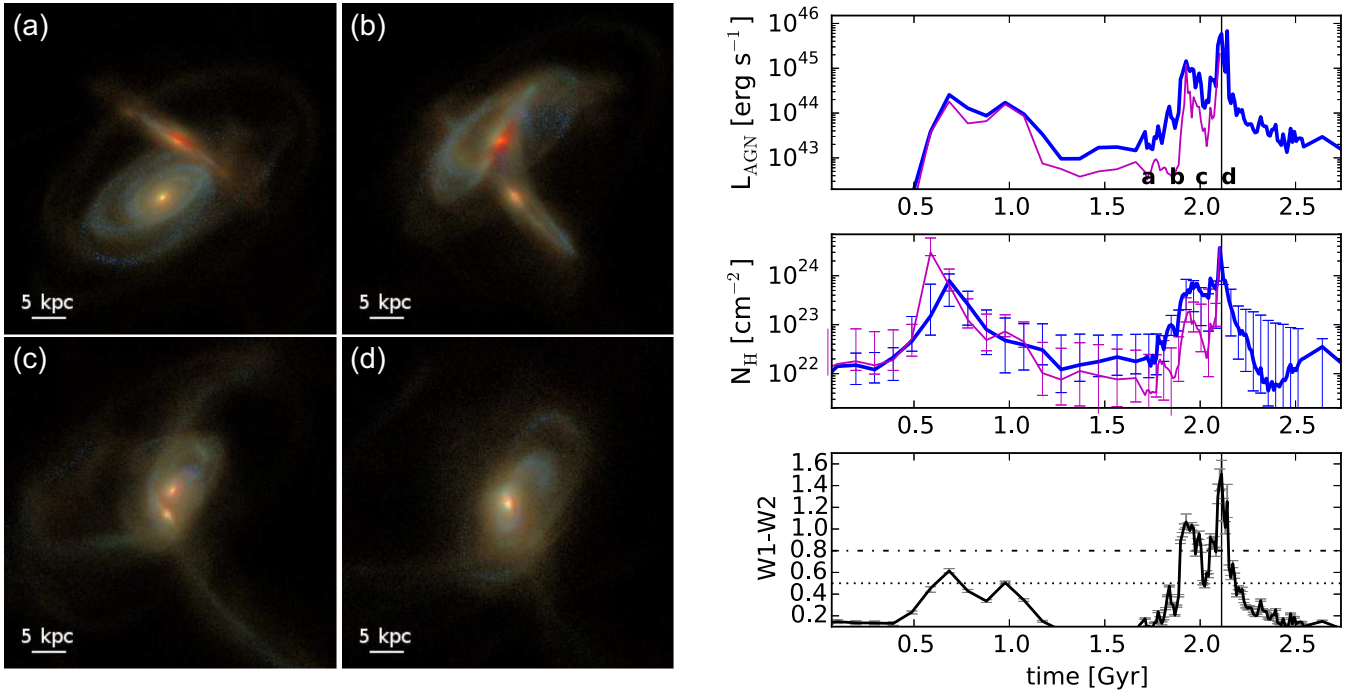
The evolution of a high-resolution, gas-rich, major merger simulation is shown in Fig. 1.<sup>1</sup> The bolometric AGN luminosity and nuclear column density ( $N_{\text{H}}$ , calculated along the line of sight to each BH<sup>2</sup>) peak during final coalescence of the galaxies, supporting the idea that luminous, obscured AGN are preferentially triggered in late-stage galaxy mergers. The *WISE* mid-IR colours (which are calculated globally for the entire galaxy) closely trace this luminous, obscured AGN phase, with  $W1 - W2 > 0.8$  for 180 Myr during the late stages of the merger (defined from the time when the BH separation falls below 10 kpc, to 100 Myr after the BH merger). This corresponds to 46 per cent of the total late-stage merger phase.

By the time of BH merger, most of the galaxies’ initial gas content has been consumed by star formation (and BH accretion). Shortly after the BH merger, AGN feedback efficiently heats up and removes much of the remaining cold gas reservoir from the nucleus of the merger remnant, causing a precipitous decline in AGN luminosity, central gas column density, and star formation rate (SFR). Such AGN feedback episodes provide a means of regulating growth of both the galaxy bulge and the central BH (e.g. Wyithe & Loeb 2003; Di Matteo et al. 2005; Hopkins et al. 2006), and they input substantial energy and metal-enriched gas into the circumgalactic medium (e.g. Hani et al. 2017). In the corresponding simulated optical images (Fig. 1), significant reddening of the nuclear regions is apparent. Disturbed morphological merger signatures are also seen throughout the late merger phase. It is worth noting, however, that the most prominent tidal features have already begun to fade in the last image, just 30 Myr after the BH merger, when the AGN is at its peak luminosity.

Fig. 2 shows the mid-IR SEDs for the four late-stage merger snapshots corresponding to the images in Fig. 1, including the AGNx0

<sup>1</sup>We use data from one of the high-resolution simulations here, to illustrate the morphological merger signatures in greater detail, but the qualitative and quantitative results are very similar for the fiducial-resolution simulations used for most of our analysis. See Appendix A for a comparison of key results between fiducial- and high-resolution simulations.

<sup>2</sup>This aperture for the  $N_{\text{H}}$  calculation is chosen to be consistent with the size scale ( $2.8 \times \epsilon_{\text{grav}}$ ) below which gravitational forces are mediated by the softening kernel. This quantity, which is 64 pc for the high-resolution simulations (as in Fig. 1) and 136 pc for the fiducial simulations, can be considered the effective spatial resolution of the simulations.



**Figure 1.** Evolution of an obscured AGN in the high-resolution simulation of the A1A0 merger. Left-hand panels: SDSS *ugz* images, generated from *SUNRISE* simulations, show the disturbed merger morphology during the late stages of the merger. In the first snapshot, the BH separation has just fallen below 10 kpc, and the final snapshot occurs just after the BH merger. The *ugz* filter combination is chosen to enhance contrast between the dust-obscured nuclei and star-forming regions. Right-hand panels: from top to bottom, the evolution of the bolometric AGN luminosity  $L_{\text{AGN}}$ , line-of-sight gas column density  $N_{\text{H}}$ , and *WISE*  $W1 - W2$  colour throughout the merger is shown. The vertical line denotes the time of BH merger. In the  $L_{\text{bol}}$  and  $N_{\text{H}}$  plots, the blue and magenta curves correspond to each BH prior to merger, and the blue curve shows the post-merger evolution. In the  $N_{\text{H}}$  and  $W1 - W2$  plots, the error bars show the range of values over all viewing angles.  $N_{\text{H}}$  is calculated along the line of sight to each BH, in an aperture 64 pc in size (consistent with the effective spatial resolution of the high-resolution simulations). *WISE* colours are calculated for the entire galaxy. For clarity, error bars are plotted for only a subset of snapshots. The dashed and dotted lines in the  $W1 - W2$  plot denote single-colour cuts of 0.5 and 0.8, used in the literature and in this work. Note that higher time resolution is used for the *SUNRISE* calculations in the late phases of the merger. The AGN luminosity and column density peak during the galaxies’ coalescence, when the galaxies are morphologically disturbed; this luminous, obscured AGN phase is closely traced by red  $W1 - W2$  colours.

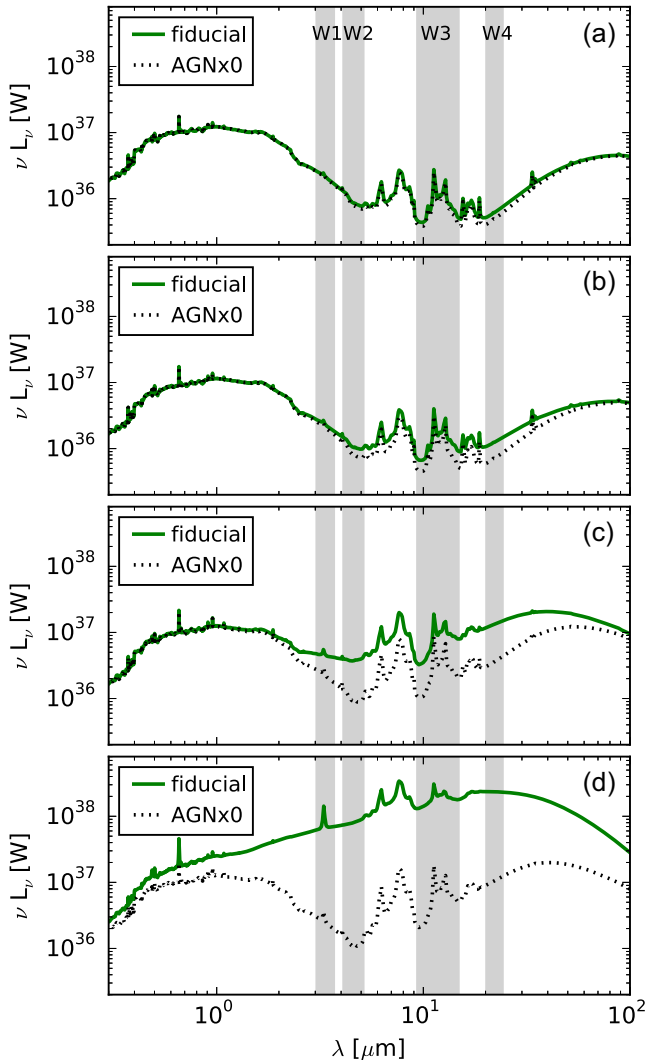
SEDs (for consistency with Fig. 1, we use data from the same high-resolution simulation here, though the SEDs are very similar in the fiducial-resolution simulation). In snapshot ‘a’, the merging galaxies are just coming together for their final coalescence. The total AGN luminosity is low,  $2 \times 10^{43} \text{ erg s}^{-1}$ , and the SED is overwhelmingly dominated by stellar emission. The AGN luminosity increases as the coalescence proceeds, both in absolute terms and relative to the total host luminosity. In snapshot ‘c’, the AGN contributes 40 per cent of the total luminosity, and the *WISE*  $W1 - W2$  colour is 0.8. At longer mid-IR wavelengths, prominent PAH emission and  $9.7 \mu\text{m}$  silicate absorption are apparent, reflecting the marginally dominant contribution of stellar emission to the total SED. At the peak of AGN activity (snapshot ‘d’), where  $L_{\text{AGN}} = 7 \times 10^{45} \text{ erg s}^{-1}$ , the AGN clearly dominates the total SED, which has a red slope from the near-IR through the  $12 \mu\text{m}$  *WISE*  $W3$  band. Here, the AGN overwhelms most of the signatures of ongoing dusty star formation; the PAH emission and silicate absorption apparent in the AGNx0 simulation (and in the earlier fiducial snapshots) are much less prominent. The  $W1 - W2$  colour is 1.3 at the AGN peak, which, given the 90 per cent contribution of the AGN to the total SED, is similar to the  $W1 - W2$  colour of the intrinsic AGN SED template.

The simulation shown in Figs 1 and 2 is the high-resolution version of the A1A0 merger, with 10 times higher mass resolution (and  $10^{1/3}$  times higher spatial resolution) than the fiducial simulations. The higher spatial resolution reveals morphological features such as tidal tails, star-forming regions, and dust-reddened nuclei

in great detail in the images. The qualitative trends in  $L_{\text{AGN}}$ ,  $N_{\text{H}}$ , and  $W1 - W2$  seen in Fig. 1 are generic to all of the major, gas-rich merger simulations; namely, we see a peak in activity soon after the galaxies’ first pericentric passage, and a larger peak during final coalescence. The minor merger in our sample (A2A0, or simulation 5) and the gas-poor mergers (6 and 7) trigger significantly less star formation and AGN activity. In particular, the E1E0 gas-poor merger (7) never exceeds the  $L_{\text{AGN}} > 10^{44} \text{ erg s}^{-1}$  AGN threshold used in much of our analysis. Unless otherwise specified, the results below refer to the fiducial-resolution simulation suite; in Appendix A, we demonstrate the consistency of results between fiducial- and high-resolution simulations.

### 3.1 (Mid-IR) AGN lifetimes in merging galaxies

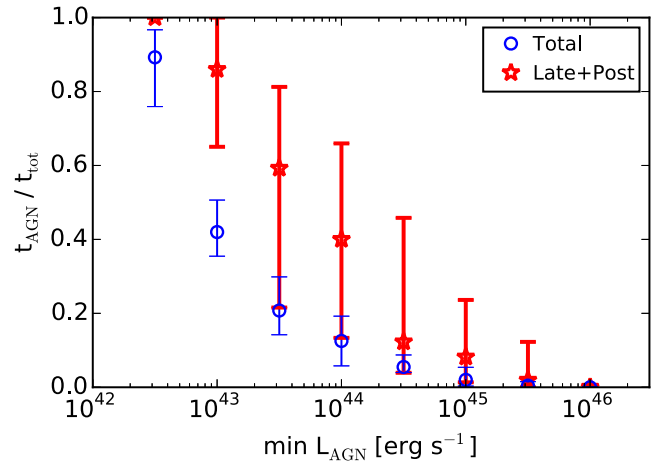
For much of our subsequent analysis, we divide each merger simulation into three phases: ‘Early’, ‘Late’, and ‘Post-Merger’, based on the BH separation. The ‘Early’ merger phase is defined as all times when the BH separation is  $> 10 \text{ kpc}$ , the ‘Late’ phase is defined for BH separations  $0 < a_{\text{sep}} < 10 \text{ kpc}$ , and the ‘Post-Merger’ phase begins at the time of the BH merger. Unless otherwise noted, the BH separation  $a_{\text{sep}}$  refers to the *projected* separation, averaged over all viewing angles, and the Post-Merger phase is restricted to the first 100 Myr of post-BH-merger evolution. Because BH binary inspiral time-scales are uncertain at sub-resolution scales (and are



**Figure 2.** Optical and IR SEDs are shown for the high-resolution A1A0 simulation snapshots corresponding to the optical images in Fig. 1. Green solid lines give the fiducial simulation SED, and black dotted lines give corresponding SED in the AGNx0 simulation. The grey shaded regions denote the four *WISE* bands (from left to right: W1, W2, W3, and W4). In the first two snapshots, the system is still dominated by star formation, and prominent PAH and  $9.7\ \mu\text{m}$  silicate absorption features are apparent. In snapshots ‘c’ and ‘d’, we see the reddening of the mid-IR colours with increasing AGN contribution, particularly in the last snapshot. The W1 – W2 colours in each snapshot are 0.1 (‘a’), 0.2 (‘b’), 0.8 (‘c’), and 1.3 (‘d’).

quite difficult to constrain observationally), we also combine the latter two phases (‘Late+Post’) for some of our analysis.

Using these definitions, we show in Fig. 3, the fraction of time that an AGN is active in the Late+Post merger phases. We refer to this as the AGN duty cycle and show how it varies with the minimum observable AGN luminosity, and also how it compares with the AGN duty cycle for the entire merger simulation (including the Early phase). While the total duty cycle depends directly on the simulation duration, which is not a physical quantity, it is useful in relative terms when compared to the Late+Post-stage duty cycle. As expected, AGN are on average much more active in the Late+Post merger phases. When low-luminosity AGN are included ( $L_{\text{AGN}} > 10^{43}\ \text{erg s}^{-1}$ ), the BHs are typically active for  $\gtrsim 85$  per cent of the Late+Post-merger phase. For moderate-to-high-luminosity



**Figure 3.** The AGN duty cycle ( $t_{\text{AGN}}/t_{\text{tot}}$ ) is shown for advanced mergers (Late+Post phases; red stars), versus the minimum bolometric luminosity used to define an AGN. For comparison, blue circles show the duty cycle for the entire evolution of the simulations. The data represent the median for all eight simulations in our suite, and error bars denote the interquartile range. AGN at all luminosities are much more likely to be found in the Late and Post merger phases than in the Early phase, and AGN with luminosities of at least  $10^{43}\ \text{erg s}^{-1}$  are active for  $\gtrsim 80$  per cent of the Late+Post merger phases.

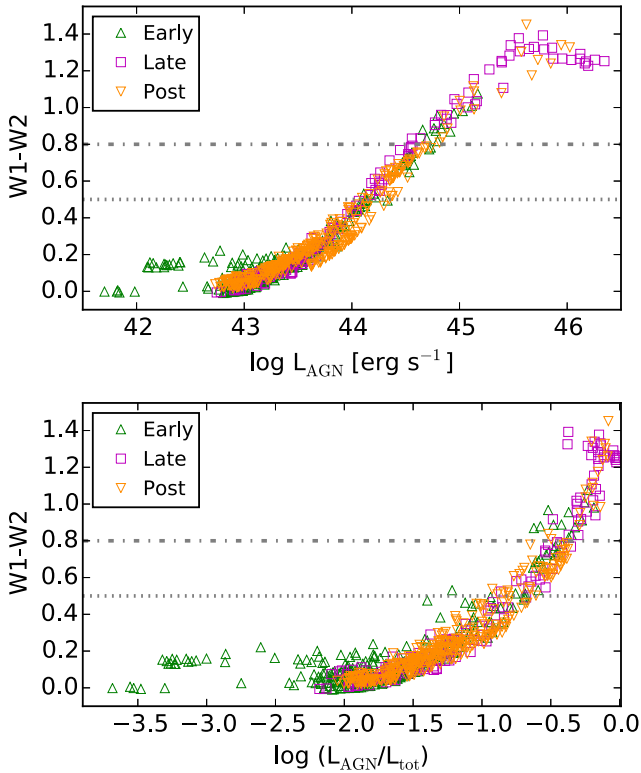
AGN ( $L_{\text{AGN}} > 10^{44-45}\ \text{erg s}^{-1}$ ), the AGN duty cycles are naturally shorter, and they depend more strongly on the merger parameters (particularly the initial gas fraction). The median AGN duty cycle at quasar luminosities ( $L_{\text{AGN}} > 10^{45}\ \text{erg s}^{-1}$ ) is  $\sim 7$  per cent in the Late+Post merger phase, with a maximum of  $\sim 50$  per cent (in simulations 0 and 1). These duty cycles correspond to quasar lifetimes ranging from 10 to 150 Myr in advanced mergers.

We find a strong correlation between AGN luminosity and *WISE* mid-IR colours; as expected, the global W1 – W2 colour is reddest when the AGN dominates the total SED of the galaxy. To quantify this, we define the bolometric AGN fraction as  $f_{\text{AGN}} \equiv L_{\text{AGN}}/L_{\text{tot}}$ , where  $L_{\text{AGN}}$  is the total *intrinsic* bolometric AGN luminosity (summed over both BHs if they have not yet merged) and  $L_{\text{tot}}$  is the total emergent bolometric luminosity of the system. Fig. 4 shows the correlation between  $f_{\text{AGN}}$  and the W1 – W2 colour for all eight fiducial simulations, with each point representing a different snapshot. The correlation has very little scatter across the range of merger environments and merger stages in our simulation suite, demonstrating clearly why mid-IR colour selection is an effective AGN diagnostic. The data are also colour coded by merger stage, which illustrates that the highest luminosities (and reddest mid-IR colours) occur in the Late- and Post-Merger phases.

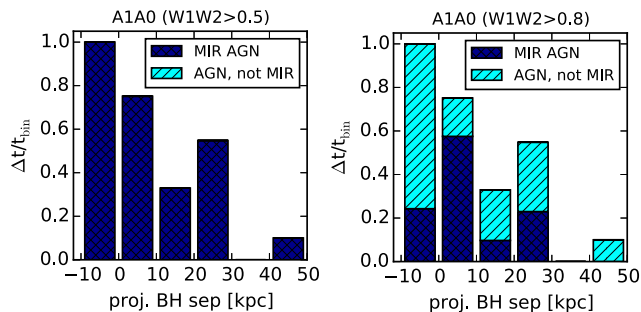
We see that  $f_{\text{AGN}} \gtrsim 0.3$  is needed to achieve W1 – W2  $> 0.8$ ; that is, the BH must contribute at least 30 per cent to the total bolometric luminosity to be selected as a mid-IR AGN via a single-colour W1 – W2  $> 0.8$  cut. Lower luminosity AGN that would be missed by a W1 – W2  $> 0.8$  cut still contribute a non-negligible fraction of the total luminosity,  $f_{\text{AGN}} \gtrsim 0.1$ ; these systems with  $f_{\text{AGN}} = 0.1-0.3$  have bolometric luminosities in the range  $L_{\text{AGN}} = 6 \times 10^{43} - 6 \times 10^{44}\ \text{erg s}^{-1}$ . Fig. 4 shows that a more lenient W1 – W2  $> 0.5$  colour cut is sensitive to these moderate-luminosity AGN with  $f_{\text{AGN}} > 0.1$ .

Fig. 5 shows the bolometric and mid-IR AGN duty cycles versus projected BH separation for an example simulation (A1A0), where we take  $L_{\text{AGN}} > 10^{44}\ \text{erg s}^{-1}$ . Here, the AGN duty cycle is calculated





**Figure 4.** In the top panel, the simulated *WISE*  $W1 - W2$  colour is shown versus the AGN luminosity for all eight mergers in the simulation suite. In the bottom panel, the AGN luminosity is shown as a fraction of the total bolometric luminosity in the host ( $f_{\text{AGN}} = L_{\text{AGN}}/L_{\text{tot}}$ ). Each point represents a single snapshot, where the *WISE* colour is averaged over all sightlines. The green upward triangles, red squares, and orange downward triangles denote snapshots in the Early-, Late-, and Post-Merger stages, respectively. The grey dotted and dot-dashed lines denote the single-colour cuts considered in this work ( $W1 - W2 > 0.5$  and  $0.8$ , respectively). The global *WISE* colour is strongly correlated with the AGN contribution to the total host SED, with the reddest  $W1 - W2$  colours produced predominantly by luminous AGN in the Late- and Post-Merger stages.



**Figure 5.** For the A1A0 merger simulation, the AGN lifetime is shown versus projected separation, as a fraction of the total time in each separation bin (i.e. the AGN duty cycle). The total height of each bar (dark blue + cyan) is the total (bolometric) AGN lifetime, where an AGN is defined as  $L_{\text{AGN}} > 10^{44} \text{ erg s}^{-1}$ . The dark blue portion of the bar denotes the *WISE*-selected AGN duty cycle for  $W1 - W2 > 0.5$  (left-hand panel) and for  $W1 - W2 > 0.8$  (right-hand panel). The leftmost ‘negative separation’ bin denotes the Post-merger phase, capped at 100 Myr. AGN are preferentially triggered in the Late- and Post-Merger stages (the two leftmost bins), and virtually all of this activity would be captured with a  $W1 - W2 > 0.5$  colour cut. With  $W1 - W2 > 0.8$ , however, these mid-IR AGN would be identified only about half of the time.

as a fraction of the total time spent in each separation bin, and mid-IR AGN duty cycles are shown for *WISE*  $W1 - W2 > 0.5$  and  $0.8$ . As we saw in Figs 3 and 4, AGNs (including mid-IR AGN) are more likely to be found in the late stages of the merger. However, Fig. 5 makes clear that a more stringent *WISE*  $W1 - W2 > 0.8$  selection criterion will miss most of the AGN lifetime in advanced mergers. In this example, a gas-rich, major merger where an AGN is active for all of the first 100 Myr of post-merger evolution, only 25 percent of this AGN phase would be selected with  $W1 - W2 > 0.8$ , versus 100 percent for  $W1 - W2 > 0.5$ . We find similar results if we compare these colour cuts for even lower luminosity AGN,  $L_{\text{AGN}} > 10^{43} \text{ erg s}^{-1}$ . Conversely, the most luminous AGNs ( $L_{\text{AGN}} > 10^{45} \text{ erg s}^{-1}$ ) are selected with high completeness with either single-colour criterion, reflecting the fact that more stringent criteria will preferentially select the highest luminosity AGN.

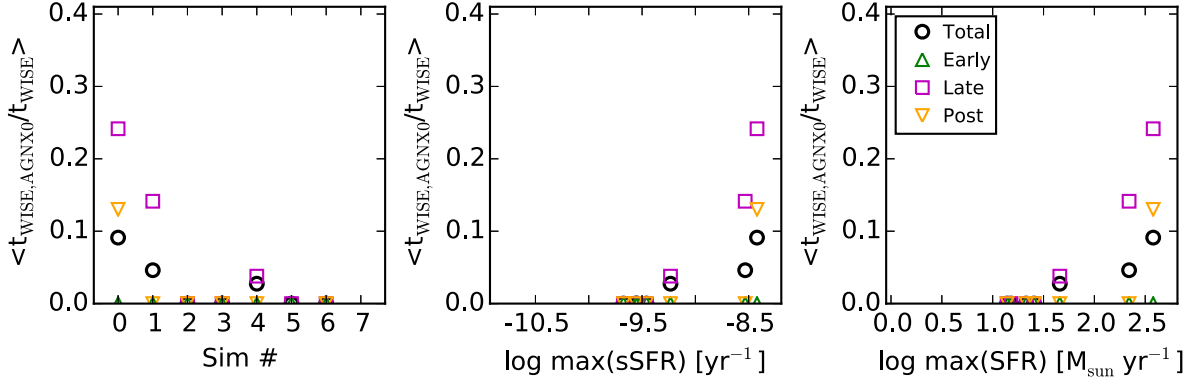
### 3.2 Reliability of mid-IR AGN selection in mergers

The primary concern with using a less stringent, more complete mid-IR selection criterion is the probable trade-off in the reliability of AGN selection, with significant contamination from star-forming galaxies. Specifically, dust-enshrouded nuclear starbursts of sufficient intensity may mimic the red mid-IR SED slope characteristic of AGN-heated dust near its sublimation temperature ( $\sim 1500 \text{ K}$ ). (At high redshift,  $z \gtrsim 1-1.5$ , the optical/near-IR peak of the stellar SED also begins to contaminate the mid-IR colours; we discuss results for  $z > 0$  in more detail below.) By comparing the set of fiducial *SUNRISE* simulations with the corresponding AGNx0 simulations, in which the AGN luminosity is artificially set to zero, we can directly quantify the contamination of the mid-IR SED by star-formation-heated dust.

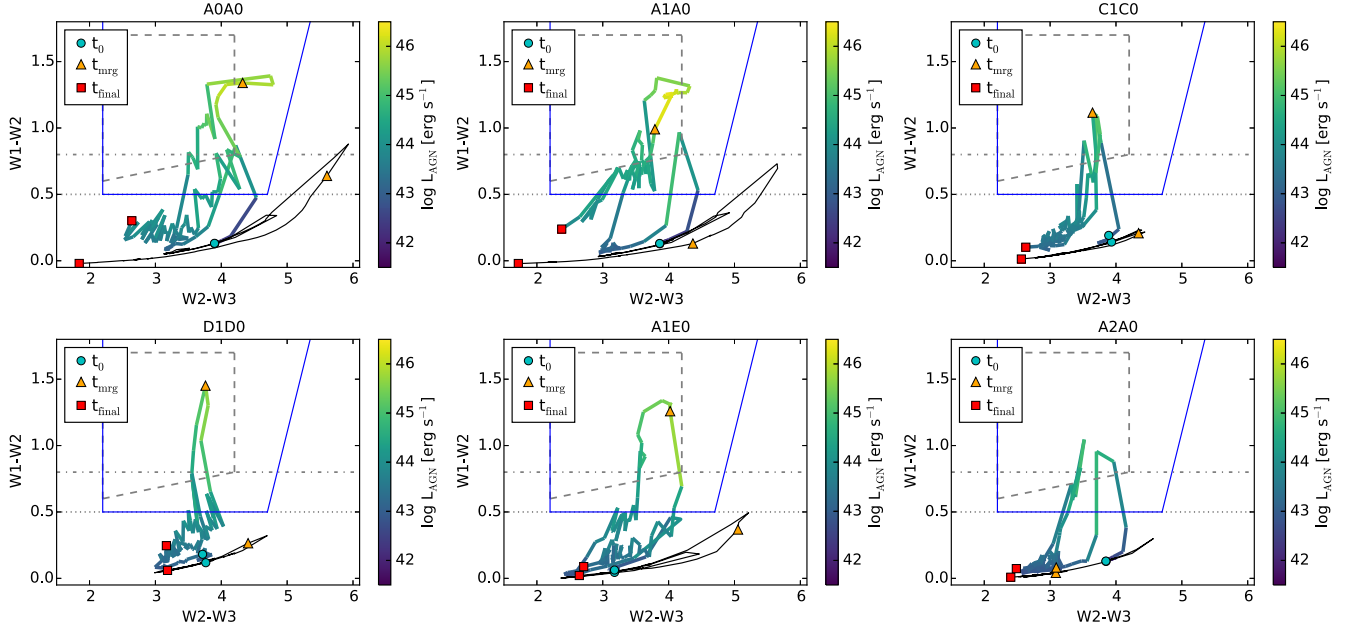
Fig. 6 compares directly the mid-IR-selected lifetime of the fiducial versus the AGNx0 runs, thereby quantifying the contamination of mid-IR colours by star formation. We calculate the *WISE*-selected lifetime (for each merger stage) in each AGNx0 simulation relative to the *WISE* lifetime in the corresponding fiducial simulation ( $t_{\text{WISE, AGNx0}}/t_{\text{WISE}}$ ). This measures the fraction of the mid-IR AGN lifetime when star formation alone would have been capable of producing the red *WISE* colours, if the AGN were not present. This fraction is shown in Fig. 6 for  $W1 - W2 = 0.5$ , for each simulation by number, and also versus the maximum SFR and sSFR in each simulation.

Only 3/8 merger simulations ever achieve  $W1 - W2 > 0.5$  in the AGNx0 run, and the median contamination fraction ( $t_{\text{WISE, AGNx0}}/t_{\text{WISE}}$ ) is zero for any single-colour cut of  $W1 - W2 \geq 0.4$ . One of these simulations (4, or A1E0) has  $t_{\text{WISE, AGNx0}}/t_{\text{WISE}} \sim 4$  percent in the Late-merger phase; this corresponds to a single simulation snapshot. The other two simulations with non-zero contamination have very high maximum SFRs of  $200-400 M_{\odot} \text{ yr}^{-1}$  and sSFRs of  $\sim 10^{-8.5} \text{ yr}^{-1}$ . Empirical studies of galaxy pairs and galaxy mergers have shown that, at least at low to moderate redshift, such extreme starbursts are found almost exclusively in major mergers that are ULIRGs (e.g. Sanders & Mirabel 1996; Veilleux et al. 2009). These are short-lived and rare in the local Universe compared to less-extreme luminous IR galaxies (LIRGs,  $L_{\text{IR}} > 10^{11} L_{\odot}$ ) with modest SFR enhancements (Scudder et al. 2012; Ellison et al. 2013a). Even in these extreme starbursts, red  $W1 - W2 > 0.5$  colours could only be produced by star formation alone for  $\sim 15-25$  percent of the *WISE* lifetime in the Late-merger stage, or 5–9 percent of the total *WISE* lifetime.

We now examine the evolution of the fiducial and AGNx0 simulations in *WISE* colour-colour space. Fig. 7 shows the evolution



**Figure 6.** For each simulation, and each merger phase, the ‘contamination’ fraction ( $t_{\text{WISE,AGN}0}/t_{\text{WISE}}$ ) of mid-IR colours by star formation in the AGN $x_0$  simulations is shown. Specifically, this is the amount of time in each AGN $x_0$  simulation for which  $W1 - W2 > 0.5$  are produced solely by star formation heating, relative to the total  $W1 - W2 > 0.5$  lifetime in the corresponding fiducial simulation. All quantities are averaged over seven viewing angles. The left-hand panel shows results by simulation number as defined in Table 2, and the middle and right-hand panels show the contamination versus the maximum sSFR and SFR, respectively. 3/8 simulations have a non-zero lifetime with red  $W1 - W2$  colours from star formation, occurring only during intense starbursts and constituting at most 15–25 per cent of the Late-merger phase. When these starburst-induced red  $W1 - W2$  phases occur in the AGN $x_0$  simulation, the corresponding fiducial simulation *always* has a simultaneously active AGN, so this is not ‘true’ contamination of the  $W1 - W2$  AGN selection.

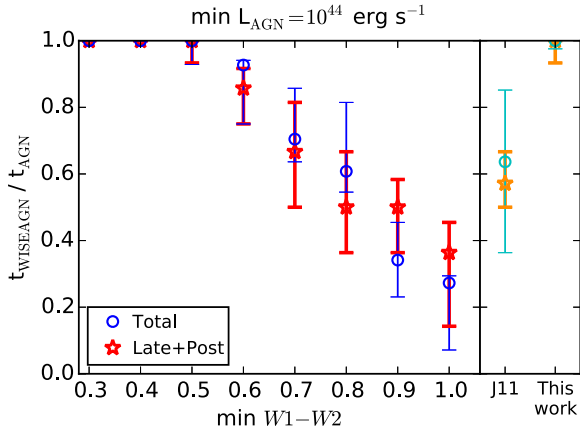


**Figure 7.** The evolution of merging galaxies in  $W1 - W2$  versus  $W2 - W3$  is shown for the gas-rich merger simulations (0–5). The thick colour curves show the evolution for the fiducial simulations, where the colour scale denotes the bolometric AGN luminosity as shown in the colour bar. The thin black curves show the evolution for the corresponding AGN $x_0$  simulations (in which the AGN luminosity is artificially set to zero for the radiative transfer calculation). The cyan circles denote the initial snapshot, the orange triangles denote the time of BH merger, and the red squares denote the final snapshot at the end of the simulation. The grey dotted and dot-dashed lines show single-colour  $W1 - W2$  selection cuts used in the literature and in this work. The grey dashed lines denote the two-colour ‘wedge’ selection criteria of J11, and the blue solid lines denote the two-colour cut motivated by this work (equation 1, Section 3.2). The  $W1 - W2$  colour typically peaks near the time of BH merger, and in the pure starburst (AGN $x_0$ ) simulations it rarely exceeds 0.5. The red  $W2 - W3$  colours produced in purely starburst systems are also clearly distinguishable from the AGN in colour-colour space.

of  $W1 - W2$  versus  $W2 - W3$  colour for four selected mergers, where the colour and grey-scale curves denote the fiducial and AGN $x_0$  simulations, respectively. In the fiducial simulations, the  $W1 - W2$  colour is low at the start of the simulations when the BH accretion rates are low, and then it rapidly reddens during peaks in AGN activity (generally after the galaxies’ first pericentric passage and during their final coalescence; cf. Fig. 1). Stellar and AGN feedback regulate these phases, and by the end of the simulation, the mid-IR colours are again dominated by stellar emission. The

reddest  $W1 - W2$  colours are usually found during the final coalescence of the galaxies (for reference, the time of BH merger is marked with an orange triangle). Note, however, that the elevated  $W1 - W2$  colours during these AGN phases often fall outside the standard  $W1 - W2$  selection cuts. At times, this includes highly obscured ULIRGs ( $L_{\text{IR}} > 10^{12} L_{\odot}$ ), with red  $W2 - W3$  colours that place them outside the two-colour J11 wedge. In the A0A0 simulation, this reduces the completeness of the J11 two-colour selection to 60 per cent in the Late-merger phase. Finally, as in Fig. 6, Fig. 7





**Figure 8.** The median mid-IR AGN selection completeness – that is, the fraction of AGN that would be selected with a given mid-IR colour cut – is shown for all eight mergers in the simulation suite. The left-hand side of the plots show selection completeness versus minimum  $W1 - W2$  for a single-colour cut. The right-hand side of the plots show the same quantities for the two-colour cuts ( $W1 - W2$  and  $W2 - W3$ ) of **J11** and in this work (see Section 3.2 for details). The blue and cyan circles denote the median *WISE* AGN fraction for the whole simulation, and the red and orange stars denote the median fraction for the Late+Post-merger phases only. Error bars denote the interquartile range. For moderate-luminosity AGN ( $L_{\text{bol}} > 10^{44} \text{ erg s}^{-1}$ ), the completeness decreases steadily with stricter  $W1 - W2$  colour cuts. The two-colour cut proposed in this work (Section 3.2) has a median completeness of 100 per cent for moderate-to-high-luminosity, merger-triggered AGN.

makes clear that star formation alone is rarely able to produce red, AGN-like  $W1 - W2$  colours in the host rest frame.

Motivated by our findings, we define a new two-colour cut that leverages the higher completeness of a more lenient  $W1 - W2 > 0.5$  cut for moderate-luminosity AGN. In addition to including many lower luminosity AGN, these selection criteria will identify high-luminosity AGN in ULIRGs that co-exist with starbursts and have red  $W1 - W2$  and  $W2 - W3$ , while excluding pure starburst galaxies with very red  $W2 - W3$  (cf. Fig. 7). We also set a lower bound on  $W2 - W3$  of 2.2 mag, as in **J11**; none of our simulated galaxies have  $W2 - W3$  colours bluer than this at  $z < 1$ . We propose the following criteria (in Vega magnitudes, as used throughout this work) for *WISE* colour selection of AGN:

$$[W1 - W2] > 0.5,$$

$$[W2 - W3] > 2.2,$$

$$\text{and } [W1 - W2] > 2.0 \times [W2 - W3] - 8.9 \quad (1)$$

These criteria are shown in Fig. 7 (blue lines) along with the other mid-IR selection criteria considered in this work. As Fig. 8 shows, this two-colour cut yields very high completeness and reliability (median 100 per cent), for simulated merger-triggered AGN with  $L_{\text{AGN}} > 10^{44} \text{ erg s}^{-1}$ .

As mentioned in Section 2.2, we have also repeated our analysis of mid-IR colour selection for SEDs redshifted up to  $z = 1$  for all simulations, and up to  $z = 4$  for select simulation snapshots. For the redshifted SEDs, we find that contamination of the mid-IR AGN selection by starbursts is negligible for  $z < 1$  in our merger-triggered sample, for all one- and two-colour selection criteria considered in this work except for  $W1 - W2 > 0.5$  (which has at most 5–10 per cent contamination). At  $z \sim 1$ –1.5, the optical/near-IR peak of stellar emission begins to shift into the  $W1$  and  $W2$  bands, greatly increasing contamination of the mid-IR AGN selection criteria (see

also the SED templates of Assef et al. 2013). As our study focuses on lower redshift merging systems, in particular those with  $z \lesssim 0.5$ , where mid-IR AGN selection completeness is highest, the contamination of  $W1 - W2$  in high-redshift systems by star formation does not affect our conclusions based on rest-frame SEDs. Moreover, the  $W2 - W3$  colour of star-forming galaxies becomes very blue ( $< 2.2$ ) at  $2 \lesssim z \lesssim 4$ , such that the two-colour AGN criteria considered in this work do not suffer from such contamination in this redshift range.

Our discussion of the contamination of mid-IR AGN selection by star formation has thus far neglected a critical point: the quantity  $t_{\text{WISE, AGN}0}/t_{\text{WISE}}$  is not actually a true measure of the level of contamination. In reality, star formation and AGN fuelling will often happen simultaneously, especially in gas-rich, late-stage major mergers. Thus, even when star formation heating alone is *sufficient* to produce red  $W1 - W2$  colours, the more pertinent question is how often this happens *when an AGN is not simultaneously present*. In other words, we must calculate  $t_{\text{WISE, AGN}0}/t_{\text{WISE}}$  for *only* the AGN0 snapshots in which the corresponding fiducial snapshot does not simultaneously have an AGN. We denote this new quantity as  $t_{\text{WISE, SF}}/t_{\text{WISE}}$ . For *every* merger simulation, we find  $t_{\text{WISE, SF}}/t_{\text{WISE}} = 0$ . That is, *all* of the merger-triggered starbursts that produce red *WISE* colours are coincident with merger-triggered AGN. Thus, in our simulations there is no ‘true’ contamination of the mid-IR AGN colour selection, even for  $W1 - W2 > 0.5$ , or for the two-colour criterion defined above.

Of course this does not mean that real samples of *WISE*-selected AGN have no contamination from starburst galaxies (e.g. Jarrett et al. 2011; Stern et al. 2012; Mateos et al. 2012, 2013; Assef et al. 2013). As mentioned above, contamination increases sharply for  $1 \lesssim z \lesssim 2$ . We also do not model isolated galaxies, where secular processes may trigger less strongly correlated star formation and AGN fuelling, nor do we consider galaxies with BH masses that deviate substantially from the BH–bulge correlations. Even when star formation and AGN fuelling are correlated, AGN are generally variable on very short times-scales compared to the starburst lifetime (and the time resolution of the simulations); although IR variability time-scales are longer than for optical or X-ray variability, this can decrease the amount of simultaneity. And clearly our finite-simulation suite does not span the entire parameter space of merging systems. But for merger-triggered AGN at  $z \lesssim 1$ , we see that mid-IR colour selection is very effective for the selection criteria considered here, which are more lenient than those typically used in the literature.

### 3.3 Completeness of mid-IR AGN selection in mergers

<sup>44–45</sup>In Fig. 8, we quantify the dependence of AGN selection completeness on the mid-IR colour criteria used, for moderate-to-high-luminosity AGN ( $L_{\text{AGN}} > 10^{44} \text{ erg s}^{-1}$ ). The median fraction of AGN that would be selected with a given single-colour  $W1 - W2$  cut is shown, for all simulations and for the Late+Post-merger phase only. We also show the selection completeness for the two-colour ( $W1 - W2$  and  $W2 - W3$ ) cut of **J11**, as well as the two-colour cut proposed in this work (Section 3.2).

At high AGN luminosities ( $L_{\text{AGN}} > 10^{45} \text{ erg s}^{-1}$ ), *all* mid-IR colour cuts considered are essentially 100 per cent complete in merging systems. The **J11** two-colour criteria will miss some of the most luminous AGN in ULIRGs, where their red  $W2 - W3$  colours push them outside the selection region. The two-colour cut defined in Section 3.2 is chosen to include these luminous ULIRGs.

For moderate-luminosity AGN ( $L_{\text{AGN}} > 10^{44} \text{ erg s}^{-1}$ ), mid-IR selection completeness is much more sensitive to the criteria used. Naturally, completeness decreases with increasingly colour cuts. We see that  $\sim 40\text{--}50$  per cent of AGN with  $L_{\text{AGN}} > 10^{44} \text{ erg s}^{-1}$  will be missed with a  $W1 - W2 > 0.8$  criterion, even in the late stages of gas-rich major mergers. Similarly, the J11 two-colour wedge is  $\sim 55\text{--}65$  per cent complete for merger-triggered AGN at these luminosities. In these cases, the selection completeness for moderate-luminosity AGN varies substantially between different mergers, and between the Early-, Late-, and Post-merger phases. However, no clear systematic trends with merger stage or merger parameters are apparent; Fig. 4 suggests that this variation is driven in large part by the AGN luminosity relative to the host. Gas-rich mergers that produce luminous AGN will typically produce starbursts as well, thereby increasing the total luminosity of the host. Preferential nuclear obscuration during late-stage mergers also means that simultaneous star formation throughout the galaxy can more easily dilute the AGN signatures.

A more lenient AGN selection criterion of  $W1 - W2 > 0.5$  yields a median completeness of 100 per cent for moderate-luminosity AGN ( $L_{\text{AGN}} > 10^{44} \text{ erg s}^{-1}$ ) in our simulated merger sample. As Fig. 8 shows, this high completeness is independent of the merger parameters or merger stage. The two-colour cut proposed in Section 3.2 similarly has a median completeness of 100 per cent for these AGN luminosities.

In addition to this analysis of the rest-frame simulated SEDs, we have calculated the completeness of mid-IR selection for SEDs redshifted up to  $z = 1.0$  (Fig. 9). Here, we have additionally imposed flux limits of  $W1 < 18.50$  and  $W2 < 17.11$ , following Assef et al. (2013) as detailed in the figure caption. We find little difference in the mid-IR selection completeness for our simulated merger sample for  $z \lesssim 0.5$ . For moderate-luminosity AGN, the median completeness drops below 90 per cent at  $z = 0.5$  for the two-colour cut proposed in this work. The completeness for advanced mergers (in the Late+Post phases) drops less sharply with redshift, reflecting the redder colours of these obscured merging nuclei. Completeness for the  $W1 - W2 > 0.8$  criterion is much lower overall, but follows a roughly similar redshift trend.

For luminous AGN ( $> 10^{45} \text{ erg s}^{-1}$ ; lower panels of Fig. 9), both the two-colour cut and the  $W1 - W2 > 0.8$  cut are 100 per cent complete out to  $z = 0.6$ . At higher redshift, the completeness of the stricter single-colour cut drops, while the completeness of the two-colour cut proposed in this work remains at  $\sim 100$  per cent until  $z = 1$ . At  $z = 1$ , the completeness of the two-colour cut is 80 per cent, versus 40–50 per cent for the  $W1 - W2 > 0.8$  cut. Our results are therefore robust for moderate-luminosity AGN at least to  $z \sim 0.5$  and for high-luminosity AGN at  $z < 1$ . Moreover, we see that the two-colour cut is more effective at identifying merger-triggered AGN with high completeness than stricter cuts often used in the literature, even for high-luminosity AGN.

### 3.4 Mid-IR selection of dual AGN

The galaxy merger/AGN connection suggests that simultaneously active dual AGN (AGN pairs with  $\lesssim 1\text{--}10 \text{ kpc}$  separations) are a natural consequence of merger-triggered BH fuelling. Identification of such AGN pairs provides an unambiguous confirmation of a late-stage merger, acting as a sort of ‘clock’ that reveals the merger phase. They can also provide insight about the types of nuclear environments that are especially conducive to rapid BH fuelling. Indeed, dual AGN activation becomes more likely as the merger

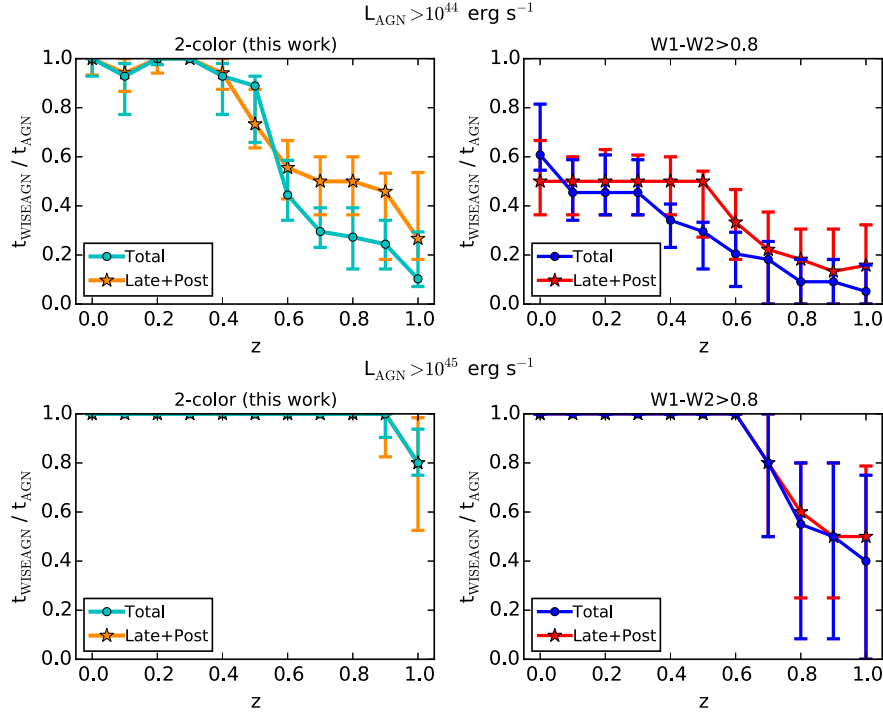
progresses; Ellison et al. (2011) found a statistical excess of paired AGN that increases as pair separation decreases, and other studies have found similar trends (Koss et al. 2012; Satyapal et al. 2017). Although these spatially resolvable AGN pairs are many orders of magnitude outside the gravitational wave (GW) dominated regime of BH inspiral, such objects are also the most readily accessible precursors to the massive BH binaries that will become GW sources.

AGN pairs have been notoriously elusive, however; until recently only a handful were known, most of which were discovered serendipitously. Optical spectroscopic selection of AGN with double-peaked narrow lines yielded a large sample of candidates, and at least 10 per cent of these do indeed appear to host dual AGN (e.g. Comerford et al. 2012; Liu, Shen & Strauss 2012; Comerford et al. 2015). However, the majority of double-peaked narrow lines arise from outflows or other gas kinematics, and the fraction of time that dual AGN induce such features is intrinsically short (e.g. Blecha et al. 2013b; Müller-Sánchez et al. 2015).

Preferential obscuration of optical AGN signatures in mergers is another hindrance to identifying dual AGN via optical spectroscopy. In the hard X-ray-selected AGN sample of Koss et al. (2010), which reveals a much higher incidence of single AGN in mergers than optically-selected AGN samples, an even stronger excess of dual X-ray-selected AGN is found. The dual AGN fraction is highest in the closest BH pairs, which are also the most luminous (Koss et al. 2012). Similarly, the strong enhancement in IR-selected AGN found in late-stage mergers (Satyapal et al. 2014) further suggests that obscured nuclear environments in late-stage mergers are likely sites for dual AGN. X-ray follow-up of these WISE-selected AGN in late-stage mergers has already revealed a high fraction of candidate duals (Satyapal et al. 2017; Ellison et al. 2017), and near-IR coronal line emission detected in some objects provides further evidence for buried AGN in these systems (Satyapal et al. 2017).

Using our merger simulations, we quantify the expected fraction of dual AGN in each merger stage, in terms of both bolometric AGN luminosity and mid-IR colour selection. Fig. 10 shows the fraction of time at each separation when both BHs are simultaneously active, separated by whether or not the dual AGN would be selected via a  $W1 - W2 > 0.5$  WISE colour cut. Also shown is the time for which only a single WISE AGN is present. We see that dual AGN duty cycle is strongly peaked at the smallest BH separations,  $< 3 \text{ kpc}$ . Thus, even though dual BHs spend less time at small separations than at larger separations, late-stage mergers are much more likely to contain dual AGN.

Also notable is the fact that, for  $L_{\text{AGN}} > 10^{44} \text{ erg s}^{-1}$ , mid-IR AGN selection identifies virtually all dual AGN systems. Fig. 11 shows the fraction of the dual AGN lifetime for which the system would be selected as a WISE AGN, for each merger phase. The median completeness of dual AGN selection is near 100 per cent for all single- and two-colour mid-IR selection criteria considered in this work. For the more stringent cuts in particular ( $W1 - W2 > 0.8$  and the J11 two-colour cut), the dual selection completeness is significantly higher than the single-AGN completeness (Fig. 8), which is typically  $\sim 50\text{--}60$  per cent. The higher completeness for WISE selection of dual AGN largely owes to the fact that the conditions conducive to fuelling both BHs simultaneously (i.e. a central, dense reservoir of cold gas) are also conducive to fuelling luminous, obscured AGN, which produce red mid-IR colours. It is perhaps unsurprising, then, that mid-IR colour selection is so effective at identifying merging hosts of dual AGN.



**Figure 9.** Mid-IR AGN selection completeness versus redshift is shown for the two-colour cut proposed in this work (left-hand panels) and for a single-colour  $W1 - W2 > 0.8$  cut (right-hand panels). In all cases, a minimum bolometric luminosity is imposed ( $L_{\text{AGN}} > 10^{44} \text{ erg s}^{-1}$ , top panels and  $> 10^{45} \text{ erg s}^{-1}$ , bottom panels). In addition, for  $z > 0$ , we impose flux limits of  $W1 < 18.50$  and  $W2 < 17.11$ , corresponding to the depth of the *WISE* survey in the NOAO Deep Wide-Field Survey Boötes field at  $S/N > 3$  (Assef et al. 2013). In each panel, the completeness ( $t_{\text{WISEAGN}} / t_{\text{AGN}}$ ) is given for both the total merger simulation and for the Late+Post phase only. The two-colour selection has high completeness ( $\gtrsim 80$  per cent) for moderate-luminosity AGN at  $z \lesssim 0.5$ , much higher than for a stricter  $W1 - W2 > 0.8$  cut. Even for high-luminosity AGN, the completeness of the  $W1 - W2 > 0.8$  cut declines at  $z \gtrsim 0.7$ , while the two-colour cut has  $\sim 100$  per cent completeness almost to  $z \sim 1$ .

For observed samples of *WISE*-selected AGN, we wish to know not only the *completeness* of dual AGN selection, ( $t_{\text{WISE, dual}}/t_{\text{dual}}$ ), but also the fraction of *all WISE* AGN that contain duals ( $t_{\text{WISE, dual}}/t_{\text{WISE}}$ ). Fig. 12 shows this fraction for our simulated sample of merger-triggered AGN, for the total merger and for the Late+Post-merger phases (note that the dual AGN fraction in the Post-merger phase is zero by definition). The fraction  $t_{\text{WISE, dual}}/t_{\text{WISE}}$  varies greatly between merger simulations, particularly for the advanced mergers. In the Late phase, 3/8 simulations (the minor merger A2A0 and the gas-poor simulations) have no dual AGN activity and little to no single AGN activity for  $L_{\text{AGN}} > 10^{44} \text{ erg s}^{-1}$ . Moreover, although a high fraction of advanced mergers contain dual mid-IR AGN (Fig. 10), the Late merger phase prior to BH merger is intrinsically short. Overall, we see that dual mid-IR AGN are actually somewhat more likely to be found at larger separations ( $\sim 10\text{--}30 \text{ kpc}$ ). For the mid-IR selection criteria we consider, we find that a majority (median fraction  $\sim 55\text{--}75$  per cent) of merger-triggered mid-IR AGN contain an AGN pair. And even for the Late+Post-merger phases, where a dual AGN can exist only in the Late phase, we still find that a significant fraction ( $\sim 30\text{--}40$  per cent) of mid-IR AGN are expected to be duals.

The fraction of mid-IR AGN in duals is highest for the more stringent mid-IR selection criteria ( $W1 - W2 > 0.8$  and the J11 two-colour cut), typically  $\sim 60\text{--}85$  per cent of the total *WISE* lifetime. The stricter colour cut yields a slightly higher fraction of dual AGN because of the correlations between AGN luminosity and mid-IR colours, and between AGN luminosity and dual AGN triggering. In other words, these criteria are sensitive to the most luminous AGN,

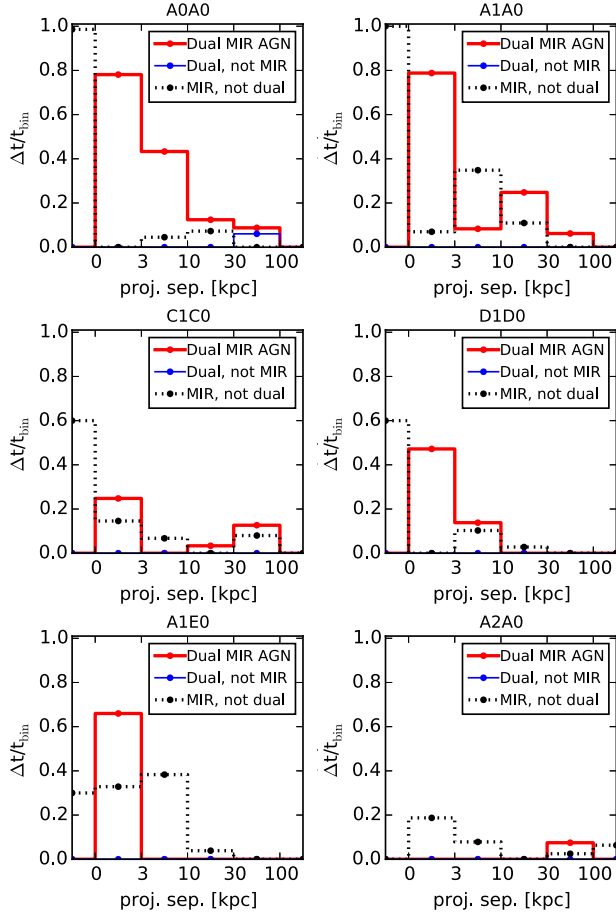
and the same merger-driven dynamics that produce these luminous AGN will increase the probability that both BHs are simultaneously active.

We reiterate that the finite time and spatial resolution of the simulations limits their sensitivity to AGN variability. Uncorrelated, large-amplitude stochastic variability in each AGN would reduce the dual active lifetime compared to our simulations. This is unlikely to be as important in the IR, where AGN are less variable and time-scales are longer, but it could affect the fraction of *WISE*-selected AGN with dual X-ray-detected nuclei, for example. None the less, our results indicate that a significant fraction of merger-triggered mid-IR AGN should contain dual active nuclei, and the early success of follow-up programs in identifying candidates supports this conclusion (Satyapal et al. 2017; Ellison et al. 2017).

Finally, we saw in Fig. 10 that *WISE*-selected dual AGN are most likely to be found at the smallest BH separations. Dual mid-IR AGN (with  $L_{\text{AGN}} > 10^{44} \text{ erg s}^{-1}$ ) are present in up to 80 per cent of mergers with projected separations  $< 3 \text{ kpc}$ . This demonstrates that many dual AGN in *WISE*-selected samples are likely still unresolved. *WISE*-selected AGN in late-stage mergers are therefore promising targets for further follow-up studies with high-resolution X-ray, optical, or IR imaging and resolved spectroscopy.

### 3.5 Merger-driven obscuration of AGN

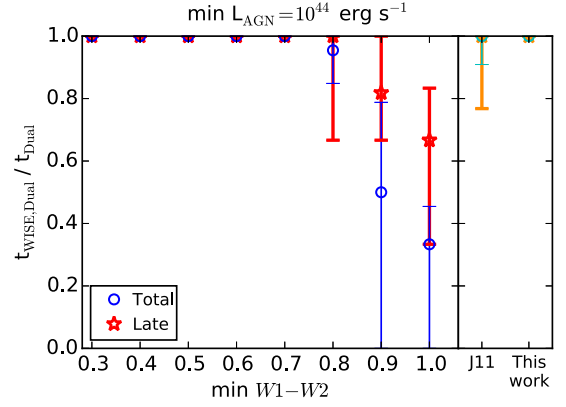
Our finding that mid-IR AGN selection is highly effective in advanced mergers supports the idea that some AGN obscuration is environmentally driven, rather than depending primarily on view-



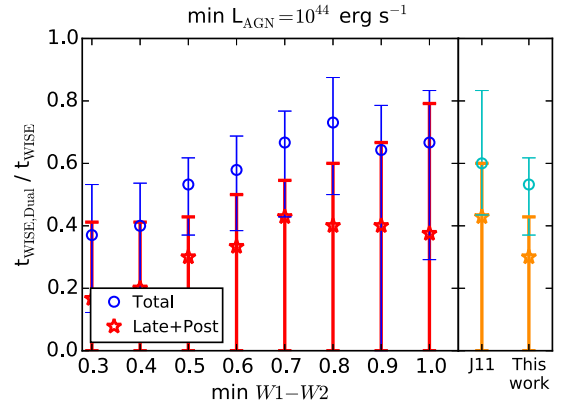
**Figure 10.** For the six gas-rich merger simulations, the mid-IR dual AGN lifetime is shown versus projected separation, as a fraction of the total time in each separation bin (i.e. the dual AGN duty cycle). The leftmost ‘negative’ separation bin denotes the post-merger phase, capped at 100 Myr. Here, a ‘dual AGN’ is defined when each BH has a bolometric luminosity  $> 10^{44}$  erg s $^{-1}$ , and the system is considered a ‘Dual MIR AGN’ if the  $W1 - W2 > 0.5$  criterion is simultaneously met. The thick red histogram denotes the *WISE*-selected dual AGN lifetime, while the dotted black histogram denotes *WISE*-selected AGN where only a *single* BH is active ( $L_{\text{bol}} > 10^{44}$  erg s $^{-1}$ ). The thin blue histogram denotes the phases where a dual AGN is present that would *not* be selected via *WISE*  $W1 - W2 > 0.5$  colours; such a phase appears in only a single bin in the top left plot. Thus, *WISE* colour selection is extremely efficient at identifying dual AGN hosts, particularly at small separations where many such duals may still be unresolved.

ing angle as in standard AGN unification scenarios. Merger-driven dynamics may create significant obscuration on galactic scales, increase the dust covering fraction of the nuclear ‘torus’ on smaller scales, or both. While we do not resolve the AGN torus scale in the simulations, we can quantify the relative amount of galactic-scale obscuration throughout the merger.

In order to obtain an upper limit on the amount of galactic-scale obscuration, we calculate  $N_{\text{H}}$  for the *total* gas density along a given line of sight – i.e. we include the cold-phase gas along with the hot phase, rather than assuming as in the radiative transfer calculations that cold-phase gas has a negligible volume filling factor. Note that, because more of the ISM is in the hot phase during the early-merger stage and long after the merger (particularly in the nuclear regions), we expect the actual evolution in  $N_{\text{H}}$  throughout the merger to be even stronger than Fig. 13 suggests.



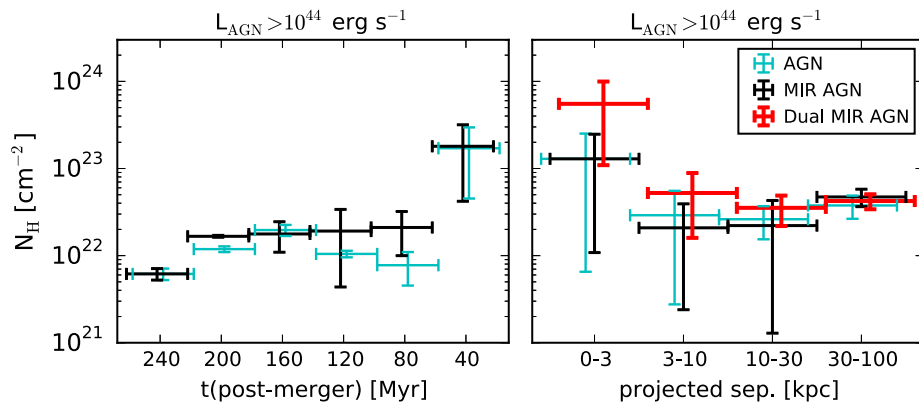
**Figure 11.** The mid-IR selection completeness for dual AGN (pairs of simultaneously active AGN) is shown in a similar manner as the single-AGN completeness in Fig. 8. Specifically, the *WISE*-selected dual AGN lifetime is plotted as a fraction of the total dual AGN lifetime. Blue circles and red stars denote the median fraction for the total merger and the Late phase, respectively, versus the  $W1 - W2$  single-colour cut assumed. The cyan circles and orange stars similarly show  $t_{\text{WISE,Dual}}/t_{\text{Dual}}$  for the two-colour cuts defined in J11 and in this work. Dual AGN are defined as systems where each BH exceeds the minimum luminosity threshold  $L_{\text{AGN}} > 10^{44}$  erg s $^{-1}$ . Mid-IR selection is even more effective at identifying dual AGN than single AGN, with a median completeness of  $\sim 100$  percent for all the selection criteria considered in this work.



**Figure 12.** The fraction of the *total* mid-IR lifetime for which a dual AGN is present ( $t_{\text{WISE,Dual}}/t_{\text{WISE}}$ ) is shown for various mid-IR selection criteria. Blue (cyan) circles denote this fraction for the total merger (for single- and two-colour *WISE* cuts, respectively), and red (orange) stars similarly denote the fraction for the Late+Post-merger phase. Note that the dual AGN lifetime is zero by definition in the Post-merger phase. Although dual mid-IR AGN are efficiently triggered advanced mergers (Fig. 10), the Late merger phase is intrinsically short. The fraction of *WISE* AGN containing duals varies greatly between mergers, particularly in the Late phase, and the short lifetime of the Late phase yields a somewhat higher probability of finding dual AGN at larger separations. None the less, we find that most merger-triggered mid-IR AGN, and  $\sim 30$ –40 per cent of mid-IR AGN in advanced mergers, are actually dual AGN with  $L_{\text{AGN}} > 10^{44}$  erg s $^{-1}$ .

Fig. 13 shows how the AGN column density varies with merger stage. We consider AGN selected via bolometric luminosity ( $L_{\text{AGN}} > 10^{44}$  erg s $^{-1}$ ) and via the two-colour *WISE* selection criterion proposed in this work. Given our findings in Section 3.4, we also examine  $N_{\text{H}}$  for *WISE*-selected *dual* AGN. In all cases, the column density along the line of sight to each BH increases in the final stage of coalescence, for BH separations  $< 3$  kpc. High obscu-





**Figure 13.** The total gas column density ( $N_{\text{H}}$ ) along the AGN line of sight is shown versus merger stage; this provides an upper limit on the amount of galactic-scale obscuration. Prior to the BH merger, the merger evolution is binned by projected BH separation, decreasing from right to left (following the convention used in the literature and throughout this work). After the BH merger, the evolution is binned by the time elapsed since the BH merger, increasing from right to left (such that time evolution can be followed from right to left across both pre- and post-merger panels).  $N_{\text{H}}$  is shown for AGN with bolometric  $L_{\text{AGN}} > 10^{44} \text{ erg s}^{-1}$  (cyan), for mid-IR AGN selected using the *WISE* two-colour cut proposed in this work (black), and for dual mid-IR AGN selected with the same colour cut (red). The column density in each bin is calculated as the time-weighted median over all viewing angles, for all eight fiducial simulations; the error bars denote the median absolute deviation. The AGN column density peaks at  $\gtrsim 10^{23} - 10^{24} \text{ cm}^{-2}$  in the late stages of the merger, at BH separations  $< 3 \text{ kpc}$  and in the first 40 Myr post merger. This demonstrates that AGN are preferentially obscured in advanced mergers, and that environmentally driven obscuration can be significant during mergers even on galactic scales.

ration persists ( $N_{\text{H}} \gtrsim 10^{23} \text{ cm}^{-2}$ ) for the first 40 Myr after the BH merger before declining significantly. The effect is most dramatic for the dual mid-IR AGN, particularly at the smallest BH separations. (Note that in the gas-rich, major merger simulations, nearly all AGN at small nuclear separations are dual mid-IR AGN; Figs 10 and A1). This further emphasizes the need for high-resolution imaging and spectroscopy with *JWST* to reveal kpc-scale dual nuclei in these preferentially obscured systems. There is substantial variation in  $N_{\text{H}}$  between mergers and for different viewing angles, but in all cases, the highest  $N_{\text{H}}$  always occurs in the bin just before or just after the BH merger. In the major, gas-rich merger simulations, we find peak  $N_{\text{H}}$  approaching the Compton-thick regime ( $\gtrsim 10^{24} \text{ cm}^{-2}$ ), which is remarkable given the limitations of finite resolution in our simulations (see also Fig. A1).

If we consider other mid-IR selection criteria, we may expect more stringent colour cuts to preferentially select AGN with higher  $N_{\text{H}}$ , owing to the associations between obscuration, luminosity, and redder mid-IR colours. We find, however, that during peak activity (during final coalescence and just after the merger) the AGN column densities are insensitive to the mid-IR criterion used during final coalescence and just after the merger. This reflects the strong tendency for advanced mergers to contain luminous, obscured (dual) AGN around the time of coalescence. At slightly earlier times (nuclear separations of 3–10 kpc), a more stringent  $W1 - W2 > 0.8$  cut does indeed select single AGN with higher median  $N_{\text{H}}$  as expected (by a factor of a few), though the dual AGN column densities are still insensitive to the colour cut. At slightly later post-merger times (40–80 Myr), however the median  $N_{\text{H}}$  is higher by a factor of 30 for AGN with  $W1 - W2 > 0.8$ .

In five of the major merger simulations, the peak column densities reach the Compton-thick regime ( $> 10^{24} \text{ cm}^{-2}$ ) during coalescence, for our upper limit calculation of  $N_{\text{H}}$ . This is notable, given that the finite resolution of the simulations does not resolve the high-density AGN ‘torus’ region that is expected to dominate the obscuring column in many, if not most, Compton-thick AGN. The degree of correlation between galactic- and torus-scale obscu-

ration remains an important open question. And it bears repeating that our calculation is an upper limit to the amount of galactic-scale obscuration, since we consider the total gas column density. None the less, our results demonstrate that galactic-scale, environmentally driven obscuration can be a significant contributor to AGN obscuration during major mergers. These findings agree well with recent X-ray and near-IR studies of merger-triggered AGN (Ricci et al. 2017; Satyapal et al. 2017; Lansbury et al. 2017; Koss et al. 2018), in which a strong association between mergers and highly obscured and Compton-thick AGN is found in the late-merger stages ( $< 1 - 10 \text{ kpc}$ ).

#### 4 SUMMARY AND CONCLUSIONS

Using hydrodynamic simulations coupled with dust radiative transfer in post-processing, we have studied the evolution of nuclear obscuration and mid-IR AGN signatures during galaxy mergers. We focus in particular on determining the completeness and reliability of mid-IR colour selection for identifying obscured (dual) AGN in advanced mergers. IR selection has revealed large populations of AGN that are hidden at other wavelengths, including a disproportionate number of AGN in ongoing mergers. Understanding the efficacy of mid-IR colour selection in a variety of merger environments and merger stages is therefore crucial for leveraging large surveys, such as the *WISE* All-Sky Survey, to understand the long-debated role of mergers in fuelling AGN.

Mid-IR-selected AGNs are much less sensitive to attenuation by gas and dust than AGN selected in optical or soft X-ray bands. None the less, we find that much of the AGN lifetime is still missed with common mid-IR selection criteria, even in the late stages of gas-rich major mergers. The AGNs must contribute at least 30–50 per cent of the total bolometric luminosity to be detected via a  $W1 - W2 > 0.8$  single-colour cut, such that 40–50 per cent of moderate-luminosity AGN ( $L_{\text{bol}} > 10^{44} \text{ erg s}^{-1}$ ) would not be identified. The J11 two-colour cut (using  $W2 - W3$  in addition to  $W1 - W2$ ) has similar completeness for these AGN.

A less stringent single-colour cut of  $W1 - W2 > 0.5$  selects merger-triggered AGN with much higher completeness (median 100 per cent for  $L_{\text{bol}} > 10^{44} \text{ erg s}^{-1}$ , with all advanced mergers  $> 75$  per cent complete). Moreover, the greater completeness of this relaxed *WISE* colour cut has a minimal trade-off in the reliability of AGN selection for  $z < 1$ . Most of our merger simulations have no contamination at all of the  $W1 - W2 > 0.5$  colour cut from star formation, and only in extreme starbursts ( $\text{sSFR} \gtrsim 10^{-8.5} \text{ yr}^{-1}$ ) does star formation alone cause any significant reddening of the *WISE*  $W1 - W2$  colour. More importantly, because nuclear starbursts and AGN are often cospatial and nearly simultaneous during the final coalescence of merging galaxies, we find that even when intense merger-triggered starbursts mimic AGN-like mid-IR colours, they are always accompanied by an AGN.

We define a new two-colour *WISE* selection criterion based on our study of AGN in merging systems (equation 1). This combines the high completeness of the  $W1 - W2 > 0.5$  single-colour cut with a  $W2 - W3$  cut that excludes pure starbursts but includes highly obscured luminous AGN in ULIRGs. We find that this more lenient selection criterion is well suited for mid-IR studies of AGN, particularly merger-triggered AGN, with high completeness out to  $z \sim 0.5$  and negligible contamination from stellar emission at  $z \lesssim 1$ .

Finally, we show that mid-IR colour selection of merger-triggered AGN is also remarkably effective at identifying systems that contain *dual* AGN. This results from the strong correlation between AGN luminosity and mid-IR colours, and between AGN luminosity and merger stage. In other words, the most luminous AGN produce the reddest mid-IR colours, and they are preferentially found in advanced mergers that have  $\lesssim \text{kpc}$ -scale BH pairs. A  $W1 - W2 > 0.5$  cut, and the two-colour cut defined in this work, select virtually *all* dual AGN (with  $L_{1,2} > 10^{44} \text{ erg s}^{-1}$ ) throughout the merger. A majority of these dual AGN would also be identified with other commonly used selection criteria ( $W1 - W2 > 0.8$  or a two-colour wedge, cf. J11). Moreover, dual AGN have the highest duty cycle at the smallest separations ( $0 < a_{\text{sep}} < 3 \text{ kpc}$ ). This indicates that many dual AGN in *WISE*-selected samples are still unresolved. Indeed, initial X-ray and near-IR spectroscopic follow-up studies have already yielded promising results (Satyapal et al. 2017; Ellison et al. 2017). A complementary adaptive optics follow-up study of hard X-ray-selected AGN also revealed a high fraction of small-scale dual nuclei in late-stage mergers (Koss et al. 2018).

Not all of the mergers in our simulation suite go through a luminous, obscured AGN phase, but this is a ubiquitous outcome in the gas-rich, major mergers. The AGN luminosity and line-of-sight column density to the BHs are well correlated, particularly in the final stages of the galaxies' coalescence when the BH fuelling and obscuration both typically reach their peak. Moreover, because we do not resolve obscuration on 'torus' scales near the AGN, the peak column densities ( $N_{\text{H}} \gtrsim 10^{24} \text{ cm}^{-2}$ ) suggest that galactic-scale obscuration during mergers can contribute significantly to the total attenuation of AGN emission, even in the Compton-thick regime. This stands in contrast to the standard AGN unification picture, in which viewing angle is the dominant factor in the degree of AGN obscuration. These obscured phases coincide with a critical stage of BH growth during which the BH can more than double its mass. This supports the expectation that the most rapid BH fuelling episodes should be preferentially obscured during late-stage mergers, indicating that merger-triggered AGN fuelling can be easily missed. Obscured AGN in late-stage mergers

are therefore ideal targets for further study in the IR, and in particular, imaging and resolved spectroscopy with *JWST* will soon be able to reveal merger dynamics on sub-kpc scales in unprecedented detail.

## ACKNOWLEDGEMENTS

We thank the anonymous referee for their helpful comments that improved the quality of our paper. This work used Extreme Science and Engineering Discovery Environment (XSEDE) resources at the San Diego Supercomputing Center and the Texas Advanced Computing Center through allocation AST130041. XSEDE is supported by National Science Foundation grant number ACI-1548562. LB acknowledges support from a National Science Foundation grant (AST-1715413). GFS appreciates support from a Giacconi Fellowship at the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-26555.

## REFERENCES

- Assef R. J. et al., 2010, *ApJ*, 713, 970  
 Assef R. J. et al., 2013, *ApJ*, 772, 26  
 Blecha L., Cox T. J., Loeb A., Hernquist L., 2011, *MNRAS*, 412, 2154  
 Blecha L., Civano F., Elvis M., Loeb A., 2013a, *MNRAS*, 428, 1341  
 Blecha L., Loeb A., Narayan R., 2013b, *MNRAS*, 429, 2594  
 Cisternas M. et al., 2011, *ApJ*, 726, 57  
 Comerford J. M. et al., 2009, *ApJ*, 698, 956  
 Comerford J. M., Gerke B. F., Stern D., Cooper M. C., Weiner B. J., Newman J. A., Madsen K., Barrows R. S., 2012, *ApJ*, 753, 42  
 Comerford J. M., Pooley D., Barrows R. S., Greene J. E., Zakamska N. L., Madejski G. M., Cooper M. C., 2015, *ApJ*, 806, 219  
 Cox T. J., Dutta S. N., Di Matteo T., Hernquist L., Hopkins P. F., Robertson B., Springel V., 2006, *ApJ*, 650, 791  
 Di Matteo T., Springel V., Hernquist L., 2005, *Nature*, 433, 604  
 Donley J. L. et al., 2012, *ApJ*, 748, 142  
 Donley J. L. et al., 2018, *ApJ*, 853, 63  
 Draine B. T., Li A., 2007, *ApJ*, 657, 810  
 Dwek E., 1998, *ApJ*, 501, 643  
 Ellison S. L., Patton D. R., Mendel J. T., Scudder J. M., 2011, *MNRAS*, 418, 2043  
 Ellison S. L., Mendel J. T., Scudder J. M., Patton D. R., Palmer M. J. D., 2013a, *MNRAS*, 430, 3128  
 Ellison S. L., Mendel J. T., Patton D. R., Scudder J. M., 2013b, *MNRAS*, 435, 3627  
 Ellison S. L., Patton D. R., Hickox R. C., 2015, *MNRAS*, 451, L35  
 Ellison S. L., Secrest N. J., Mendel J. T., Satyapal S., Simard L., 2017, *MNRAS*, 470, L49  
 Fan L. et al., 2016, *ApJ*, 822, L32  
 Gebhardt K. et al., 2000, *ApJ*, 543, L5  
 Goulding A. D. et al., 2018, *PASJ*, 70, S37  
 Groves B., Dopita M. A., Sutherland R. S., Kewley L. J., Fischera J., Leitherer C., Brandl B., van Breugel W., 2008, *ApJS*, 176, 438  
 Gültekin K. et al., 2009, *ApJ*, 698, 198  
 Hani M. H., Sparre M., Ellison S. L., Torrey P., Vogelsberger M., 2018, *MNRAS*, 475, 1160  
 Hayward C. C., Kereš D., Jonsson P., Narayanan D., Cox T. J., Hernquist L., 2011, *ApJ*, 743, 159  
 Hayward C. C., Narayanan D., Kereš D., Jonsson P., Hopkins P. F., Cox T. J., Hernquist L., 2013, *MNRAS*, 428, 2529  
 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

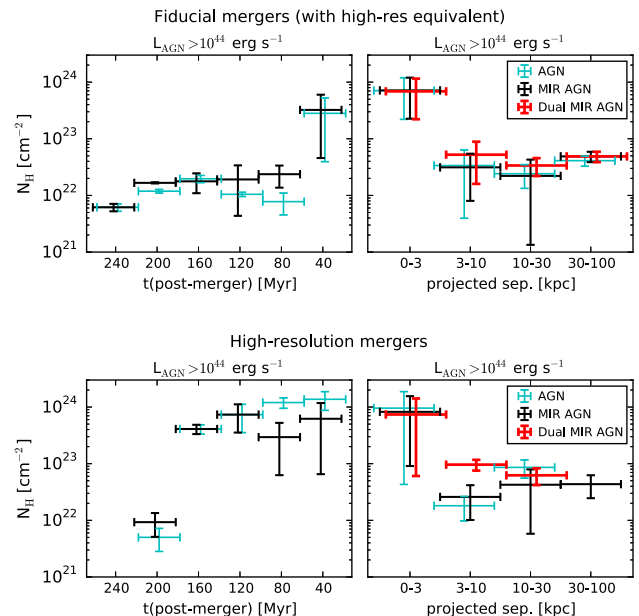
Hopkins P. F., Cox T. J., Kereš D., Hernquist L., 2008, *ApJS*, 175, 390  
 Jarrett T. H. et al., 2011, *ApJ*, 735, 112  
 Jonsson P., 2006, *MNRAS*, 372, 2  
 Jonsson P., Groves B. A., Cox T. J., 2010, *MNRAS*, 403, 17  
 Kocevski D. D. et al., 2012, *ApJ*, 744, 148  
 Kocevski D. D. et al., 2015, *ApJ*, 814, 104  
 Koss M. et al., 2018, Nature in review  
 Koss M., Mushotzky R., Veilleux S., Winter L., 2010, *ApJ*, 716, L125  
 Koss M., Mushotzky R., Treister E., Veilleux S., Vasudevan R., Trippe M., 2012, *ApJ*, 746, L22  
 Kroupa P., 2002, *Science*, 295, 82  
 Lackner C. N. et al., 2014, *AJ*, 148, 137  
 Lacy M. et al., 2004, *ApJS*, 154, 166  
 Lansbury G. B. et al., 2017, *ApJ*, 846, 20  
 Lanz L., Hayward C. C., Zezas A., Smith H. A., Ashby M. L. N., Brassington N., Fazio G. G., Hernquist L., 2014, *ApJ*, 785, 39  
 Laurent O., Mirabel I. F., Charmandaris V., Gallais P., Madden S. C., Sauvage M., Vigroux L., Cesarsky C., 2000, *A&A*, 359, 887  
 Leitherer C. et al., 1999, *ApJS*, 123, 3  
 Liu X., Shen Y., Strauss M. A., Greene J. E., 2010, *ApJ*, 708, 427  
 Liu X., Shen Y., Strauss M. A., Hao L., 2011, *ApJ*, 737, 101  
 Liu X., Shen Y., Strauss M. A., 2012, *ApJ*, 745, 94  
 Lotz J. M., Jonsson P., Cox T. J., Croton D., Primack J. R., Somerville R. S., Stewart K., 2011, *ApJ*, 742, 103  
 Magorrian J. et al., 1998, *AJ*, 115, 2285  
 Mateos S. et al., 2012, *MNRAS*, 426, 3271  
 Mateos S., Alonso-Herrero A., Carrera F. J., Blain A., Severgnini P., Caccianiga A., Ruiz A., 2013, *MNRAS*, 434, 941  
 McConnell N. J., Ma C.-P., 2013, *ApJ*, 764, 184  
 Müller-Sánchez F., Comerford J. M., Nevin R., Barrows R. S., Cooper A. C., Greene J. E., 2015, *ApJ*, 813, 103  
 Narayan R., McClintock J. E., 2008, *New Astron. Rev.*, 51, 733  
 Narayanan D. et al., 2010, *MNRAS*, 407, 1701  
 Ricci C. et al., 2017, *MNRAS*, 468, 1273  
 Richards G. T. et al., 2006, *ApJS*, 166, 470  
 Robertson B., Bullock J. S., Cox T. J., Di Matteo T., Hernquist L., Springel V., Yoshida N., 2006, *ApJ*, 645, 986  
 Roebuck E., Sajina A., Hayward C. C., Pope A., Kirkpatrick A., Hernquist L., Yan L., 2016, *ApJ*, 833, 60  
 Sanders D. B., Mirabel I. F., 1996, *ARA&A*, 34, 749  
 Sanders D. B., Soifer B. T., Elias J. H., Madore B. F., Matthews K., Neugebauer G., Scoville N. Z., 1988a, *ApJ*, 325, 74  
 Sanders D. B., Soifer B. T., Elias J. H., Neugebauer G., Matthews K., 1988b, *ApJ*, 328, L35  
 Satyapal S. et al., 2017, *ApJ*, 848, 126  
 Satyapal S., Ellison S. L., McAlpine W., Hickox R. C., Patton D. R., Mendel J. T., 2014, *MNRAS*, 441, 1297  
 Schawinski K., Simmons B. D., Urry C. M., Treister E., Glikman E., 2012, *MNRAS*, 425, L61  
 Scudder J. M., Ellison S. L., Torrey P., Patton D. R., Mendel J. T., 2012, *MNRAS*, 426, 549  
 Secrest N. J., Dudik R. P., Dorland B. N., Zacharias N., Makarov V., Fey A., Frouard J., Finch C., 2015, *ApJS*, 221, 12  
 Silverman J. D. et al., 2011, *ApJ*, 743, 2  
 Snyder G. F., Hayward C. C., Sajina A., Jonsson P., Cox T. J., Hernquist L., Hopkins P. F., Yan L., 2013, *ApJ*, 768, 168  
 Springel V., 2005, *MNRAS*, 364, 1105  
 Springel V., Hernquist L., 2003, *MNRAS*, 339, 289  
 Stern D. et al., 2005, *ApJ*, 631, 163  
 Stern D. et al., 2012, *ApJ*, 753, 30  
 Tacconi L. J. et al., 2008, *ApJ*, 680, 246  
 Urrutia T., Lacy M., Becker R. H., 2008, *ApJ*, 674, 80  
 Urry C. M., Padovani P., 1995, *PASP*, 107, 803  
 Veilleux S. et al., 2009, *ApJS*, 182, 628  
 Veilleux S., Kim D.-C., Sanders D. B., 2002, *ApJS*, 143, 315  
 Villforth C. et al., 2014, *MNRAS*, 439, 3342  
 Villforth C. et al., 2017, *MNRAS*, 466, 812  
 Vito F. et al., 2018, *MNRAS*, 474, 4528

Weston M. E., McIntosh D. H., Brodwin M., Mann J., Cooper A., McConnell A., Nielsen J. L., 2017, *MNRAS*, 464, 3882  
 Wyithe J. S. B., Loeb A., 2003, *ApJ*, 595, 614

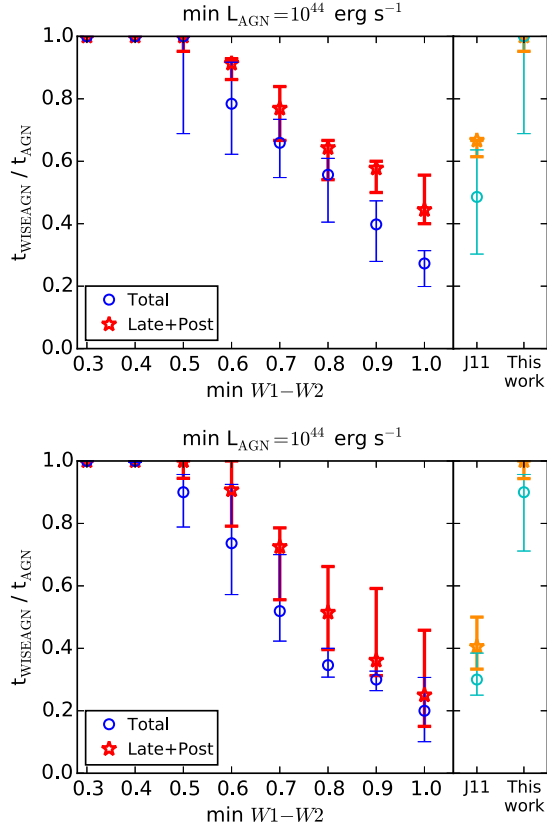
## APPENDIX A: NUMERICAL RESOLUTION

For a subset of four gas-rich mergers (denoted in boldface in Table 2), we run simulations with  $10\times$  higher mass resolution and  $10^{1/3}\times$  higher spatial resolution. Overall, we find that the qualitative evolution of key quantities is unchanged at higher resolution, and the total amount of star formation during each simulation is consistent to within 10 per cent. The peak SFR, peak AGN luminosity, peak  $N_{\text{H}}$ , and total BH mass accreted are consistent within a factor of  $\sim 3\text{--}4$ ; the largest difference in peak  $L_{\text{AGN}}$  occurs in the D1D0 simulation, where the peak AGN luminosity is six times lower in the high-resolution version. These modest variations are to be expected given the stochastic nature of BH accretion and the fluctuations in instantaneous quantities in a highly dynamic environment. For the most part, they do not show any systematic trends.

We do find slightly higher AGN column densities before the merger in the high-resolution runs, and significantly higher AGN column densities after the merger. This is largely because, in each case, we have calculated  $N_{\text{H}}$  along an aperture that corresponds to the effective spatial resolution of the simulation: 136 and 64 pc for the fiducial- and high-resolution simulations, respectively. In other words, a smaller (and thus denser) nuclear region is being resolved in the high-resolution simulations. This indicates that AGN may



**Figure A1.** Similar to Fig. 13, the top and bottom plots show the evolution of nuclear column density  $N_{\text{H}}$  (along the line of sight) to the AGN for the four fiducial-resolution counterparts to the high-resolution runs, and for the high-resolution simulations, respectively. As before, time evolution proceeds from right to left in these plots. A  $2.2\times$  smaller aperture is used for the high-resolution  $N_{\text{H}}$  calculation, to reflect the higher spatial resolution in both the GADGET and SUNRISE simulations. Although the AGN in the high-resolution runs have higher  $N_{\text{H}}$  after the BH merger for this subsample of gas-rich major mergers, overall the nuclear column density is consistent between the fiducial- and high-resolution simulations, within the variation between individual simulations, and the differences in  $N_{\text{H}}$  do not affect our conclusions.



**Figure A2.** As in Fig. 8, the completeness of various mid-IR AGN selection criteria is shown, for the high-resolution simulations (bottom panel) and their fiducial-resolution counterparts (top panel). Completeness is slightly lower on average in the higher resolution simulations, owing to a small increase in dust self-absorption around buried AGN. But the mid-IR selection completeness is consistent between the high- and fiducial-resolution simulations, within the level of variation between individual simulations.

be even more preferentially obscured in mergers than our fiducial results suggest. However, when we perform similar analysis for the BH column densities at all time-steps (not just the AGN), we find that the differences in  $N_H$  at higher resolution are subdominant to the variation with galaxy merger parameters. Moreover, we find that our conclusions in this work do not depend on these variations in  $N_H$  with simulation resolution.

Fig. A2 shows the completeness of mid-IR selection criteria as in Fig. 8, for high resolution versus fiducial resolution. For the more stringent  $W1 - W2$  colour cuts, the high-resolution simulations have marginally lower median completeness, primarily because the higher  $N_H$  increases the self-absorption of emission from the hottest, AGN-heated dust. The slight increase in dust self-absorption around the buried AGN at higher resolution produces excess emission at longer mid-IR wavelengths, and thus redder  $W2 - W3$  colours, which slightly lowers completeness for the two-colour selection criteria as well. This is a very small effect, however; here again the differences between high and fiducial resolution are within the level of variation between simulations.

Even in the absence of a buried AGN, the high-resolution AGNx0 simulations produce slightly more compact nuclear starbursts, which result in longer lifetimes with  $W1 - W2 > 0.5$  from star formation alone. The fiducial- and high-resolution results are still consistent within the level of variation between simulations, and the median ‘contamination’ for  $W1 - W2 > 0.5$  is  $\sim 0.1$  in either case. (Note that the simulation subset used in our resolution study includes two of the three fiducial mergers that exceed  $W1 - W2 = 0.5$ , so this sample has higher median contamination than the full fiducial suite. Most importantly, the ‘true’ contamination – i.e. the amount of mid-IR contamination by star formation in the AGNx0 simulations when an AGN is *not* simultaneously present in the fiducial runs – is still zero for all selection criteria considered in this work.

This paper has been typeset from a  $\text{\LaTeX}$  file prepared by the author.