Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

Mon. Not. R. Astron. Soc. **424,** 951–970 (2012) doi:10.1111/j.1365-2966.2012.21254.x

How to distinguish starbursts and quiescently star-forming galaxies: the 'bimodal' submillimetre galaxy population as a case study

Christopher C. Hayward,[1,2]★ Patrik Jonsson,[2] Dušan Kereš,[3,4]† Benjamin Magnelli,[5] Lars Hernquist[2] and T. J. Cox[6]

[1]*Heidelberger Institut für Theoretische Studien, Schloss–Wolfsbrunnenweg 35, 69118 Heidelberg, Germany*
[2]*Harvard–Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA*
[3]*Department of Physics, Center for Astrophysics and Space Sciences, University of California at San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA*
[4]*Department of Astronomy and Theoretical Astrophysics Center, University of California Berkeley, Berkeley, CA 94720, USA*
[5]*Max–Planck–Institut für Extraterrestrische Physik, Postfach 1312, 85741 Garching, Germany*
[6]*Carnegie Observatories, 813 Santa Barbara Street, Pasadena, CA 91101, USA*

Accepted 2012 May 4. Received 2012 April 23; in original form 2012 March 5

ABSTRACT

In recent work, we have suggested that the high-redshift ($z \sim 2$–4) bright submillimetre galaxy (SMG) population is heterogeneous, with major mergers contributing both at early stages, where quiescently star-forming discs are blended into one submm source ('galaxy-pair SMGs'), and at late stages, where mutual tidal torques drive gas inflows and cause strong starbursts. Here we combine hydrodynamic simulations of major mergers with 3D dust radiative transfer calculations to determine observational diagnostics that can distinguish between quiescently star-forming SMGs and starburst SMGs via integrated data alone. We fit the far-infrared (FIR) spectral energy distributions of the simulated galaxies with the optically thin single-temperature modified blackbody, the full form of the single-temperature modified blackbody and a power-law temperature distribution model. The effective dust temperature, $T_{\rm d}$, and power-law index of the dust emissivity in the FIR, β, derived can significantly depend on the fitting form used, and the intrinsic β of the dust is not recovered. However, for all forms used here, there is $T_{\rm d}$ above which almost all simulated galaxies are starbursts, so a $T_{\rm d}$ cut is very effective at selecting starbursts. Simulated merger-induced starbursts also have higher $L_{\rm IR}/M_{\rm gas}$ and $L_{\rm IR}/L_{\rm FUV}$ than quiescently star-forming galaxies and lie above the star formation rate–stellar mass relation. These diagnostics can be used to test our claim that the SMG population is heterogeneous and to observationally determine what star formation mode dominates a given galaxy population. We comment on applicability of these diagnostics to ultraluminous IR galaxies (ULIRGs) that would not be selected as SMGs. These 'hot-dust ULIRGs' are typically starburst galaxies lower in mass than SMGs, but they can also simply be SMGs observed from a different viewing angle.

Key words: radiative transfer – stars: formation – galaxies: high-redshift – galaxies: starburst – infrared: galaxies – submillimetre: galaxies.

1 INTRODUCTION

1.1 The two modes of star formation

Star formation is one of the fundamental processes driving galaxy formation: it depletes the gas content of galaxies, enriches the interstellar medium (ISM) with metals and deposits energy and momentum via supernovae, stellar winds and radiation pressure. Furthermore, the light emitted by stars encodes much information about the current physical properties of a galaxy and the galaxy's formation history. Thus understanding star formation is crucial for understanding galaxy formation and evolution.

An important step towards understanding the star formation processes that built up galaxies across cosmic time is determining where and when most stars are formed, be it in disc galaxies or in starbursts triggered by, for example, galaxy mergers, which are short-lived but can dramatically alter a galaxy's properties. An increasing amount of observational evidence supports the notion that there are two

★E-mail: christopher.hayward@h-its.org
†Hubble Fellow.

modes of star formation (e.g. Elbaz et al. 2011; Rodighiero et al. 2011; Wuyts et al. 2011a,b; Nordon et al. 2012), typically referred to as quiescently star forming or quiescent[1] (that occurring in normal disc galaxies) and starburst (found in unstable discs and merging galaxies at the first passage and coalescence, though whether a starburst is induced in the latter depends on factors such as gas content, orbit and mass ratio of the progenitors; e.g. Cox et al. 2008). One clear difference is that gas depletion time-scales of starbursts are significantly lower than those of quiescently star-forming galaxies. Star formation in starbursts tends to be dominated by the nuclear regions, whereas quiescent star formation is more extended. Additionally, starbursts may obey a different global Kennicutt–Schmidt (KS) relation (Schmidt 1959; Kennicutt 1998a) from quiescently star-forming disc galaxies: Daddi et al. (2010) and Genzel et al. (2010) argue that the normalization of the KS relation for starbursts is ∼4–10 times greater than that for quiescently star-forming discs. However, this conclusion strongly depends on the bimodality of the values adopted for the CO–H_2 conversion factor, so the data in fact may be consistent with a single KS law normalization (Narayanan et al. 2012). The ratio of infrared (IR) luminosity to molecular gas mass, which is a measure of the global star formation efficiency (SFE) of the system, is larger in starbursts than in quiescently star-forming galaxies, though the magnitude of the difference is also sensitive to the CO–H_2 conversion factor. Furthermore, the relationship between star formation rate (SFR) and dust mass may also show a bimodal behaviour (da Cunha et al. 2010).

At a given redshift, most galaxies lie on a tight relation between SFR and stellar mass ($M_\star$; Daddi et al. 2007; Noeske et al. 2007a,b; Rodighiero et al. 2010; Karim et al. 2011). The relation arises because star formation is supply limited, so, on average, SFR correlates well with cosmological gas accretion rates, which are well correlated with halo mass (Kereš et al. 2005, 2009; Faucher-Giguère et al. 2011). In this picture, starbursts are transient events that cause a galaxy to move significantly above the SFR–$M_\star$ relation for a short (∼50–100 Myr) time. During the burst the gas is rapidly consumed, the SFR declines sharply, and the galaxy returns to the SFR–$M_\star$ relation or is quenched (i.e. drops significantly below the relation), depending on factors such as merger mass ratio, orbit and gas fraction.

Though there is some observational support for two star formation modes, the underlying physics is not fully understood. Thus further detailed observations are crucial. However, it can be difficult to observationally determine which mode of star formation dominates a given galaxy population; this complicates efforts to understand the underlying physics. This is especially a problem at high redshift because of stricter observational limitations, and lessons learned from the local Universe may not apply to high-redshift galaxies. For example, at high redshift gas accretion rates are significantly higher than locally (e.g. Kereš et al. 2005), so gas fractions (Erb et al. 2006; Tacconi et al. 2006, 2010; Daddi et al. 2010) and SFRs (Daddi et al. 2007; Noeske et al. 2007a,b) of galaxies at fixed galaxy mass increase rapidly with redshift. Consequently, even a typical star-forming galaxy at $z \sim 2$ can reach ultraluminous IR galaxies (ULIRG luminosities; e.g. Daddi et al. 2005, 2007; Hopkins et al. 2008, 2010; Dannerbauer et al. 2009). It would thus be useful to have simple observational diagnostics that can be used to determine which mode of star formation dominates a given galaxy or galaxy population. This is one of the goals of this paper.

1.2 The 'bimodality' of the SMG population

Submillimetre galaxies (SMGs; Smail, Ivison & Blain 1997; Barger et al. 1998; Hughes et al. 1998; Eales et al. 1999; see Blain et al. 2002, for a review) are a class of high-redshift (median $z \sim 2.3$–2.6; Chapman et al. 2005; Yun et al. 2012) galaxies notable for their extreme luminosities (bolometric luminosity $L_{\rm bol} \sim 10^{12}$–$10^{13}\,{\rm L}_\odot$; e.g. Kovács et al. 2006; Magnelli et al. 2010, 2012), almost all of which is emitted in the IR. Since they seem to be powered by star formation rather than active galactic nuclei (AGN; Alexander et al. 2005a,b, 2008; Menéndez-Delmestre et al. 2007, 2009; Valiante et al. 2007; Pope et al. 2008; Younger et al. 2008, 2009c), they have inferred SFRs of $\sim 10^2$–$10^3\,{\rm M}_\odot\,{\rm yr}^{-1}$ (e.g. Coppin et al. 2008), much greater than those of even the most extreme local galaxies.

It has long been known that during major mergers tidal torques drive significant amounts of gas inwards, resulting in a strong burst of star formation (Hernquist 1989; Barnes & Hernquist 1991, 1996; Mihos & Hernquist 1996). Locally, ULIRGs (defined by $L_{\rm IR} > 10^{12}\,{\rm L}_\odot$) are exclusively late-stage major mergers (e.g. Sanders & Mirabel 1996; Veilleux, Kim & Sanders 2002; Lonsdale, Farrah & Smith 2006), so the strong gas inflows induced during the coalescence stage of a major merger seem necessary to power the most luminous and rapidly star-forming galaxies.

The identification of ULIRGs with late-stage major mergers in the local Universe has caused many researchers to suspect that SMGs, some of the most IR-luminous galaxies at high redshift, are also late-stage major mergers. There is much observational evidence that supports this picture (e.g. Ivison et al. 2002, 2007, 2010; Chapman et al. 2003; Neri et al. 2003; Smail et al. 2004; Swinbank et al. 2004; Greve et al. 2005; Tacconi et al. 2006, 2008; Bouché et al. 2007; Biggs & Ivison 2008; Capak et al. 2008; Younger et al. 2008, 2010; Iono et al. 2009; Engel et al. 2010; Riechers et al. 2011a,b; Bussmann et al. 2012; Magnelli et al. 2012). Furthermore, by combining hydrodynamic simulations with dust radiative transfer (RT), we have shown that simulated major mergers have observed 850-μm fluxes and typical spectral energy distributions (SEDs; Narayanan et al. 2010a; Hayward et al. 2011a; but cf. Chakrabarti et al. 2008), stellar masses (Michałowski et al. 2012) and CO properties (Narayanan et al. 2009) consistent with observed SMGs. Semi-analytic models typically also find that merger-induced starbursts (though not necessarily major mergers, as minor-merger-induced starbursts dominate in some models) account for the bulk of the SMG population (Baugh et al. 2005; Fontanot et al. 2007; Swinbank et al. 2008; Lo Faro et al. 2009; Fontanot & Monaco 2010; González et al. 2011; but cf. Granato et al. 2004). However, Davé et al. (2010) have claimed that there are not enough major mergers to account for the observed SMG population. Instead, they argue that a significant fraction of the population must be massive discs fuelled by smooth accretion and minor mergers.

In Hayward et al. (2011a, hereafter H11), we suggested a modification to the canonical picture, arguing that SMGs are not purely late-stage major mergers but rather a heterogeneous population, composed of the following: (1) late-stage major mergers undergoing strong starbursts; (2) early-stage major mergers ('galaxy-pair SMGs') which are quiescently forming stars; (3) physically unrelated galaxies blended into one submm source (Wang et al. 2011); and (4) isolated disc galaxies and minor mergers, which only contribute significantly at the fainter end of the population. One reason early-stage mergers also contribute is that observed submm flux

[1] Confusingly, the term 'quiescent' is also used to refer to galaxies that have essentially no ongoing star formation; here the term 'quiescent' always means 'quiescently star forming'.

increases rather weakly with SFR, and the starburst mode is significantly less efficient at boosting submm flux than the quiescent mode. Physically, submm flux scales more weakly in starbursts for two reasons. (1) In high-redshift mergers significant star formation occurs before the starburst induced at the coalescence of the two discs. During the starburst the 'old' stars formed pre-burst can provide a non-negligible contribution to $L_{\rm IR}$. For example, for the merger shown in the right-hand panel of fig. 1 of H11, the stars formed pre-burst account for $\sim$1/7 of the total IR luminosity at the time of the burst. Consequently, $L_{\rm IR}$ does not scale linearly with the instantaneous SFR during the burst because of the dust heating from the older stars (see e.g. section 2.5 of Kennicutt 1998b and section 3.1 of H11). (2) Driven primarily by the strong drop in dust mass during the merger, the effective dust temperature of the SED increases sharply during the starburst. This effect mitigates the increase in submm flux caused by the increased luminosity. These two effects result in a very weak scaling ($S_{850} \propto {\rm SFR}^{0.1}$, compared to $S_{850} \propto {\rm SFR}^{0.4}$ for the quiescently star-forming galaxies). For the example in H11, an increase in SFR of $\sim$16 times causes a boost of $\lesssim$2 times in the submm flux.

The other reason why early-stage mergers contribute significantly is that the large [full width at half-maximum (FWHM) $\sim$ 15 arcsec, or $\sim$120–130 kpc at $z \sim$ 2–3] beams of the single-dish submm telescopes used to detect SMGs cause the two merging discs to be blended into a single source for much of the pre-coalescence stage of the merger. Fig. 1 demonstrates the effects of blending by showing observed-frame 850-μm (assuming $z = 2$) continuum images of the merger from H11 near apocentre (first row), during the final approach (second row) and at the peak of the starburst (third row). The fourth row shows the isolated disc from H11. The first column shows the full-resolution simulated images. The others are meant to mimic the resolution attainable with various telescopes/interferometers. See the figure caption for full details.

Comparison of the columns shows the importance of blending. For instruments with resolution $\sim$15 arcsec (fourth column of Fig. 1), at all stages (even at apocentre) the two discs are blended into one submm source, and the morphologies of the galaxy-pair SMGs, the late-stage merger-induced starbursts and the isolated disc are similar. This resolution is typical of the surveys that have been done to date with Submillimetre Common-User Bolometer Array (SCUBA), AzTEC and *Herschel* SPIRE, so those surveys should contain many galaxy-pair SMGs consisting of widely separated merging discs blended into one submm source. With resolution $\sim$7 arcsec (third column of Fig. 1), the widely separated galaxy-pair SMGs (first row) can be resolved, but some mergers which are nearer coalescence but still predominantly forming stars quiescently will be blended into one source (second row). Thus wide-field surveys done with, for example, *Herschel* PACS or the Large Millimeter Telescope (LMT) can more effectively distinguish galaxy-pair SMGs from starburst SMGs and should have a higher fraction of close pairs in the far-IR (FIR)/(sub)mm than the SCUBA, AzTEC and *Herschel* SPIRE surveys. When the resolution is $\sim$0.5 arcsec (second column of Fig. 1; see Fig. 2 for zoom-ins of the boxed regions), typical of interferometers such as the SMA, the IRAM Plateau de Bure Interferometer and ALMA, all but the mergers nearest coalescence are resolved into two components, as this beam size corresponds to $\sim 3\,h^{-1}$ kpc at $z \sim$ 2–3. At such low separations, the mergers typically are undergoing strong starbursts, so the SMGs observed as single components with interferometers should be predominantly starbursts or isolated discs. Note that when widely separated merging discs are observed with some interferometers, one of the sources may be outside the field of view (FOV), so from the interferometry such SMGs would appear to be isolated discs, whereas they are actually galaxy-pair SMGs. Clearly it is difficult to disentangle the SMG subpopulations using the number of components observed or the morphology.

The blending of the discs during the early stages of a merger is effective at creating SMGs because, by treating the system as a single source, the integrated submm flux and SFR are both doubled; see sections 3.2 and 4.1 of H11, for further details. This blending, along with the stronger scaling between submm flux and SFR in the quiescent mode compared to the starburst mode, causes early-stage mergers to contribute significantly to the SMG population (details of the contribution will be presented in Hayward et al., in preparation). The brightest SMGs should still be merger-induced starbursts, but early-stage mergers must provide a significant contribution to the population.

There is already much evidence that some SMGs are early-stage mergers. For example, Engel et al. (2010) used submm interferometry to show that approximately half of their SMG sample are well-resolved binary systems. (See also Tacconi et al. 2006, 2008; Bothwell et al. 2010; Riechers et al. 2011a,b.) Two of the 12 SMGs in the Engel et al. sample consist of two well-separated, resolved components (projected separations $\gtrsim$20 kpc), and it is possible that they have missed one component of galaxy-pair SMGs with more widely separated components because of the limited FOV. In such widely separated systems the star formation induced by the tidal torques exerted by the discs upon one another is not sufficient to drive a strong starburst, so the discs would form stars at similar rates even if the companion were absent. Thus such systems should be considered physically analogous to normal disc galaxies rather than late-stage mergers because they are forming stars via the quiescent rather than starburst mode. In addition to the evidence from CO interferometry, support for the galaxy-pair contribution is provided by the frequency of multiple counterparts in the radio (e.g. Ivison et al. 2002, 2007; Chapman et al. 2005; Clements et al. 2008; Younger et al. 2009a; Yun et al. 2012) and 24-μm emission (e.g. Pope et al. 2006; Hwang et al. 2010; Yun et al. 2012) and SMGs with morphologies that appear more like discs than late-stage mergers (e.g. Bothwell et al. 2010; Carilli et al. 2010; Ricciardelli et al. 2010; Targett et al. 2011). (Note, however, that in gas-rich mergers, discs can rapidly re-form, potentially confusing interpretation of these results; Springel & Hernquist 2005; Narayanan et al. 2008b; Robertson & Bullock 2008; Hopkins et al. 2009.)

We have shown that physical arguments and observations suggest that the SMG population is a mix of quiescently star-forming galaxies and starbursts. Thus SMGs differ significantly from local ULIRGs, which are exclusively starburst- or AGN-dominated, so one should draw comparisons between the two populations with care. The heterogeneity complicates physical interpretation of the SMG population. For example, one should not apply a CO–H_2 conversion factor appropriate for starbursts to the quiescently star-forming subpopulation of SMGs. Furthermore, proper treatment of all subpopulations is key for reproducing the observed SMG number counts (Hayward et al. 2011b; Hayward et al., in preparation). However, the relative contributions of these various subpopulations is not yet observationally well determined, and predictions of the relative contributions depend sensitively on uncertain model details. We do not predict the relative contributions in the present work. Instead, we wish to determine how one can observationally distinguish between starburst-driven (late-stage merger) SMGs and those powered by quiescent star formation even when only integrated data are available. One can then use these diagnostics to observationally

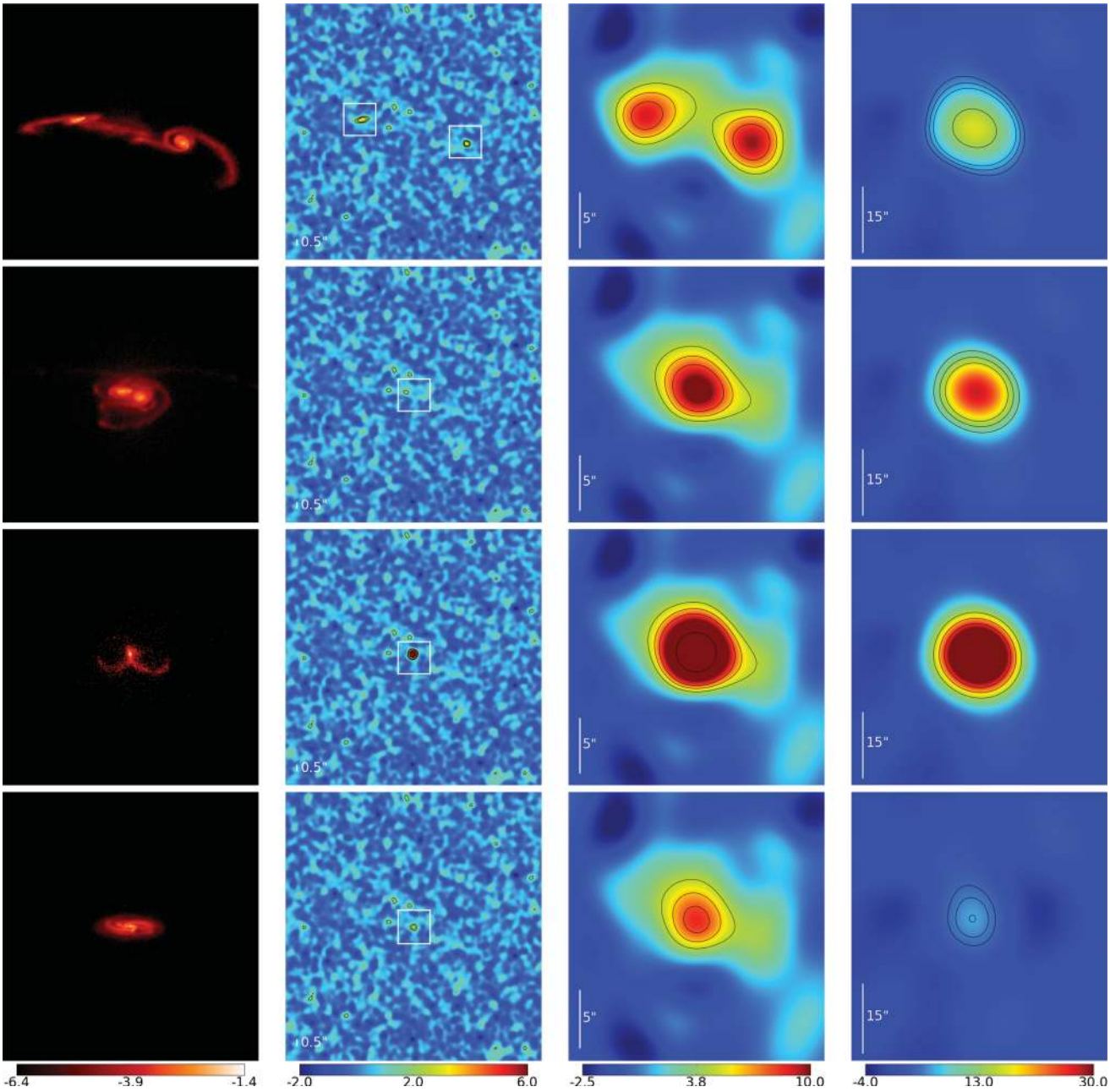

Figure 1. The first three rows show mock observed-frame 850-μm continuum images (assuming $z = 2$) of the merger from H11 near apocentre (first row; time $t = 0.28$ Gyr; nuclear separation $d_{\rm BH} = 71\,h^{-1}$ kpc; quiescently star forming), during the final approach (second row; $t = 0.63$ Gyr; $d_{\rm BH} = 13\,h^{-1}$ kpc; quiescently star forming) and at the peak of the starburst (third row; $t = 0.7$ Gyr; $d_{\rm BH} = 0.1\,h^{-1}$ kpc; starburst). The fourth row shows the isolated disc at $t = 0.28$ Gyr (quiescently star forming). The simulated images have been convolved with a Gaussian filter in order to mimic the intrinsic simulation resolution (first column; FWHM $= 0.2\,h^{-1}$ kpc) and the resolution achievable with various telescopes/interferometers. The second column has FWHM $= 0.5$ arcsec ($3\,h^{-1}$ kpc at $z = 2$), the resolution roughly characteristic of (sub)mm interferometers such as the SMA, the IRAM Plateau de Bure Interferometer and ALMA. (Note that we have not made any attempt to model the interferometric process.) See Fig. 2 for zoom-ins of the boxed regions. Third column: FWHM $= 7$ arcsec ($42\,h^{-1}$ kpc), representative of the resolution available with, for example, *Herschel* PACS and the current configuration of the LMT. Fourth column: FWHM $=$ 15 arcsec ($89\,h^{-1}$ kpc), representative of, for example, *Herschel* SPIRE and the JCMT. The fields of view are 23.5 arcsec ($140\,h^{-1}$ kpc) for the first three columns and 59 arcsec ($351\,h^{-1}$ kpc) for the last. For all but the first column, Gaussian noise with realistic amplitude has been added. The same noise map is used for all images in a given column. When present, scale bars show the beam FWHM, and contours correspond to 3σ, 4σ, 5σ, 10σ and 20σ. The first column has logarithmic scaling; all others are linear. Units are arbitrary.

constrain the relative contributions of starbursts and quiescently star-forming galaxies to the SMG population, thereby testing the bimodality we claim exists.

The rest of the paper is organized as follows. We describe our simulation methodology in Section 2. Section 3 presents multiple observational diagnostics that can be used to distinguish between quiescently star-forming galaxies and starbursts from integrated data alone, including the luminosity–effective dust temperature relation (Section 3.1), SFE (Section 3.2), IR excess (IRX; Section 3.3) and the SFR–$M_\star$ relation (Section 3.4). In Section 4 we discuss

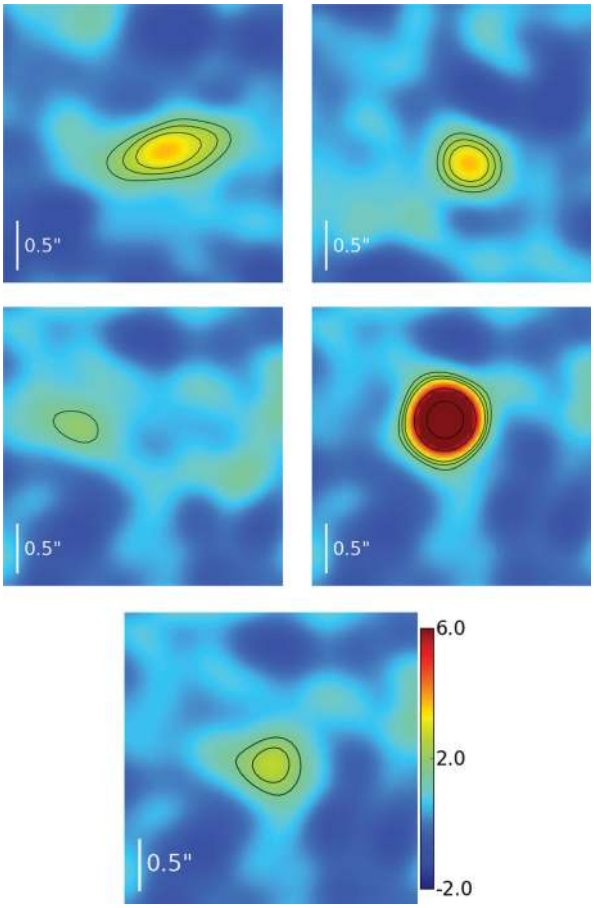

Figure 2. Zoom-ins of the regions marked with white boxes in the second column of Fig. 1. The upper-left (right) panel shows the left (right) region from the first-row image of Fig. 1. The middle left corresponds to the second row, the middle right to the third row and the bottom to the fourth row of Fig. 1. The FOV of each image is 3 arcsec ($18\,h^{-1}$ kpc). See the caption of Fig. 1 for further details.

some implications of our work, and in Section 5 we summarize and conclude.

2 SIMULATION METHODOLOGY

We analyse GADGET-2 (Springel, Yoshida & White 2001; Springel 2005) 3D *N*-body/smoothed particle hydrodynamics[2] (SPH) simulations of mergers of equal-mass disc galaxies with the SUNRISE (Jonsson 2006; Jonsson, Groves & Cox 2010) polychromatic Monte Carlo dust RT code to calculate synthetic SEDs of the simulated galaxies. The combination of GADGET-2 and SUNRISE has been used to successfully reproduce the SEDs/colours of a variety of galaxy populations, both low redshift and high redshift, including the following: local SINGS galaxies (Kennicutt et al. 2003; Dale et al. 2007; Jonsson et al. 2010); local ULIRGs (Younger et al. 2009b); extended ultraviolet (UV) discs (Bush et al. 2010); 24-μm-selected galaxies (Narayanan et al. 2010b); massive, quiescent, compact $z \sim 2$ galaxies (Wuyts et al. 2009, 2010); and K+A/post-starburst galaxies (Snyder et al. 2011), among other populations. These successes support our application of GADGET and SUNRISE to modelling high-redshift ULIRGs.

2.1 Hydrodynamic simulations

GADGET-2[3] (Springel et al. 2001; Springel 2005) is a TreeSPH (Hernquist & Katz 1989) code that computes gravitational interactions via a hierarchical tree method (Barnes & Hut 1986) and gas dynamics via SPH (Gingold & Monaghan 1977; Lucy 1977; Springel 2010a). The formulation of SPH used explicitly conserves both energy and entropy when appropriate (Springel & Hernquist 2002). Radiative heating and cooling is included in GADGET-2 following Katz, Weinberg & Hernquist (1996). Star formation is implemented using a volume-density-dependent KS law (Kennicutt 1998a), $\rho_{\rm SFR} \propto \rho_{\rm gas}^{1.5}$, with a minimum density threshold. The assumed KS index results in a surface-density-dependent KS law consistent with observations of $z \sim 2$ discs (Krumholz & Thompson 2007; Narayanan et al. 2008a, 2011a), suggesting that it is reasonable to use this prescription in our simulations of $z \sim 2$ mergers. The gas is enriched with metals assuming each particle behaves as a closed box, so those gas particles with higher SFRs are more rapidly metal enriched.

We use the subresolution two-phase ISM model of Springel & Hernquist (2003). In this model, cold, dense clouds are embedded in a diffuse, hot medium. Supernova feedback (Cox et al. 2006), radiative heating and cooling, and star formation control the exchange of energy and mass in the two phases. A simple model for black hole (BH) accretion and AGN feedback (Di Matteo, Springel & Hernquist 2005; Springel, Di Matteo & Hernquist 2005) is included. BH sink particles with initial mass $10^5\,{\rm M}_\odot$ are included in both initial disc galaxies. They accrete via Eddington-limited Bondi–Hoyle accretion (Hoyle & Lyttleton 1939; Bondi & Hoyle 1944). The luminosity of each BH is calculated from the accretion rate $\dot{M}_{\rm BH}$ assuming the radiative efficiency appropriate for a Shakura & Sunyaev (1973) thin disc, 10 per cent. Thus $L_{\rm bol} = 0.1\dot{M}c^2$. 5 per cent of the luminosity emitted by the BHs is deposited into the surrounding ISM.

The simulations are initialized in the following manner. Exponential discs with initial gas fraction $f_{\rm g} = M_{\rm gas}/(M_\star + M_{\rm gas}) = 0.8$[4] are embedded in dark matter haloes described by a Hernquist (1990) profile. The progenitor discs are scaled to $z \sim 3$ as described in Robertson et al. (2006) so that the mergers occur at $z \sim 2$. We have selected galaxy masses representative of the SMG population (e.g. Michałowski et al. 2012). The gravitational softening lengths are $200\,h^{-1}$ pc for the dark matter particles and $100\,h^{-1}$ pc for the stellar, gas and BH particles. Each disc galaxy is composed of 6×10^4 dark matter particles, 4×10^4 stellar particles, 4×10^4 gas

[2] Recent works (Agertz et al. 2007; Springel 2010b; Bauer & Springel 2011; Sijacki et al. 2011) suggest that the SPH method has several significant flaws, including artificially suppressing fluid instabilities, preventing efficient gas stripping of infalling structures and artificially damping turbulent eddies in the subsonic regime. Comparison of cosmological simulations run with GADGET-2 to otherwise identical simulations run with the more accurate moving-mesh code AREPO (Springel 2010b) shows that significant differences exist even though all physics incorporated in the simulations is identical (Kereš et al. 2011; Torrey et al. 2011; Vogelsberger et al. 2011). However, preliminary comparison of idealized merger simulations of the sort presented here run with GADGET-2 and AREPO suggests that the two methods yield similar results (Hayward et al., in preparation). Thus we expect our results to be robust in this regard.

[3] A public version of GADGET-2 is available at http://www.mpa-garching.mpg.de/ volker/gadget/index.html.

[4] Note, however, that we discard all snapshots with $f_{\rm g} < 0.4$ for reasons described below.

Table 1. Simulation parameters.

Name	M_{200} [a] ($\times 10^{12}\,h^{-1}\,\mathrm{M}_\odot$)	$M_{\star,\mathrm{init}}$ [b] ($\times 10^{10}\,h^{-1}\,\mathrm{M}_\odot$)	$M_{\mathrm{gas,init}}$ [c] ($\times 10^{10}\,h^{-1}\,\mathrm{M}_\odot$)	R_{init} [d] (h^{-1} kpc)	R_{peri} [e] (h^{-1} kpc)	θ_1 [f] (°)	ϕ_1^f (°)	θ_2^f (°)	ϕ_2^f (°)
b6i	6.2	5.3	22	70	6.7	0	0	71	30
b6j	6.2	5.3	22	70	6.7	−109	90	71	90
b6k	6.2	5.3	22	70	6.7	−109	−30	71	−30
b6l	6.2	5.3	22	70	6.7	−109	30	180	0
b6m	6.2	5.3	22	70	6.7	0	0	71	90
b6n	6.2	5.3	22	70	6.7	−109	−30	71	30
b6o	6.2	5.3	22	70	6.7	−109	30	71	−30
b6p	6.2	5.3	22	70	6.7	−109	90	180	0
b5.5i	3.2	2.7	11	57	5.3	0	0	71	30
b5.5j	3.2	2.7	11	57	5.3	−109	90	71	90
b5.5k	3.2	2.7	11	57	5.3	−109	−30	71	−30
b5.5l	3.2	2.7	11	57	5.3	−109	30	180	0
b5.5m	3.2	2.7	11	57	5.3	0	0	71	90
b5.5n	3.2	2.7	11	57	5.3	−109	−30	71	30
b5.5o	3.2	2.7	11	57	5.3	−109	30	71	−30
b5.5p	3.2	2.7	11	57	5.3	−109	90	180	0
b5i	1.6	1.4	5.6	44	4.0	0	0	71	30
b5j	1.6	1.4	5.6	44	4.0	−109	90	71	90
b5k	1.6	1.4	5.6	44	4.0	−109	−30	71	−30
b5l	1.6	1.4	5.6	44	4.0	−109	30	180	0
b5m	1.6	1.4	5.6	44	4.0	0	0	71	90
b5n	1.6	1.4	5.6	44	4.0	−109	−30	71	30
b5o	1.6	1.4	5.6	44	4.0	−109	30	71	−30
b5p	1.6	1.4	5.6	44	4.0	−109	90	180	0

[a]Virial mass of each progenitor; [b]initial stellar mass of each disc; [c]initial gas mass of each disc;
[d]initial separation of the discs; [e]pericentric passage distance; [f]orientation of each disc's spin axis in spherical coordinates.

particles and one BH particle. Two identical discs are initialized on parabolic orbits with initial separation $R_{\mathrm{init}} = 5R_{\mathrm{vir}}/8$ and pericentric distance twice the disc scale length (Robertson et al. 2006). We analyse only equal-mass mergers because these simulations provide sufficient examples of quiescently star-forming galaxies (during the early stages) and starbursts (near coalescence). The differences between quiescent and starburst modes are insensitive to orbit and merger mass ratio. (Not all mergers induce strong starbursts, but those are irrelevant for our present purposes.) The physical parameters of the simulated major mergers are summarized in Table 1.

2.2 Radiative transfer

We use the 3D Monte Carlo dust RT code SUNRISE[5] (Jonsson 2006; Jonsson et al. 2010) in post-processing to calculate the far-UV (FUV)–mm SEDs of each simulated merger at 10-Myr intervals. We briefly describe the SUNRISE calculation here, but the reader is encouraged to see Jonsson et al. (2010) for full details. SUNRISE uses the stellar and BH particles from the GADGET-2 simulations as radiation sources. Each stellar particle with age >10 Myr is treated as a single-age stellar population and assigned a STARBURST99 (Leitherer et al. 1999) SED template appropriate for its age and metallicity. Stellar particles with age <10 Myr are assigned a template from Groves et al. (2008), which includes emission from the H II regions surrounding the clusters. We do not include the photodissociation regions for the reasons discussed in detail in section 2.2.1 of H11. The stars in the initial discs are assigned ages by assuming that the population was formed at a constant rate equal to the SFR of the initial snapshot. The gas and stars present in the initial discs have metallicity $Z = 0.015$, which results in the galaxies being roughly on the $z \sim 2$ mass–metallicity relation during the starburst. The BH particles are assigned SEDs using the luminosity-dependent templates of Hopkins, Richards & Hernquist (2007). These templates are derived from observations of unreddened quasars, so they include the intrinsic power-law emission and reprocessed hot dust emission from the torus.

SUNRISE calculates the dust distribution by projecting the GADGET-2 gas-phase metal density on to a 3D adaptive mesh refinement grid using the SPH smoothing kernel and assuming a dust-to-metal density ratio of 0.4 (Dwek 1998; James et al. 2002). A minimum cell size of $55\,h^{-1}$ pc is used; this is sufficient to ensure the SEDs are converged to within $\lesssim$10 per cent at all wavelengths. Grain compositions, size distributions and optical properties are given by the Milky Way $R = 3.1$ dust model of Weingartner & Draine (2001) as updated by Draine & Li (2007). The opacity curve for this model has a power-law slope in the FIR $\beta \approx 2$.

To perform the RT we use 10^7 photon packets for each stage, $\sim$10 times the number of grid cells. This limits Monte Carlo noise to less than a few per cent. SUNRISE randomly emits the photon packets from the sources and randomly draws interaction optical depths using the appropriate probability distributions. At the interaction optical depth a fraction of the photon packet's intensity is absorbed; the remainder is scattered into a direction randomly drawn using the scattering phase function. This is repeated until the photon packet leaves the grid or its intensity drops below a minimum value.

The energy absorbed by the dust is reradiated in the IR. SUNRISE assumes all dust (except for half of the polycyclic aromatic hydrocarbons with grain size <100 Å; see Jonsson et al. 2010, for details) is in thermal equilibrium, so the dust temperature is calculated by setting the luminosity absorbed by each grain equal to the energy emitted by the grain. The equilibrium temperature of

[5] SUNRISE is publicly available at http://code.google.com/p/sunrise/.

a grain depends on the local radiation field heating the grain and its absorption cross-section, so there are in principle $n_{\rm cells}\, n_{\rm grain\,sizes}$ different dust temperatures in a given SUNRISE calculation. This is important to keep in mind when one considers fitting IR SEDs with modified blackbodies, as discussed below.

In high-density environments the ISM can be optically thick in the IR; this is common in the central regions of the late-stage mergers modelled here. Consequently, one must account for attenuation of the dust re-emission (a.k.a. dust self-absorption). Furthermore, since the IR emission absorbed heats the dust, one must iterate the dust temperature calculation and RT of the dust emission until the dust temperatures are converged. SUNRISE uses a reference field technique similar to that of Juvela (2005) to perform this iteration. We encourage the interested reader to see Jonsson et al. (2010) and Jonsson & Primack (2010) for details.

The SUNRISE calculation yields spatially resolved, multiwavelength (for these simulations there are 120 wavelengths sampling the UV–mm range) SEDs for each galaxy snapshot observed from seven different viewing angles distributed uniformly in solid angle. The data are analogous to data yielded by integrated field unit (IFU) spectrographs. For this work we spatially integrate to calculate integrated SEDs for the system. When calculating observed flux densities we assume the simulated galaxies are at redshift $z = 2$.

3 OBSERVATIONAL DIAGNOSTICS TO DISTINGUISH BETWEEN STAR FORMATION MODES

In this section we present multiple observational diagnostics that can distinguish between quiescently star-forming galaxies (for SMGs this includes galaxy pairs and isolated discs) and starbursts induced at the final coalescence of major mergers. We present diagnostics that rely only on integrated broad-band photometry and CO line intensities (to determine gas mass). We also assume that sufficiently accurate redshifts are known. Spatially resolved data, such as that provided by (sub)mm interferometers and near-IR (NIR) IFU spectrographs, can potentially provide more diagnostic power but come at a much greater observational cost. However, even with high-resolution NIR data it can be difficult to distinguish between disc galaxies and gas-rich mergers in which discs re-form shortly after the final coalescence (Robertson & Bullock 2008), and because of the high attenuation of SMGs NIR observations may not probe the central starburst regions. The FIR and (sub)mm are able to probe much deeper into the central regions, but the spatial resolution available at these wavelengths is much coarser unless an interferometer is used, so typically only integrated FIR and (sub)mm photometry are available for high-redshift galaxies. (However, ALMA will soon change this situation drastically.) Finally, mergers and starbursts are not equivalent (see above), so identifying an object as a merger based on morphology does not ensure it is also a starburst. Thus diagnostics that make use of only integrated data will continue to be crucial for distinguishing between star formation modes and understanding the properties of high-redshift galaxies.

We have identified the starburst phase by defining the baseline SFR as the minimum SFR that occurs between the first passage and coalescence and selecting all snapshots where the instantaneous SFR is >3 times that baseline SFR. This factor is chosen so that the star formation induced by the merger dominates that which would occur in the discs even if they were not merging. Increasing (decreasing) the threshold would result in less (more) sources identified as bursts and amplify (diminish) the differences between modes that we describe below. The snapshots that meet the starburst criterion are labelled 'starburst' and plotted as blue squares. Since the mutual gravitational torques are subdominant at the first passage relative to internal instabilities, the galaxies are primarily quiescently star forming prior to the starburst induced at coalescence. We thus label all snapshots before the starburst phase 'quiescent' and plot them as black circles. All snapshots after the starburst phase are neglected, as these are typically AGN or spheroids with relatively little ongoing star formation. We will investigate the FIR properties of obscured AGN in detail in future work.

Our focus is the bright SMG population, defined by $S_{850} > 5$ mJy, so throughout this work simulated SMG data points are plotted with larger symbols (black circles for quiescently star-forming and blue squares for starburst SMGs). However, in order to comment on the applicability of the diagnostics to the high-redshift ULIRG population in general, data for simulated galaxies that would not be selected as SMGs are also plotted. Small black circles (blue squares) correspond to quiescently star-forming (starburst) ULIRGs with $S_{850} < 5$ mJy.

We have neglected all snapshots with greater than 40 per cent gas fraction in order to eliminate the early parts of the simulations when the discs are still stabilizing and to remain consistent with (albeit uncertain; e.g. Bothwell et al. 2010) observational constraints (Tacconi et al. 2006, 2008). Such high initial gas fractions are required to maintain sufficient gas until the time of coalescence because our simulations do not include any additional gas supply beyond what the galaxies start with. We have checked that the results are qualitatively the same when we use an initial gas fraction of 60 per cent and include all snapshots. All quantities plotted are totals for the entire system because we wish to present observational diagnostics based on integrated data alone. In all figures, if a simulated galaxy subpopulation has sample size $N > 100$, only a randomly selected subsample of 100 objects is plotted; this is done to improve the readability of the plots, as a subsample of 100 is sufficient to show the distribution of the simulated data. Finally, note that we plot data from idealized simulations without applying any weighting to account for cosmological abundances. Thus the exact distribution of data in the various diagnostic plots is not necessarily representative of the SMG population. What is meaningful, however, is when subpopulations occupy distinct regimes in a diagnostic plot; this is a clear prediction for how the star formation modes should differ and how one can observationally disentangle the classes in order to determine their relative contributions to a given galaxy population.

3.1 Luminosity–effective dust temperature relation

FIR galaxy SEDs are often described in terms of a 'dust temperature' obtained via fitting a single-temperature (single-T) modified blackbody to the FIR SED (Hildebrand 1983). The single-T modified blackbody form is

$$S_\nu = S_0 \left(1 - {\rm e}^{-(\lambda/\lambda_1)^{-\beta}}\right) B_\nu(T_{\rm d}), \quad (1)$$

where S_ν is the flux density at rest-frame frequency ν, S_0 is the normalization, λ_1 is the wavelength at which the effective optical depth $\tau_\lambda = 1$, β is the effective slope of the emissivity in the FIR, $T_{\rm d}$ is the effective dust temperature[6] and $B_\nu(T_{\rm d})$ is the Planck function.

[6] Note, however, that in both the simulations and reality, there is *always* a distribution of dust temperatures, as the temperature of a grain depends on the local radiation field and the size of the grain. The effective dust temperature fitted in this manner may provide a good approximation to the

The form is motivated by a simple model: if it is assumed that the entire mass of dust has *physical* temperature equal to $T_{\rm d}$, the source has constant source function $B_\nu(T_{\rm d})$. Assuming also that the source is uniform, with projected area A and angular size $\mathrm{d}\Omega$, and that the opacity curve is a power law in the FIR,

$$\kappa_\lambda = \kappa_0(\lambda/\lambda_0)^{-\beta}, \tag{2}$$

the optical depth is

$$\tau_\lambda = \frac{\kappa_\lambda M_{\rm d}}{A} = \frac{\kappa_0 M_{\rm d}}{\mathrm{d}\Omega D^2}\left(\frac{\lambda}{\lambda_0}\right)^{-\beta}, \tag{3}$$

where D is the distance to the source, $M_{\rm d}$ is the dust mass and κ_0 is the opacity at wavelength λ_0.[7] (For simplicity we have assumed $z \ll 1$ here. In practice, we transform our simulation photometry to the rest frame before fitting.) Under these assumptions, the solution of the RT equation is

$$S_\nu = \mathrm{d}\Omega\left(1 - \mathrm{e}^{-\tau_\lambda}\right) B_\nu(T_{\rm d}) \tag{4}$$

$$= \mathrm{d}\Omega\left[1 - \exp\left(\frac{-\kappa_0 M_{\rm d}}{\mathrm{d}\Omega D^2}\left(\frac{\lambda}{\lambda_0}\right)^{-\beta}\right)\right] B_\nu(T_{\rm d}). \tag{5}$$

Note that the validity of equation (5) is limited to the IR wavelengths where the opacity curve has the assumed power-law form. Typically it is assumed that optical depths in the FIR are small, so $[1 - \exp(-(\lambda/\lambda_1)^{-\beta})] \approx (\lambda/\lambda_1)^{-\beta}$. Thus,

$$S_\nu = S_0'\lambda^{-\beta} B_\nu(T_{\rm d}), \tag{6}$$

where $S_0' = S_0\lambda_1^{\beta}$. We shall refer to the form given in equation (6) as the single-T optically thin (OT) modified blackbody. We refer the reader to the appendix of H11 for more details about the modified blackbody forms.

The simple model described above provides motivation for the fitting forms given in equations (1) and (6). However, because the simple assumptions inherent in the model (specifically the single dust temperature and assumed geometry) are certainly not true for real galaxies, one should not assume a priori that the parameters derived via fitting FIR SEDs with equation (1) or (6) can be used to infer the physical quantities present in equation (5). Indeed, the parameters determined using this fitting method should not be interpreted physically for multiple reasons. For example, β and $T_{\rm d}$ are degenerate (e.g. Sajina et al. 2006), and the $T_{\rm d}$ and β one derives from the fitting depend strongly on observational noise (Shetty et al. 2009a), dust temperature variations along the line of sight (Shetty et al. 2009b) and the wavelength range spanned by the photometry included in the fit (Magnelli et al. 2012, hereafter M12). Even in the case of a uniform radiation field grains of different sizes will have different temperatures, so there is no reason to expect β derived from the single-T modified blackbody to be equal to the intrinsic power-law index of the dust emissivity in the FIR. (However, if a distribution of dust temperatures is permitted, it is perhaps possible to recover the true β from fitting the SED.) Furthermore, adding a significant component of very cold dust can mimic the effect of high effective optical depth, so τ and the temperature distribution are also degenerate (Papadopoulos et al. 2010b). Finally, even the fitting method can cause incorrect conclusions: when traditional non-hierarchical χ^2 minimization is used, observational noise can introduce a spurious anticorrelation between $T_{\rm d}$ and β even when the true values of $T_{\rm d}$ and β are positively correlated; it is possible to avoid this pitfall by utilizing a hierarchical Bayesian method (Kelly et al. 2012). Thus, at the least, one should use caution when attempting to infer physical conditions from the parameters derived by modified blackbody fitting, as we demonstrate below.

However, even if the model parameters cannot be interpreted physically, the modified blackbody forms often provide acceptable descriptions of the FIR SEDs of galaxies. Thus in order to compare to observations in a meaningful way and provide testable predictions from our models, we have fit the FIR SEDs of our simulated galaxies with both the OT and full forms of the modified blackbody. We have also used a more sophisticated model which assumes a power-law distribution of temperatures. The results for each fitting form are discussed in turn below.

3.1.1 Single-T, OT modified blackbody

Fig. 3 shows the distribution of our simulations on the $T_{\rm d}$–$L_{\rm IR}$ plot when the single-T OT modified blackbody form (equation 6) is used. The top panel shows $T_{\rm d}$ derived from fitting the simulated *Herschel* SPIRE (Griffin et al. 2010) 250-, 350- and 500-μm photometry, SCUBA (Holland et al. 1999) 850-μm photometry and AzTEC 1.1-mm photometry (Wilson et al. 2008); in the bottom panel, the simulated *Herschel* PACS (Poglitsch et al. 2010) 160-μm photometry is also included in the fit. We have excluded the PACS 70- and 100-μm points because often an acceptable fit cannot be obtained if those points are used. Possible reasons for this are that, at those wavelengths, stochastically heated grains contribute significantly to the SED and the assumption of optical thinness is violated most. When performing the Levenberg–Marquardt least-squares fit, we have assumed $10 \leq T_{\rm d} \leq 100$ K, $1.0 \leq \beta \leq 2.5$ and constant fractional flux error of 10 per cent. The median effective dust temperatures for the each galaxy type are given in Table 2, where the galaxies have been divided into two bins, $L_{\rm IR} < 10^{13}\,{\rm L}_\odot$ and $L_{\rm IR} \geq 10^{13}\,{\rm L}_\odot$.

The key trends to take away from Fig. 3 are that effective dust temperature correlates with luminosity, and the most luminous, hottest SMGs are almost exclusively starbursts. When the PACS photometry is (not) used, almost all SMGs with $T_{\rm d} \gtrsim 40$ (35) K are starbursts. Thus one can use a cut in $T_{\rm d}$ to cleanly select starbursts from the overall SMG population. Furthermore, for the SMG population the cut $L_{\rm IR} > 10^{13}\,{\rm L}_\odot$ can also be used to cleanly select starburst SMGs.

When the non-SMG population is also considered, the scatter in the $T_{\rm d}$–$L_{\rm IR}$ relation is increased, as there is a significant population of galaxies at the lower $L_{\rm IR}$ end with relatively high $T_{\rm d}$ that are missed by the SMG selection (see also fig. 1 of M12). These simulated galaxies correspond to the observed 'hot-dust ULIRGs' (Chapman et al. 2004, 2008; Casey et al. 2009, 2011; Magnelli et al. 2010; M12); we discuss the differences between the SMG and non-SMG subpopulations in detail below (Section 4.3). At a given $L_{\rm IR}$ the non-SMGs have higher effective dust temperatures than the SMGs because of the bias inherent in the SMG selection.[8] As for

luminosity-weighted average dust temperature, but to our knowledge this correspondence has not been demonstrated.

[7] It is important to distinguish between λ_0 and λ_1. The former is just part of the parametrization of the opacity curve of the dust (see equation 2). The latter depends on the dust mass and effective area of the source in addition to the dust opacity. λ_1 can be obtained from equation (3) by setting $\tau_\lambda = 1$ and solving for λ.

[8] Note that this is not necessarily reflected in the tables of median temperatures because the non-SMGs in a given bin tend to have lower $L_{\rm IR}$ than the SMGs; thus the comparison of the median $T_{\rm d}$ values for the bins is not the same as comparing the values for populations at fixed $L_{\rm IR}$.

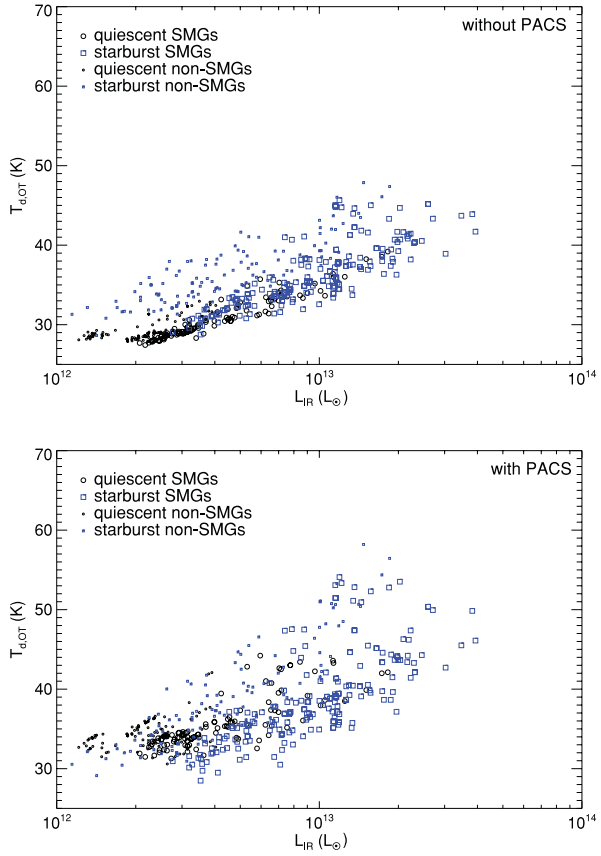

Figure 3. Effective dust temperature $T_{\rm d}$ (K) derived from fitting the single-T OT modified blackbody (equation 6) to the simulated SPIRE, SCUBA and AzTEC photometry (top) and PACS 160-μm, SPIRE, SCUBA and AzTEC photometry (bottom) versus total IR luminosity $L_{\rm IR}$ ($\rm L_\odot$). For the SMG population, effective dust temperature correlates strongly with luminosity, and the sources in the high-$T_{\rm d}$, high-$L_{\rm IR}$ region of the plot are almost exclusively merger-induced starbursts. The SMG selection is biased against galaxies with hotter SEDs, so, for a given luminosity, the non-SMGs have higher $T_{\rm d}$ than the SMGs. At the lower luminosities there is a significant population of galaxies missed by the SMG selection; these cause the $T_{\rm d}$–$L_{\rm IR}$ relation for the full population to have larger scatter than that for the SMG population. Note also that adding the PACS 160-μm data results in higher $T_{\rm d}$ and increased scatter.

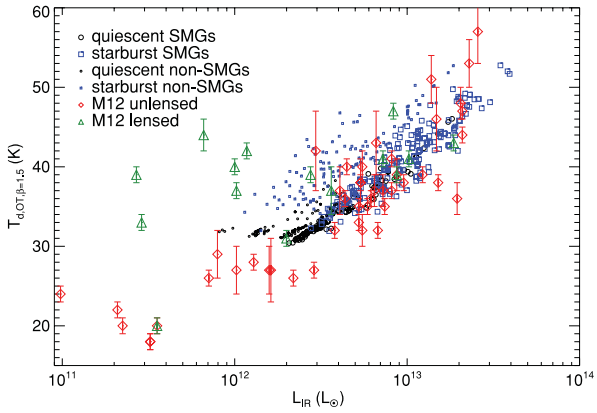

Figure 4. Effective dust temperature (K) yielded by fitting the single-T OT form of the modified blackbody with $\beta = 1.5$ fixed versus $L_{\rm IR}$ ($\rm L_\odot$). In order to compare directly to M12, only the PACS 160-μm, SPIRE, SCUBA and AzTEC photometry have been used in the fit. The large black circles (blue squares) are the simulated quiescently star-forming (starburst) SMGs, and the small black circles (blue squares) are the simulated quiescently star-forming (starburst) non-SMGs. The red diamonds (green triangles) with error bars are the unlensed (lensed) SMGs ($S_{850} > 3$ mJy) from M12. The agreement between the simulated and observed unlensed SMGs is very good for the entire luminosity range sampled by our simulations, giving confidence that the FIR SEDs of the simulated galaxies are reasonable.

the SMG population, almost all non-SMGs with $T_{\rm d} \gtrsim 40$ (35) K when the PACS 160-μm point is (not) used in the fit are starbursts. Thus the $T_{\rm d}$ cut is an effective means for selecting starbursts from the overall galaxy population even at luminosities where there are a mix of starbursts and quiescently star-forming galaxies. This is consistent with the observations of Magnelli et al. (2010) and M12.

Note also that inclusion of the PACS photometry results in both increased dust temperature and larger scatter. This effect occurs because we allow β to vary; if β is fixed, then $T_{\rm d}$ is relatively similar whether or not the PACS photometry is included (Magdis et al. 2010; M12). The fact that the results of IR SED fitting can depend strongly on the regions of the SED sampled is another reason one must use caution when interpreting the derived parameters. This is especially significant when comparing low- and high-redshift samples, as in this case a given observed data point can correspond to significantly different rest-frame wavelengths for different redshifts.

In Fig. 4 we compare our simulations to the observations of M12. The effective temperature plotted here has been determined by fitting the single-T OT form with fixed $\beta = 1.5$, which is what M12 have done. Only the PACS 160-μm, SPIRE, SCUBA and AzTEC photometry have been used for the fits. For the range in $L_{\rm IR}$ covered by the simulations, the agreement between the simulated and observed SMGs is very good. This agreement gives confidence that the FIR SEDs of the simulated galaxies are reasonable. However, there are a few galaxies that lie below the region spanned by the simulations. This is consistent with the conclusion of Jonsson et al. (2010) that real galaxies form a more diverse family of SEDs than are generated by the simulations.

3.1.2 Full form of the single-T modified blackbody

There is increasing evidence that the simple single-T OT modified blackbody form given in equation (6) provides a poor fit to the FIR SEDs of simulated (H11) and observed (Lupu et al. 2010; Conley et al. 2011; M12; Sajina et al., submitted) high-redshift ULIRGs

Table 2. Median $T_{\rm d}$ (K) for the OT single-T modified blackbody.

Galaxy type	Without PACS 160 μm		With PACS 160 μm		Sample size	
	$L_{\rm IR} < 10^{13}\,\rm L_\odot$	$L_{\rm IR} \geq 10^{13}\,\rm L_\odot$	$L_{\rm IR} < 10^{13}\,\rm L_\odot$	$L_{\rm IR} \geq 10^{13}\,\rm L_\odot$	$L_{\rm IR} < 10^{13}\,\rm L_\odot$	$L_{\rm IR} \geq 10^{13}\,\rm L_\odot$
Quiescent SMGs	30.1	36.4	34.2	42.0	954	8
Starburst SMGs	33.4	38.1	35.1	41.0	409	292
Quiescent non-SMGs	28.9	–	34.1	–	648	–
Starburst non-SMGs	35.2	42.0	36.4	50.2	741	21

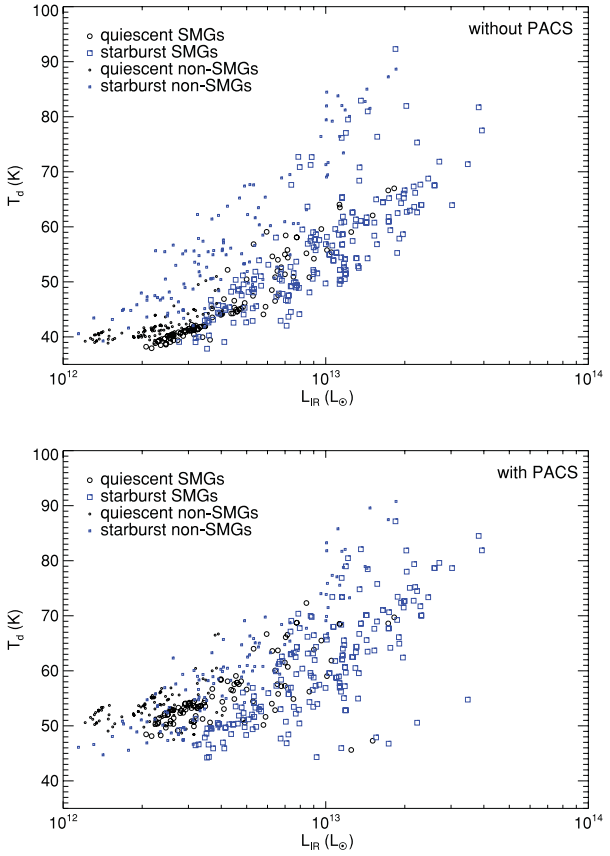

Figure 5. Same as Fig. 3, but the effective dust temperature has been derived by fitting the full form of the modified blackbody (equation 1). The $T_{\rm d}$ values derived using this form are systematically higher than when optical thinness is assumed (see text for details), but all trends – except for the increased scatter when PACS 100- and 160-μm data are included – seen in Fig. 3 still hold.

on the Wien side of the SED. Instead, more sophisticated forms, such as equation (1) or multiple-temperature models (e.g. Dale & Helou 2002; Clements, Dunne & Eales 2010; Kovács et al. 2010), must be used. Equation (1) allows non-negligible optical depths in the FIR, and allowing β to vary mimics the effect of a temperature distribution (e.g. Shetty et al. 2009b; Clements et al. 2010). The multiple-temperature models can account for non-negligible optical depths in the IR and multiple temperatures of dust, both of which are physically more valid assumptions than those implicit in the single-T OT blackbody model, but require at least one parameter more than equation (1).

Fig. 5 shows the $T_{\rm d}$–$L_{\rm IR}$ plot when we derive the effective dust temperature $T_{\rm d}$ by fitting the full form of the modified blackbody, equation (1), to the simulated SPIRE+SCUBA+AzTEC (top) and PACS+SPIRE+SCUBA+AzTEC (bottom) photometry. The PACS 70-μm data have not been used because at $z = 2$ observed-frame 70 μm corresponds to rest-frame 23 μm, a region of the SED dominated by emission from stochastically heated grains. If the 70-μm data are included, an acceptable fit is often not possible. We have assumed $10 \leq T_{\rm d} \leq 100$ K, $1.0 \leq \beta \leq 2.5$ and 10 per cent flux uncertainty. λ_1 is allowed to vary freely, but in practice it is always greater than ∼375 μm. The median effective dust temperatures are given in Table 3.

Almost all trends demonstrated by Fig. 3 hold here: effective dust temperature correlates with luminosity, and the most luminous, hottest sources [$T_{\rm d} \gtrsim 70$ (60) K when the PACS data are (not) used] are almost exclusively starbursts. Almost all galaxies with $L_{\rm IR} > 10^{13}\,{\rm L}_{\odot}$ are starbursts, but, again, a $T_{\rm d}$ cut is significantly better than an $L_{\rm IR}$ cut at selecting starbursts from the non-SMG population. There is a significant population of galaxies at the lower luminosity end that are missed by the SMG selection because of their relatively high effective dust temperatures. Finally, inclusion of the PACS photometry again results in higher effective dust temperature.

Comparison of Figs 3 and 5 and the median values given above shows that assuming optical thinness results in systematically lower $T_{\rm d}$ than when equation (1) is used. This occurs because $[1 - \exp(-(\lambda/\lambda_1)^{-\beta})] < (\lambda/\lambda_1)^{-\beta}$ for all $\lambda > 0$. Thus, for fixed $T_{\rm d}$ and β, the assumption of optical thinness will systematically overpredict the flux at frequencies for which $\tau \gtrsim 1$. As a result, for a given SED and fixed β, $T_{\rm d}$ derived from equation (6) will be lower than that derived from equation (1). This effect has been demonstrated when fitting SEDs of high-redshift ULIRGs (Lupu et al. 2010; Conley et al. 2011; Sajina et al., submitted), and it shows that one should use caution when attempting to interpret $T_{\rm d}$ physically.

3.1.3 *Power-law T-distribution model*

We have also fitted a subset of the simulated galaxies' SEDs assuming a power-law temperature distribution with a low-temperature cut-off following Kovács et al. (2010). The fitting method is summarized here, but the reader is referred to Kovács et al. (2010) and M12 for full details of the model. In this model, the dust has a distribution of physical temperatures given by $\mathrm{d}M_{\rm d}/\mathrm{d}T \propto T^{-\gamma}$ for $T > T_{\rm c}$ and $M_{\rm d}(T) = 0$ otherwise. The effective optical depth τ_λ is parametrized as

$$\tau_\lambda = \kappa_0 \left(\frac{\lambda}{\lambda_0}\right)^{-\beta} \frac{M_{\rm d}}{\pi R_{\rm eff}^2}, \qquad (7)$$

where $R_{\rm eff}$ can be thought of as an effective radius of the source. Because of the added parameter γ, we have always used the PACS 100- and 160-μm, SPIRE, SCUBA and AzTEC data. Following M12, we have assumed that single values of β, γ and $R_{\rm eff}$ can be used for all sources, and we have used a subset of 20 simulated galaxies to fix those parameters in the following manner. We first gridded the (β, γ, $R_{\rm eff}$) parameter space. For each point in the grid we fit all 20 sources allowing $T_{\rm d}$ and $M_{\rm d}$ to vary. We summed the χ^2 values of the individual fits for each parameter combination and chose the parameter combination with the lowest total χ^2 value. As above, the fractional flux error assumed is 10 per cent. The parameters we determined in this manner are (β, $R_{\rm eff}$, γ) = (1.6, 2 kpc, 8.7). Kovács et al. (2010) found that the parameters (β, $R_{\rm eff}$, γ) = (1.5, 1 kpc, 6.7) provided the best description of their sample of $z \sim 2$ starbursts. For the SMG sample of M12, the best-fitting parameters are (β, $R_{\rm eff}$, γ) = (2.0 ± 0.2, 3 ± 1 kpc, 7.3 ± 0.3). The values of β and $R_{\rm eff}$ derived from our simulations lie between those from the two observational studies. However, our temperature distribution is steeper than those found by both Kovács et al. and M12, perhaps because our simulations do not yet include stochastically heated very small grains and thus may underestimate the amount of dust at high temperatures. We defer a detailed comparison of simulated and observed ULIRG SEDs and the derived dust parameters to future work.

The resulting $T_{\rm c}$–$L_{\rm IR}$ plot is shown in Fig. 6. Note that the temperature plotted here is the low-temperature cut-off, which is also the temperature of most of the dust because of the steepness of the power-law distribution. The median $T_{\rm c}$ values are given in Table 4.

Table 3. Median T_d (K) for the full form of the single-T modified blackbody.

Galaxy type	Without PACS 100 and 160 μm		With PACS 100 and 160 μm		Sample size	
	$L_{IR} < 10^{13}\,L_\odot$	$L_{IR} \geq 10^{13}\,L_\odot$	$L_{IR} < 10^{13}\,L_\odot$	$L_{IR} \geq 10^{13}\,L_\odot$	$L_{IR} < 10^{13}\,L_\odot$	$L_{IR} \geq 10^{13}\,L_\odot$
Quiescent SMGs	42.7	63.5	60.4	69.6	954	8
Starburst SMGs	48.6	61.3	57.2	70.4	409	292
Quiescent non-SMGs	40.9	–	60.1	–	648	–
Starburst non-SMGs	53.3	80.1	60.0	77.8	741	21

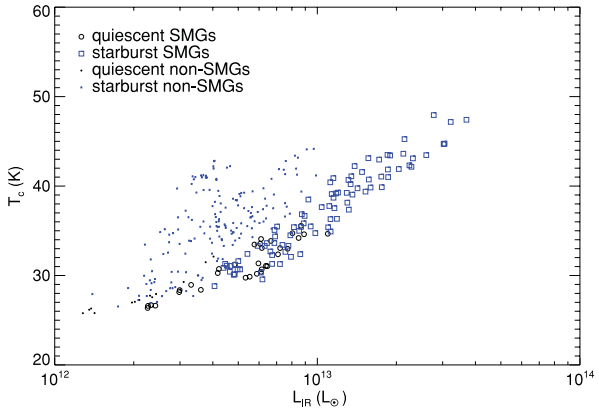

Figure 6. Same as Fig. 3, but the effective dust temperature has been derived by assuming a power-law distribution of dust temperatures following Kovács et al. (2010; the temperature plotted is the low-temperature cut-off of the power-law distribution – see text for details). Again, the dust temperatures differ, but the trends are insensitive to the manner in which the effective dust temperature is derived. All sources with $T_c > 36$ K are starbursts.

As seen for both single-T fitting forms, for SMGs there is a clear correlation between effective dust temperature and luminosity, and the starbursts are the most luminous and have the highest values of T_c. An effective dust temperature cut is very effective here: all galaxies with $T_c > 36$ K are starbursts. Furthermore, at the lower luminosity end, there is again a population of starbursts missed by the SMG selection because of their relatively high effective dust temperatures. Consequently, a T_c cut can select a larger subset of the starbursts than can an L_{IR} cut.

M12 argue that the cut $T_c \gtrsim 25$ K can very effectively separate starbursts (identified observationally by their offset from the SFR–$M_\star$ relation) from quiescently star-forming SMGs, as we also see in our simulations. However, the value of T_c that separates simulated starburst SMGs from quiescently star-forming SMGs differs significantly from that derived by M12. The primary reason for this discrepancy is that the (β, R_{eff}, γ) values derived for our simulations differ from those for the M12 sample. If the simulated SEDs are fitted with the (β, R_{eff}, γ) values from M12, the resulting T_c values are significantly lower, the SED fits are still acceptable and the agreement between the observations and simulations is much improved. See section 6 of M12 for further details.

3.1.4 Summary of results that are independent of the fitting form

Figs 3, 5 and 6 all show that the T_d–L_{IR} plot is an excellent way to select starburst SMGs from the general population. In all three figures there is a clear correlation between T_d and L_{IR}, which agrees with observations of both local and high-redshift ULIRGs (e.g. Kovács et al. 2006; Amblard et al. 2010; Chapman et al. 2010; Hwang et al. 2010; Magdis et al. 2010; Magnelli et al. 2010; M12). Though there is some overlap between the quiescently star-forming and starburst subpopulations, the most luminous, hottest SMGs are almost exclusively starbursts. Note that both the T_d–L_{IR} correlation and the separation between the populations are independent of the fitting method used, though the specific temperature values above which there are no quiescently star-forming galaxies differ (as expected because of the systematic difference in temperatures yielded by the two methods). Thus our simulations make the clear, robust prediction that the most luminous galaxies will have the hottest SEDs and will typically be late-stage merger-induced starbursts.

In all cases there is a subpopulation of hot-dust ULIRGs which have relatively high T_d for a given IR luminosity. Such galaxies are not present in the SMG subpopulation because of the bias of the SMG selection. The SMG selection bias results in an apparent relatively tight correlation between T_d and L_{IR}, so an L_{IR} cut is roughly as effective as a T_d cut at selecting a subset of starburst SMGs from the general SMG population. When the hot-dust ULIRGs are included, the scatter in the T_d–L_{IR} relation for $L_{IR} \lesssim 10^{13}\,L_\odot$ is increased significantly. For $L_{IR} \lesssim 10^{13}\,L_\odot$, a T_d cut can select starbursts, whereas an L_{IR} cut cannot. Thus if one wishes to select starbursts from a given galaxy population, a T_d cut is preferred because it can isolate a larger subsample of starbursts with a wider range in L_{IR}.

Fig. 7 shows T_d derived from fitting the single-T OT modified blackbody form (top), full form of the single-T modified blackbody (middle) and power-law T-distribution model (bottom) to the PACS, SPIRE, SCUBA and AzTEC photometry (with some PACS

Table 4. Median low-temperature cut-off T_c (K) for the power-law dust temperature distribution model.

Galaxy type	With PACS 100 and 160 μm		Sample size[a]	
	$L_{IR} < 10^{13}\,L_\odot$	$L_{IR} \geq 10^{13}\,L_\odot$	$L_{IR} < 10^{13}\,L_\odot$	$L_{IR} \geq 10^{13}\,L_\odot$
Quiescent SMGs	31.0	34.7	31	1
Starburst SMGs	32.6	40.9	41	46
Quiescent non-SMGs	27.6	–	15	–
Starburst non-SMGs	36.0	–	164	–

[a]The sample sizes here differ from those above because only a subset of the simulated galaxy SEDs have been fitted with the power-law T-distribution form.

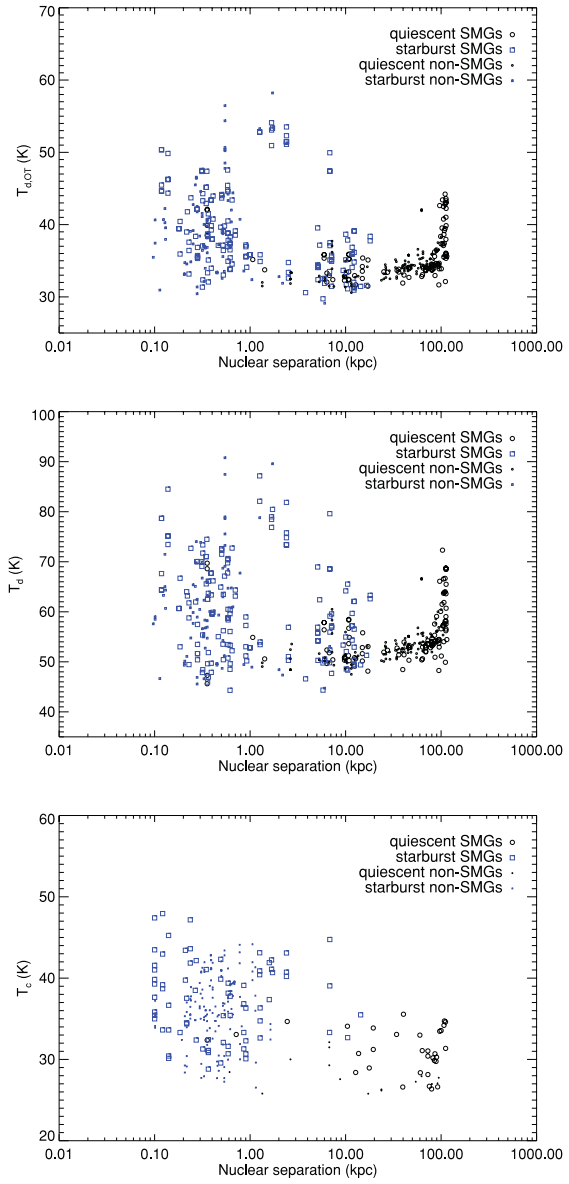

Figure 7. Effective dust temperature derived from fitting the single-T OT modified blackbody (equation 6; top), full modified blackbody (equation 1; middle) and power-law T-distribution model (bottom) to the PACS+SPIRE+SCUBA+AzTEC integrated photometry versus nuclear separation for snapshots where the BHs have not coalesced. T_d is anti-correlated with nuclear separation, but there is a large scatter in effective temperature at a given nuclear separation. The objects with the highest effective dust temperatures ($T_{d,OT} \gtrsim 35$ K, $T_d \gtrsim 55$ K and $T_c \gtrsim 36$ K) are almost exclusively late-stage merger-induced starbursts with $d_{BH} \lesssim 10$ kpc. The simulations with high T_d at $d_{BH} \sim 100$ kpc are those where L_{IR} – and thus T_d – is still high (but instantaneous SFR is not) because of the burst of star formation that can occur at first passage.

data excluded for each form, as described above) versus separation of the central BHs (also known as nuclear separation; d_{BH}). The nuclear separation serves as a proxy for the merger stage, but it is important to keep in mind that it does not decrease monotonically as the merger progresses. The starburst galaxies have systematically lower d_{BH} than the quiescently star-forming galaxies because the tidal torques which drive the starburst are strongest at the final coalescence of the two discs. The typical T_d values increase with decreasing d_{BH}, though there is large scatter at a given d_{BH}, especially for $d_{BH} \lesssim 10$ kpc. This occurs because, for a given simulation, L_{IR} is anticorrelated with d_{BH} (because the starburst is strongest at low nuclear separations), and, as seen above, the most luminous galaxies are also the hottest. Interestingly, for a given nuclear separation the SMGs and non-SMGs span a similar range in T_d. The simulated mergers with $d_{BH} \sim 100$ kpc and relatively high T_d are those observed shortly after the first passage. At this time L_{IR} is high, and the starburst mode is important even though the galaxies do not meet the definition of starburst because the instantaneous SFR has dropped.

3.1.5 *Comparison of fitting forms*

In the top panel of Fig. 8 we show the rest-frame SED of one of the simulated starbursts viewed from a single viewing angle. The overplotted data points are the simulated photometry as described in the caption. The four lines are fits to the photometry using the models discussed above. The bottom panel shows the ratio of the model SED derived from fitting the photometry to the actual SED for each of the fitting methods. The values of T_d and β for the best-fitting models are given in Table 5.

All forms except the single-T OT modified blackbody with fixed $\beta = 1.5$ are able to recover the simulated photometry to within ∼10 per cent, which is the level of uncertainty assumed when performing the χ^2 minimization. The single-T OT form with fixed β can only recover the photometry to within ∼30 per cent. Though the more complicated forms are successful at recovering the photometry used for the fitting, they have varying levels of success describing the SED beyond the wavelengths spanned by the photometry. As explained above, we expect the OT modified blackbody to underpredict the SED on the Wien side of the SED. Indeed, this model underpredicts the SED shortwards of the 100-μm point, and the underprediction is more severe when β is fixed. The full form of the modified blackbody fares better, but it also underpredicts the SED for rest-frame wavelength $\lambda_{rest} \lesssim 20$ μm. The power-law temperature distribution model fares best at the shortest wavelengths, but it overpredicts the SED at $\lambda_{rest} \sim 15$–25 μm. At the longest wavelengths, the single-T OT and power-law T-distribution models are most accurate because they have relatively low β values and thus less steeply declining SEDs.

Table 5 shows that the derived parameters T_d and β, which are often interpreted as a physical dust temperature and the intrinsic power-law index of the emissivity of the dust grains in the FIR, vary for the different fitting forms. Since observed galaxies typically have $20 \lesssim T_d \lesssim 80$ K and $1 \lesssim \beta \lesssim 2$, the variation among the best-fitting values is very significant. Consequently, it is difficult to interpret the fitted parameters physically, as clearly the intrinsic properties of the dust do not vary with the method used to fit the SED.[9] If the fitted effective dust temperatures have a physical meaning, they may correspond to different physical temperatures (e.g. the single-T modified blackbody may recover the luminosity-weighted dust temperature, whereas the power-law model may better recover the mass-weighted temperature). None of the fitted β values recovers the intrinsic β of the dust, which is ∼2 for the dust model we use. This is not unexpected, because a distribution of physical dust temperatures will change the slope of the SED and thus the β in

[9] Note that the issues discussed here are independent of observational noise, as we have added none to our simulated SEDs. Observational noise further complicates interpretation of the derived parameters (Shetty et al. 2009a; Kelly et al. 2012).

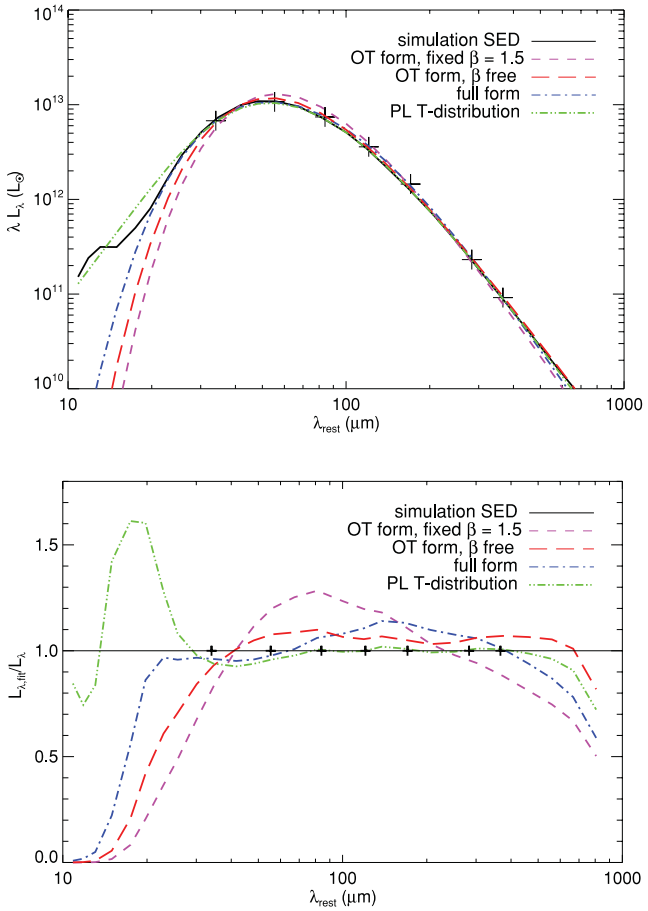

Figure 8. Top: the black line is an example rest-frame FIR SED, λL_λ ($L_\odot$) versus $\lambda_{\rm rest}$ (μm), of a simulated starburst. The crosses are the PACS 100- and 160-μm photometry, SCUBA 250-, 350- and 500-μm photometry, SCUBA 850-μm photometry and AzTEC 1.1-mm photometry calculated from the simulated galaxy's SED. The other lines are best fits to the photometric points for different fitting forms: the single-T, OT modified blackbody (equation 6) with fixed $\beta = 1.5$ (magenta dashed), the single-T OT modified blackbody with β allowed to vary (red long-dashed), the full form of the single-T modified blackbody (equation 1; blue dash dot) and the power-law T-distribution (Kovács et al. 2010; green double-dot–dashed). Bottom: ratio of the fitted L_λ to the actual L_λ versus $\lambda_{\rm rest}$ (μm). The crosses overlaid on the black solid line, which denotes $L_{\lambda,\rm fit} = L_\lambda$, show the wavelengths of the photometric points used for the fits. The values of $T_{\rm d}$ and β for the best-fitting models are given in Table 5. All forms except the single-T OT form with fixed β recover the photometry within the 10 per cent uncertainty. However, the derived dust temperature and β depend strongly on the fitting method, so the model parameters should not be interpreted physically even though the models adequately describe the data.

Table 5. $T_{\rm d}$ and β for the best-fitting models shown in Fig. 8.

Model	$T_{\rm d}$ (K)	β
Single-T OT form, fixed β	47	1.5
Single-T OT form	53	1.1
Single-T full form	69	1.5
Power-law T-distribution	49	1.2

a single-T model; non-negligible optical depths in the IR further complicate the picture.

Clearly it is necessary to determine how the fitted parameters relate to intrinsic properties of the dust, but we defer further exploration of this complex topic to future work. We only wish to stress that it may be necessary to use forms more sophisticated than the single-T OT modified blackbody to fit IR SEDs and that the parameters derived from the fits should not be interpreted physically. Instead, the models should be thought of as useful ways to encapsulate the data with a few parameters and thus compare galaxy SEDs in a simple way by comparing the fitted parameters. Put another way, the $T_{\rm d}$–$L_{\rm IR}$ plot still contains useful information about SED variation even if $T_{\rm d}$ is not a physical dust temperature, and differences in $T_{\rm d}$ amongst galaxies reflect real differences in the galaxies' SEDs.

3.2 Star formation efficiency

Since starbursts form stars much more efficiently than quiescently star-forming galaxies, one should be able to distinguish between them via some observational indicator of SFE. To be consistent with the literature (e.g. Daddi et al. 2010; Genzel et al. 2010), we define 'star formation efficiency' as SFE $\equiv L_{\rm IR}/M_{\rm gas}$. Note, however, that $L_{\rm IR}$ does not necessarily track the instantaneous SFR because, in addition to recently formed stars, older stars and AGN can also contribute to $L_{\rm IR}$ (see e.g. section 2.5 of Kennicutt 1998a and section 3.1 of H11, for details), so this should be considered only an approximate indicator of SFE. Furthermore, inferring the total molecular gas mass from CO observations is notoriously difficult, as the CO–H_2 conversion factor $X_{\rm CO} = N_{\rm H2}/L'_{\rm CO}$ depends on the giant molecular cloud surface density and the kinetic temperature and velocity dispersion within clouds (Narayanan et al. 2011b; Shetty et al. 2011a,b; Narayanan et al. 2012; Papadopoulos et al. 2012). As a result, $X_{\rm CO}$ is expected to be a factor of ∼2–10 lower in starbursts than in disc galaxies (Narayanan et al. 2011b). The uncertainty surrounding $X_{\rm CO}$ complicates efforts to determine how much the SFE of starbursts and quiescently star-forming galaxies differs (e.g. Papadopoulos et al. 2012). It would be best for us to predict the CO line luminosity for our simulated galaxies, as has been done by, for example, Narayanan et al. (2009), and, ideally, to self-consistently track formation and destruction of molecular gas (see e.g. Robertson & Kravtsov 2008). However, doing so requires introduction of another code – and the corresponding complexities and uncertainties – in addition to the two employed, so we feel this is best left to future work (though we briefly discuss how the molecular gas emission might be used as a diagnostic in Section 4.4). Thus when calculating SFE here we use the total gas mass instead of the molecular gas mass.

In Fig. 9 we plot SFE versus $L_{\rm IR}$ (top) and nuclear separation (bottom); the median values are given in Table 6. At a given $L_{\rm IR}$, the starburst SMGs have SFE up to five times greater than that of the quiescently star-forming galaxies. The starbursts not selected as SMGs have the highest SFE, whereas the quiescently star-forming SMGs have the lowest values; the discrepancy can be as much as ∼10 times. The bottom panel of Fig. 9 demonstrates that for both the SMGs and non-SMGs, the SFE increases as the merger advances and is highest for mergers nearest coalescence. At a given $d_{\rm BH}$ the SMGs and non-SMGs span a similar range in SFE. The objects with high SFE at $d_{\rm BH} \sim 100$ kpc are mergers observed shortly after the first passage. For these galaxies $L_{\rm IR}$ is still high, causing high SFE, but instantaneous SFR is not, so they do not meet the definition of starburst.

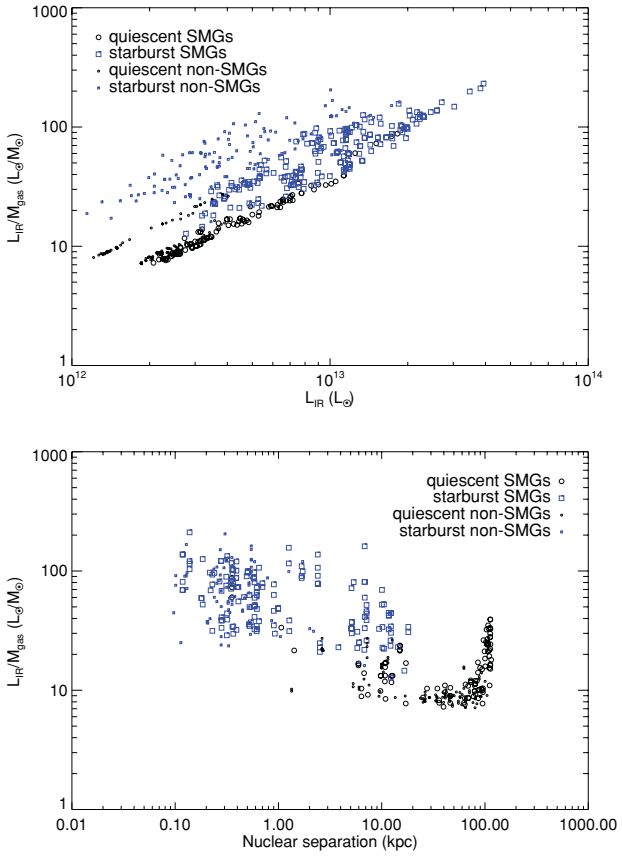

Figure 9. 'Star formation efficiency' SFE $\equiv L_{\rm IR}/M_{\rm gas}$ ($L_\odot/M_\odot$) versus $L_{\rm IR}$ ($L_\odot$; top) and nuclear separation (kpc; bottom). SFE and $L_{\rm IR}$ are totals for the system in order to mimic the effects of blending in the FIR and (sub)mm. At fixed $L_{\rm IR}$, the SFE is characteristically higher by a factor of a few for starburst SMGs than for quiescently star-forming SMGs, and the starbursts (quiescently star-forming galaxies) missed by the SMG selection have SFE greater than the starburst (quiescently star-forming) SMGs. The spread in SFE at fixed $L_{\rm IR}$ between the most extreme starburst non-SMGs and the quiescently star-forming SMGs can be as much as ∼10 times. The starbursts near coalescence have the highest SFE. The objects with high SFE at $d_{\rm BH} \sim 100$ kpc are the same as the high-$T_{\rm d}$ objects with $d_{\rm BH} \sim 100$ kpc, mergers observed shortly after the first passage when $L_{\rm IR}$ is still high, but instantaneous SFR is not.

Table 6. Median SFE values ($L_\odot/M_\odot$).

Galaxy type	$L_{\rm IR} < 10^{13}\,L_\odot$	$L_{\rm IR} \geq 10^{13}\,L_\odot$
Quiescent SMGs	13.6	60.3
Starburst SMGs	37.1	91.2
Quiescent non-SMGs	9.4	–
Starburst non-SMGs	51.2	123.4

The simulations qualitatively reproduce the systematic offset in SFE between quiescently star-forming galaxies and starbursts shown in fig. 1 of Daddi et al. (2010). However, for a given $L_{\rm IR}$ the SFE values of the simulated starbursts are significantly (∼5 times) lower than the observed values. Part of this is because our measure of the SFE in the simulations uses the total gas mass, not the molecular gas mass. In the simulations at the peak of the starburst, cold gas within 5 kpc of the centre, which we take as a rough approximation for molecular gas that would be probed by observations (see Narayanan et al. 2009), is typically less than half of the total gas mass even though it accounts for effectively all of the star formation. Thus if we were to calculate SFE for our simulated starbursts using molecular gas mass rather than total gas mass, the values for the simulations would be a factor of ∼2–3 higher, which would account for a large part of the discrepancy between the simulated and observed starbursts.

The SFE of the simulated starburst SMGs is only a factor of ∼2–3 greater than that of the simulated quiescently star-forming galaxies, whereas the observed difference is ∼4–10 times. For the starbursts nearest coalescence, however, the difference can be as great as 10 times (see bottom panel of Fig. 9). One possible reason the SFE discrepancy is lower than observed is that, because of the set-up of our simulations, the pre-burst, quiescently star-forming discs have systematically higher gas fractions than those discs near coalescence. Since the star formation law implemented in the simulations has a non-linear dependence on gas density, the SFE will increase with gas fraction. Despite the possible discrepancies described above, the simulations make the robust prediction that those systems with largest SFE at a given $L_{\rm IR}$ will be merger-induced starbursts.

3.3 IR excess

In Fig. 10 we plot the total IR luminosity divided by the rest-frame FUV luminosity, $L_{\rm IR}/L_{\rm FUV}$ – referred to as the IRX – versus $L_{\rm IR}$ (top) and nuclear separation (bottom); the median values are given in Table 7. Both IRX and $L_{\rm IR}$ are calculated for the total system. The IRX serves as a measure of the level of obscuration of a galaxy. IRX increases with $L_{\rm IR}$, as has previously been both observed (e.g. Wang & Heckman 1996; Buat & Burgarella 1998; Buat et al. 1999, 2005, 2007, 2009; Adelberger & Steidel 2000; Hopkins et al. 2001; Bell 2003; Reddy et al. 2010) and demonstrated by simulations (Jonsson et al. 2006). The quiescently star-forming galaxies tend to have lower IRX than the starbursts, primarily because the starbursts are typically more luminous. At a given $L_{\rm IR}$, the starbursts and quiescently star-forming discs have very similar IRX. Jonsson et al. (2006) have previously demonstrated this result using similar simulations, and they argue that the correlation arises because both SFR and dust optical depth correlate with density. The bottom panel shows that IRX increases as nuclear separation decreases. This adds further evidence in support of the arguments of Jonsson et al. (2006): The galaxies are more compact at coalescence than during the pre-burst, infalling-disc stage. The resulting higher densities lead to higher SFR and thus greater $L_{\rm IR}$. Furthermore, the stars formed in the starburst, which dominate the luminosity, are centrally concentrated and thus typically more obscured than stars distributed throughout the initial discs. Thus $L_{\rm IR}$ and IRX both increase as nuclear separation decreases. This finding is consistent with that from studies of AGN obscuration using similar simulations (Hopkins et al. 2005a,b, 2006).

The solid line in the top panel of Fig. 10 is the relation given by constant $L_{\rm FUV} = 4 \times 10^{10}\,L_\odot$. Note that, assuming this value, $L_{\rm FUV}$ is less than 1 per cent of the bolometric luminosity. At the bright end this relation approximates that of the simulated galaxies reasonably well, suggesting that the light observed in the rest-frame FUV is decoupled from the bolometric luminosity. This what we expect when the luminosity of a galaxy is dominated by a deeply dust-enshrouded starburst and the only light observed in the FUV is from stars located outside the heavily obscured nuclear region (see also Jonsson et al. 2006). At lower luminosities $L_{\rm FUV}$ decreases as $L_{\rm IR}$ decreases, causing the simulated galaxies to lie above the $L_{\rm FUV} = 4 \times 10^{10}\,L_\odot$ relation. In particular, the quiescently

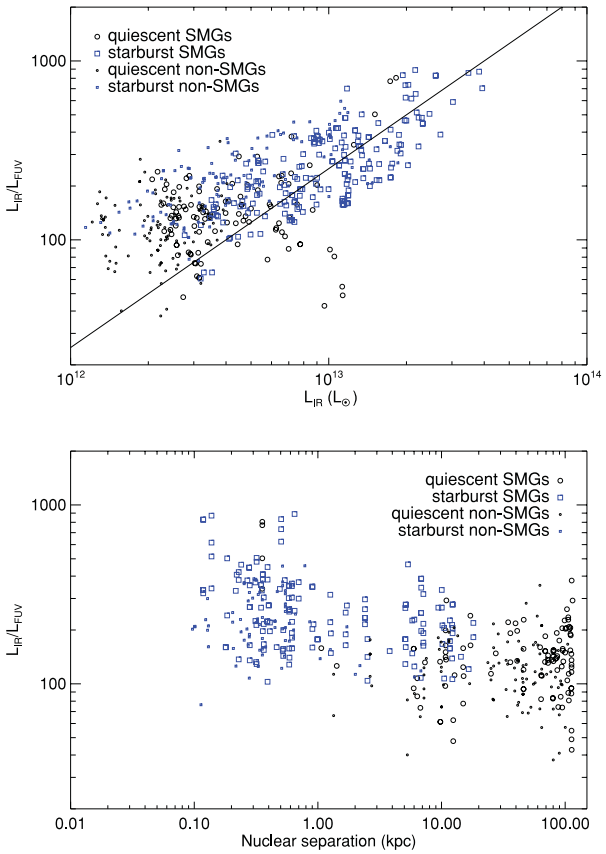

Figure 10. IR excess (IRX $\equiv L_{\rm IR}/L_{\rm FUV}$) versus $L_{\rm IR}$ ($L_\odot$; top) and nuclear separation (kpc; bottom). Both IRX and $L_{\rm IR}$ are calculated for the total system. For the starbursts, IRX is correlated with $L_{\rm IR}$ and anticorrelated with nuclear separation. The starbursts are typically more obscured (have higher IRX) than the quiescently star-forming galaxies. However, for fixed $L_{\rm IR}$ the starbursts and quiescently star-forming discs have similar IRX values, so the reason starbursts typically have higher attenuation is because they are typically more luminous than the quiescently star-forming discs. Interestingly, the correlation is reasonably well approximated by constant $L_{\rm FUV} = 4 \times 10^{10}\,{\rm L}_\odot$ (solid line in the top panel) at the bright end.

Table 7. Median IRX values.

Galaxy type	$L_{\rm IR} < 10^{13}\,{\rm L}_\odot$	$L_{\rm IR} \geq 10^{13}\,{\rm L}_\odot$
Quiescent SMGs	144.9	340.0
Starburst SMGs	185.4	332.3
Quiescent non-SMGs	120.1	–
Starburst non-SMGs	199.9	373.3

star-forming galaxies tend to lie above the relation because they are much less obscured than the starbursts.

3.4 SFR–$M_\star$ relation

As discussed in Section 1.1, if the tight SFR–$M_\star$ relation observed is set by the gas supply rate, then galaxies that lie above the relation must be undergoing transient events that temporarily boost their SFR above what can be sustained over long time periods. Major mergers are one type of event that can cause galaxies to move

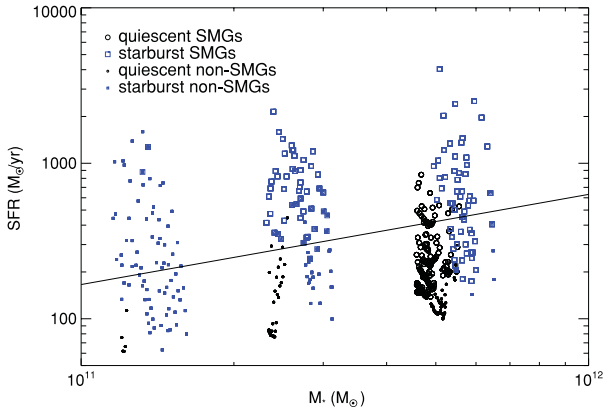

Figure 11. SFR ($M_\odot\,{\rm yr}^{-1}$) versus $M_\star$ ($M_\odot$), where both are totals for the system. The solid line is the relation for the $z = 2.0$–2.5 bin of Karim et al. (2011). Simulated galaxies above the observed relation are almost exclusively starbursts, and the starbursts typically lie above the relation.

above the relation. We plot the SFR–$M_\star$ relation for our simulated galaxies in Fig. 11 along with the observed relation for $z = 2.0$–2.5 from Karim et al. (2011; solid line). Almost all objects above the observed relation are starbursts, and a large fraction of the starbursts are above the relation. For a given $M_\star$, the starbursts can have SFR values $\gtrsim$10 times those of the quiescently star-forming galaxies.

Note that some of the simulated galaxies, including starbursts, lie significantly below the relation. This is partially caused by the idealized set-up of the simulations: since there is no cosmological gas accretion, some of the simulated galaxies have gas fractions lower than what is expected from observations and cosmological simulations. For fixed galaxy size and total mass, the observed KS relation implies that SFR scales superlinearly with gas fraction, so simulated galaxies with gas fractions lower than real galaxies will have significantly lower SFR than observed. Note also that the pre-burst discs have systematically higher gas fraction because the gas fraction decreases monotonically with time in the simulations. This biases the quiescently star-forming galaxies' SFR high relative to the starbursts, so the magnitude of the difference between starbursts and quiescently star-forming discs' SFR shown in Fig. 11 should be taken as a lower limit. Furthermore, discrepancies in the SFR–$M_\star$ relation of the simulations and that observed may occur because the SFR derived using a given diagnostic is not equivalent to the instantaneous SFR of the simulated galaxies, as even the observed relations can differ depending on what SFR diagnostic is used to derive them.

Observationally, whether SMGs lie on the SFR–$M_\star$ relation depends on the measured $L_{\rm IR}$ used to infer the SFR and the inferred $M_\star$. The latter is especially difficult to determine: different authors have inferred masses differing by a factor of $\sim$6 for the same SMGs (Michałowski, Hjorth & Watson 2010a; Michałowski, Watson & Hjorth 2010b; Hainline et al. 2011; Michałowski et al. 2012). If the Michałowski et al. (2012) masses are used, SMGs lie much closer to the SFR–$M_\star$ relation than they do if the Hainline et al. (2011) masses are used. However, in both cases a subset of SMGs are significant outliers from the relation; this conclusion is consistent with our claim that the SMG population is a mix of both quiescently star-forming and starburst galaxies.

4 DISCUSSION

4.1 The need to distinguish star formation modes

If one wishes to understand star formation, it is crucial to look beyond the local Universe, because the SFR density of the Universe was greatest at $z \sim 2$–3 (e.g. Madau et al. 1996; Steidel et al. 1996; Hopkins 2004; Hopkins & Beacom 2006; Hopkins et al. 2010; Karim et al. 2011; Magnelli et al. 2011). Furthermore, the bulk of the star formation at those redshifts was obscured (e.g. Bouwens et al. 2011; Magnelli et al. 2011), so studying IR-luminous galaxies at those redshifts is crucial. Unfortunately, galaxies become fainter and physical resolution poorer as one moves from $z \sim 0$ to higher redshift, so observations of high-redshift galaxies are significantly less detailed than for local galaxies. It is thus tempting to use wisdom gleaned from detailed observations of local galaxies to guide the interpretation of observations of high-redshift galaxies. This is perfectly acceptable if the only difference between local galaxies and those at $z \sim 2$–3 is that the latter are farther away. However, this is clearly not the case, so one must apply local-Universe-derived wisdom with caution.

Assuming what is true locally is also true at $z \sim 2$–3 can be problematic. For example, as discussed in Section 3.2, locally it seems that the CO–H_2 conversion factor X_{CO} differs for ULIRGs (i.e. merger-induced starbursts) and quiescently star-forming disc galaxies (e.g. Solomon et al. 1997; Downes & Solomon 1998). If one wishes to, for example, study possible evolution of the KS law with redshift, then, lacking other options, it is necessary to assume some CO–H_2 conversion factor for the high-redshift galaxy populations observed. Choosing an appropriate X_{CO} requires determining whether the high-redshift galaxies are analogous to local merger-induced starbursts or quiescently star-forming discs. For example, since it is commonly thought that SMGs are merger-driven starbursts, Daddi et al. (2010) and Genzel et al. (2010) use the starburst X_{CO} value for SMGs. If, however, SMGs are a mix of quiescently star-forming, early-stage mergers and late-stage, merger-induced starbursts, as we have argued in H11 and above, then a single X_{CO} value is not appropriate for the population. In this case, the use of the ULIRG X_{CO} value will artificially accentuate the apparent differences between SMGs and more typical galaxies at high redshift.

These diagnostics can also be used to distinguish the quiescently star-forming subpopulations of SMGs (blended galaxy pairs and isolated discs) from the starbursts in order to test the claims of our model and to understand the true nature of the population. We argue that galaxy pairs must contribute significantly to the SMG population (Hayward et al. 2011b; Hayward et al., in preparation) because of the weak scaling of submm flux with SFR in starbursts and the significantly longer duration of the galaxy-pair phase. However, given the modelling uncertainties, it is crucial to observationally determine the relative contributions.

Furthermore, if one wishes to understand which mode of star formation dominates the SFR density of the Universe, one must be able to separate the modes. Even when one can clearly identify mergers (e.g. by the presence of tidal features), one cannot assume that those galaxies are dominated by merger-induced star formation, as the SFR elevation caused by the mutual tidal torques is significant only near coalescence and, depending on the gas content and bulge fraction of the progenitors, possibly the first passage. (See Hopkins & Hernquist 2010; Hopkins et al. 2010, for further discussion of the distinction between star formation in mergers and merger-induced star formation.) The problem is amplified at higher redshifts when mergers cannot be easily identified.

4.2 An observational road map to determine what star formation mode powers high-redshift ULIRGs

Fortunately, the integrated SED of a galaxy contains much information about the star formation mode powering it, so it is possible to use the diagnostics we have presented here to observationally disentangle what star formation mode dominates high-redshift ULIRGs. This can be achieved by applying the diagnostics to the results of FIR and (sub)mm wide-field surveys. Since the diagnostics rely on integrated data alone, they are robust to blending in the FIR and (sub)mm, but the beam sizes at different wavelengths should be similar because otherwise blending will be more severe at the wavelengths where resolution is poorer. The FIR and (sub)mm data are enough to use the T_d–L_{IR} relation as a diagnostic. Some of the other diagnostics require data at shorter wavelengths (to determine L_{FUV} and $M_\star$), so care must be taken to include all sources that contribute to the FIR emission, not just one component of a multiple component system. Keeping in mind the caveats discussed above, one can also estimate gas masses and use the SFE as a diagnostic.

Once ALMA SMG surveys are available, it will be simple to divide SMGs into single-component and multiple-component subclasses. This information can be combined with L_{IR} and T_d values (preferably derived using the power-law T-distribution model) to effectively divide the SMG population into the various subpopulations: galaxy-pair SMGs will be resolved into two components, and those with smaller separations between components should typically have higher L_{IR} and T_d values. The merger-induced starbursts and isolated discs will only have one component, but the former can be distinguished by their higher T_d values. Observations at rest-frame UV–NIR wavelengths can be used to apply the IRX and SFR–$M_\star$ diagnostics to further check the classification done in the above manner, as starbursts will have higher values of IRX and lie above the SFR–$M_\star$ relation. If gas masses are available, one can also use the SFE as a diagnostic.

Spatial information beyond the number of components can further aid the classification. Though IFU spectrograph data provide high-resolution kinematics, high-redshift ULIRGs can be optically thick well into the IR, so even rest-frame NIR observations may not probe the central regions and thus must be interpreted with caution.[10] Molecular gas emission, on the other hand, suffers significantly less dust attenuation, so (sub)mm interferometry with, for example, ALMA can provide a direct view of the central regions (but see Papadopoulos, Isaak & van der Werf 2010a; Papadopoulos et al. 2010b). The close pairs and isolated discs are most effectively distinguished via interferometry because the former will be resolved into two components with disc-like kinematics (see e.g. Engel et al. 2010), whereas the latter will only show one disc. The merger-induced starbursts may show more disordered kinematics, but it is not always simple to distinguish mergers shortly after coalescence from discs (see Robertson & Bullock 2008). Luckily, the starbursts can be distinguished from discs using the diagnostics presented here.

By combining various data sets in the manner described above, it should be possible to separate high-redshift ULIRGs into starburst and quiescently star-forming subpopulations. In addition, mergers can be roughly divided into widely separated pairs and close pairs, which are both quiescently star-forming and starbursts near

[10] Rothberg & Fischer (2010) have demonstrated that at $\lambda = 0.65$ μm, dust obscures the nuclear discs of young stars in local (U)LIRGs; since the galaxies studied here are more obscured, this effect should be more severe for high-redshift ULIRGs.

coalescence. One can then more efficiently target sources for detailed follow-up, focusing on, for example, different stages in the merger process. Furthermore, by applying this technique to SMGs one can compare the sizes of the subpopulations to test our claim that the SMG population is heterogeneous.

4.3 Physical differences between SMGs and hot-dust ULIRGs

Throughout this work we have compared the properties of quiescently star-forming and starburst SMGs with simulated galaxies that would not be selected by the SMG selection ('hot-dust ULIRGs'; Chapman et al. 2004, 2008; Casey et al. 2009, 2011; Magnelli et al. 2010; M12). At a given $L_{\rm IR}$, the hot-dust ULIRGs tend to have higher effective dust temperature than the SMGs; this is simply a consequence of the SMG selection because, for fixed β and $L_{\rm IR}$, submm flux can be decreased only by increasing $T_{\rm d}$. Furthermore, at a given $L_{\rm IR}$ the hot-dust ULIRGs have higher SFE values than the SMGs, suggested by Chapman et al. (2008), but the IRX values are similar.

The relative locations of the SMGs and hot-dust ULIRGs on the SFR–$M_\star$ diagram (Fig. 11) provide insight into the physical differences between these two galaxy classes: the most massive galaxies are almost all selected as SMGs, regardless of whether they are quiescently star forming or starbursts, because they are luminous enough to have $S_{850} > 5$ mJy for any reasonable $T_{\rm d}$. At intermediate masses almost all of the starbursts are selected as SMGs, but the quiescently star-forming galaxies are not. At the lowest masses simulated almost no galaxies are selected as SMGs. This reflects the fact that galaxy mass is an important driver of the observed submm flux because it affects both the SFR and the dust mass (see also H11; Michałowski et al. 2012). The smaller galaxies can be very luminous in the IR if they are undergoing a strong starburst, but this results in a relatively hot SED (because $T_{\rm d}$ tends to increase sharply during starbursts; H11) and causes them to be hot-dust ULIRGs rather than SMGs.

Furthermore, H11 noted that the observed-frame submm flux, and thus whether a simulated starburst is selected as an SMG or considered a hot-dust ULIRG, depends on the angle from which the galaxy is viewed. Variation in the SED with viewing angle can cause $T_{\rm d}$ to vary for different lines of sight, but the variation in $T_{\rm d}$ (at most a few degrees) cannot account for the entire variation in submm flux. Instead, the primary cause is that the effective emission area depends on viewing angle. Note that in Fig. 11 there are many cases where large blue squares and small blue squares are centred at the same ($M_\star$,SFR) point. These correspond to simulated starbursts that are classified as both SMGs and hot-dust ULIRGs depending on the angle from which they are viewed. Such cases are relatively common, but there are no quiescently star-forming galaxies for which this occurs; this is most likely because in the quiescently star-forming galaxies most of the FIR emission arises from regions that are OT in the FIR, and thus the effective emitting area in the FIR does not depend strongly on viewing angle. This interpretation is supported by the results presented in Section 3.3.

4.4 Other potential diagnostics

This work has focused on diagnostics that rely on integrated UV–mm SEDs alone, but there are various other diagnostics that could potentially be used to distinguish quiescently star-forming and starburst galaxies. For example, the late-stage merger-induced starburst SMGs may have higher CO linewidths than the quiescently star-forming discs (Narayanan et al. 2009). In addition, the CO spectral line energy distributions (SLEDs) may differ, with the starbursts typically having higher mean excitation than the quiescently star-forming galaxies. Note, however, that the significant dust optical depths present in the starbursts can result in SLEDs that appear indicative of very low excitation conditions, potentially making it difficult to distinguish highly obscured high-excitation gas from unobscured low-excitation gas (Papadopoulos et al. 2010a,b). X-ray observations may also be useful: in our simulations the merger-induced starbursts typically coincide with strong AGN activity, though the peak AGN activity occurs ∼50–100 Myr after the peak SFR (Hopkins 2012). Thus the merger-induced starbursts should have stronger central X-ray sources than the quiescent discs. Furthermore, since the starbursts are typically more highly obscured, the AGN they host should have higher hardness ratios than the AGN hosted by quiescently star-forming galaxies. Finally, radio observations can provide excellent spatial resolution and thus be used to determine the number of components and separation of the nuclei.

5 CONCLUSIONS

We have combined high-resolution 3D hydrodynamic simulations of $z \sim 2$ major mergers of disc galaxies and 3D Monte Carlo dust RT calculations to investigate the differences between quiescently star-forming and starburst galaxies. We have focused on the SMG population as a case study because, as argued in H11 and elaborated in Section 1.2, the SMG population is likely a mix of quiescently star-forming galaxies and merger-induced starbursts. Our models make robust observational predictions for how quiescently star-forming galaxies on the so-called 'main sequence of star formation' should differ from merger-induced starbursts. We present multiple observational diagnostics which can distinguish quiescently star-forming and starburst SMGs based on integrated UV–mm data alone. The testable predictions and observational diagnostics presented in this work include the following.

(i) Effective dust temperature – derived from fitting a single-T OT modified blackbody, the full form of the single-T modified blackbody or a model assuming a power-law distribution of dust temperatures – correlates with $L_{\rm IR}$, and the galaxies in the high-$L_{\rm IR}$, high-$T_{\rm d}$ region of the $T_{\rm d}$–$L_{\rm IR}$ plot are almost exclusively merger-induced starbursts. A $T_{\rm d}$ cut is an effective means to select starbursts from the overall galaxy population, whereas an $L_{\rm IR}$ cut only works well for those galaxies that are also selected as SMGs.

(ii) Star formation efficiency, SFE $\equiv L_{\rm IR}/M_{\rm gas}$, is ∼2–3 times higher for starburst SMGs than quiescently star-forming SMGs at a given $L_{\rm IR}$. The starbursts not detected as SMGs have SFE ∼2 times greater than the starburst SMGs.

(iii) The IR excess, IRX $\equiv L_{\rm IR}/L_{\rm FUV}$, correlates with $L_{\rm IR}$. The starbursts have IRX a factor of ∼5 greater than quiescently star-forming galaxies, primarily because they are typically more luminous than quiescently star-forming galaxies and IRX correlates with $L_{\rm IR}$. At the bright end the correlation is approximately linear, suggesting that $L_{\rm FUV}$ is decoupled from $L_{\rm IR}$ for the most obscured starbursts.

(iv) Effective dust temperature, SFE and IRX are all inversely correlated with nuclear separation because the strength of the merger-driven starburst increases as the galaxies coalesce.

(v) The majority of the simulated starbursts lie above the SFR–$M_\star$ relation, whereas the quiescently star-forming galaxies lie close to it. The observational definition of starbursts as outliers above the relation agrees well with our definition of merger-induced

starbursts as simulated galaxies with instantaneous SFR > 3 times their baseline quiescent SFR.

One can apply these observational diagnostics to test our claim that the SMG population is a mix of quiescently star-forming galaxies and merger-induced starbursts and to constrain the relative contribution of the subpopulations. Furthermore, the tests presented here provide physically motivated ways to observationally separate quiescently star-forming galaxies and starbursts, enabling one to more cleanly study the underlying physics than when heterogeneous samples (e.g. SMGs) are used. Though we have focused on the SMG population here, we have also discussed the hot-dust ULIRGs that would not be selected by the SMG selection. These galaxies tend to be less massive than SMGs. Though they can be as luminous as SMGs, this often requires that they be starbursts, so they have relatively high $T_{\rm d}$ for their $L_{\rm IR}$ and are thus missed by the SMG selection. Hot-dust ULIRGs are also sometimes just SMGs viewed from a different angle, because the submm flux of a simulated starburst can vary significantly with viewing angle.

We have also explored how well various IR SED fitting forms describe our simulated galaxies' SEDs. For a given SED, the fitted $T_{\rm d}$ and β can vary significantly depending on what form is used, so these parameters should not be interpreted physically. However, some trends are robust to choice of fitting method, indicating that the fits still yield useful information about the SED variations amongst galaxies. In future work we will more thoroughly investigate how well the different FIR SED fitting methods describe observed and simulated galaxy SEDs, determine how the inferred dust properties relate to the actual physical dust properties and develop an improved method for fitting the FIR SEDs of large samples of galaxies.

ACKNOWLEDGMENTS

CCH thanks Emanuele Daddi for motivation for this work, Padelis Papadopoulos and Rahul Shetty for lively discussion of IR SED fitting and Shane Bussmann, Diego Muñoz and Dominik Riechers for insight into submm interferometry. We thank the anonymous referee, Caitlin Casey, Helmut Dannerbauer, Michał Michałowski, Alex Pope and Amy Stutz for helpful comments on the manuscript. We thank Volker Springel for providing the non-public version of GADGET-2 used for this work and Brant Robertson for use of his code to scale the initial conditions to high redshift. DK is supported by NASA through Hubble Fellowship grant HST-HF-51276.01-A. PJ acknowledges support by a grant from the W. M. Keck Foundation. The simulations in this paper were performed on the Odyssey cluster supported by the FAS Research Computing Group at Harvard University.

REFERENCES

Adelberger K. L., Steidel C. C., 2000, ApJ, 544, 218
Agertz O. et al., 2007, MNRAS, 380, 963
Alexander D. M., Bauer F. E., Chapman S. C., Smail I., Blain A. W., Brandt W. N., Ivison R. J., 2005a, ApJ, 632, 736
Alexander D. M., Smail I., Bauer F. E., Chapman S. C., Blain A. W., Brandt W. N., Ivison R. J., 2005b, Nat, 434, 738
Alexander D. M. et al., 2008, AJ, 135, 1968
Amblard A. et al., 2010, A&A, 518, L9
Barger A. J., Cowie L. L., Sanders D. B., Fulton E., Taniguchi Y., Sato Y., Kawara K., Okuda H., 1998, Nat, 394, 248
Barnes J., Hernquist L., 1991, ApJ, 370, L65
Barnes J., Hernquist L., 1996, ApJ, 471, 115
Barnes J., Hut P., 1986, Nat, 324, 446
Bauer A., Springel V., 2011, MNRAS, preprint (arXiv:1109.4413)
Baugh C. M., Lacey C. G., Frenk C. S., Granato G. L., Silva L., Bressan A., Benson A. J., Cole S., 2005, MNRAS, 356, 1191
Bell E. F., 2003, ApJ, 586, 794
Biggs A. D., Ivison R. J., 2008, MNRAS, 385, 893
Blain A. W., Smail I., Ivison R. J., Kneib J.-P., Frayer D. T., 2002, Phys. Rep., 369, 111
Bondi H., Hoyle F., 1944, MNRAS, 104, 273
Bothwell M. S. et al., 2010, MNRAS, 405, 219
Bouché N. et al., 2007, ApJ, 671, 303
Bouwens R. J. et al., 2011, preprint (arXiv:1109.0994)
Buat V., Burgarella D., 1998, A&A, 334, 772
Buat V., Donas J., Milliard B., Xu C., 1999, A&A, 352, 371
Buat V. et al., 2005, ApJ, 619, L51
Buat V., Marcillac D., Burgarella D., Le Floc'h E., Takeuchi T. T., Iglesias-Paràmo J., Xu C. K., 2007, A&A, 469, 19
Buat V., Takeuchi T. T., Burgarella D., Giovannoli E., Murata K. L., 2009, A&A, 507, 693
Bush S. J., Cox T. J., Hayward C. C., Thilker D., Hernquist L., Besla G., 2010, ApJ, 713, 780
Bussmann R. S. et al., 2012, ApJ, 744, 150
Capak P. et al., 2008, ApJ, 681, L53
Carilli C. L. et al., 2010, ApJ, 714, 1407
Casey C. M. et al., 2009, MNRAS, 399, 121
Casey C. M., Chapman S. C., Smail I., Alaghband-Zadeh S., Bothwell M. S., Swinbank A. M., 2011, MNRAS, 411, 2739
Chakrabarti S., Fenner Y., Cox T. J., Hernquist L., Whitney B. A., 2008, ApJ, 688, 972
Chapman S. C., Windhorst R., Odewahn S., Yan H., Conselice C., 2003, ApJ, 599, 92
Chapman S. C., Smail I., Blain A. W., Ivison R. J., 2004, ApJ, 614, 671
Chapman S. C., Blain A. W., Smail I., Ivison R. J., 2005, ApJ, 622, 772
Chapman S. C. et al., 2008, ApJ, 689, 889
Chapman S. C. et al., 2010, MNRAS, 409, L13
Clements D. L. et al., 2008, MNRAS, 387, 247
Clements D. L., Dunne L., Eales S., 2010, MNRAS, 403, 274
Conley A. et al., 2011, ApJ, 732, L35
Coppin K. et al., 2008, MNRAS, 384, 1597
Cox T. J., Dutta S. N., Matteo T. D., Hernquist L., Hopkins P. F., Robertson B., Springel V., 2006, ApJ, 650, 791
Cox T. J., Jonsson P., Somerville R. S., Primack J. R., Dekel A., 2008, MNRAS, 384, 386
da Cunha E., Charmandaris V., Díaz-Santos T., Armus L., Marshall J. A., Elbaz D., 2010, A&A, 523, A78
Daddi E. et al., 2005, ApJ, 631, L13
Daddi E. et al., 2007, ApJ, 670, 156
Daddi E. et al., 2010, ApJ, 713, 686
Dale D. A., Helou G., 2002, ApJ, 576, 159
Dale D. A. et al., 2007, ApJ, 655, 863
Dannerbauer H., Daddi E., Riechers D. A., Walter F., Carilli C. L., Dickinson M., Elbaz D., Morrison G. E., 2009, ApJ, 698, L178
Davé R., Finlator K., Oppenheimer B. D., Fardal M., Katz N., Kereš D., Weinberg D. H., 2010, MNRAS, 404, 1355
Di Matteo T., Springel V., Hernquist L., 2005, Nat, 433, 604
Downes D., Solomon P. M., 1998, ApJ, 507, 615
Draine B. T., Li A., 2007, ApJ, 657, 810
Dwek E., 1998, ApJ, 501, 643
Eales S., Lilly S., Gear W., Dunne L., Bond J. R., Hammer F., Fèvre O. L., Crampton D., 1999, ApJ, 515, 518
Elbaz D. et al., 2011, A&A, 533, A119
Engel H. et al., 2010, ApJ, 724, 233
Erb D. K., Steidel C. C., Shapley A. E., Pettini M., Reddy N. A., Adelberger K. L., 2006, ApJ, 646, 107
Faucher-Giguère C.-A., Kereš D., Ma C.-P., 2011, MNRAS, 417, 2982
Fontanot F., Monaco P., 2010, MNRAS, 405, 705
Fontanot F., Monaco P., Silva L., Grazian A., 2007, MNRAS, 382, 903
Genzel R. et al., 2010, MNRAS, 407, 2091
Gingold R. A., Monaghan J. J., 1977, MNRAS, 181, 375

González J. E., Lacey C. G., Baugh C. M., Frenk C. S., 2011, MNRAS, 413, 749
Granato G. L., De Zotti G., Silva L., Bressan A., Danese L., 2004, ApJ, 600, 580
Greve T. R. et al., 2005, MNRAS, 359, 1165
Griffin M. J. et al., 2010, A&A, 518, L3
Groves B., Dopita M. A., Sutherland R. S., Kewley L. J., Fischera J., Leitherer C., Brandl B., van Breugel W., 2008, ApJS, 176, 438
Hainline L. J., Blain A. W., Smail I., Alexander D. M., Armus L., Chapman S. C., Ivison R. J., 2011, ApJ, 740, 96
Hayward C. C., Kereš D., Jonsson P., Narayanan D., Cox T. J., Hernquist L., 2011a, ApJ, 743, 159 (H11)
Hayward C. C., Narayanan D., Jonsson P., Cox T. J., Kereš D., Hopkins P. F., Hernquist L., 2011b, in Treyer M., Wyder T., Neill J., Seibert M., Lee J., eds, ASP Conf. Ser. Vol. 440, Have Observations Revealed a Variable Upper End of the Initial Mass Function? Astron. Soc. Pac., San Francisco, p. 369
Hernquist L., 1989, Nat, 340, 687
Hernquist L., 1990, ApJ, 356, 359
Hernquist L., Katz N., 1989, ApJS, 70, 419
Hildebrand R. H., 1983, QJRAS, 24, 267
Holland W. S. et al., 1999, MNRAS, 303, 659
Hopkins A. M., 2004, ApJ, 615, 209
Hopkins P. F., 2012, MNRAS, 420
Hopkins A. M., Beacom J. F., 2006, ApJ, 651, 142
Hopkins P. F., Hernquist L., 2010, MNRAS, 402, 985
Hopkins A. M., Connolly A. J., Haarsma D. B., Cram L. E., 2001, AJ, 122, 288
Hopkins P. F., Hernquist L., Cox T. J., Di Matteo T., Martini P., Robertson B., Springel V., 2005a, ApJ, 630, 705
Hopkins P. F., Hernquist L., Cox T. J., Di Matteo T., Robertson B., Springel V., 2005b, ApJ, 630, 716
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., Hernquist L., Cox T. J., Kereš D., 2008, ApJS, 175, 356
Hopkins P. F. et al., 2009, MNRAS, 397, 802
Hopkins P. F., Younger J. D., Hayward C. C., Narayanan D., Hernquist L., 2010, MNRAS, 402, 1693
Hoyle F., Lyttleton R. A., 1939, Proc. Cambridge Philos. Soc., 35, 405
Hughes D. H. et al., 1998, Nat, 394, 241
Hwang H. S. et al., 2010, MNRAS, 409, 75
Iono D. et al., 2009, ApJ, 695, 1537
Ivison R. J. et al., 2002, MNRAS, 337, 1
Ivison R. J. et al., 2007, MNRAS, 380, 199
Ivison R. J., Smail I., Papadopoulos P. P., Wold I., Richard J., Swinbank A. M., Kneib J., Owen F. N., 2010, MNRAS, 404, 198
James A., Dunne L., Eales S., Edmunds M. G., 2002, MNRAS, 335, 753
Jonsson P., 2006, MNRAS, 372, 2
Jonsson P., Primack J. R., 2010, New Astron., 15, 509
Jonsson P., Cox T. J., Primack J. R., Somerville R. S., 2006, ApJ, 637, 255
Jonsson P., Groves B. A., Cox T. J., 2010, MNRAS, 403, 17
Juvela M., 2005, A&A, 440, 531
Karim A. et al., 2011, ApJ, 730, 61
Katz N., Weinberg D. H., Hernquist L., 1996, ApJS, 105, 19
Kelly B. C., Shetty R., Stutz A. M., Kauffmann J., Goodman A. A., Launhardt R., 2012, ApJ, 752, 55
Kennicutt R. C., 1998a, ApJ, 498, 541
Kennicutt R. C., 1998b, ARA&A, 36, 189
Kennicutt R. C. et al., 2003, PASP, 115, 928
Kereš D., Katz N., Weinberg D. H., Davé R., 2005, MNRAS, 363, 2
Kereš D., Katz N., Fardal M., Davé R., Weinberg D. H., 2009, MNRAS, 395, 160
Kereš D., Vogelsberger M., Sijacki D., Springel V., Hernquist L., 2011, preprint (arXiv:1109.4638)
Kovács A., Chapman S. C., Dowell C. D., Blain A. W., Ivison R. J., Smail I., Phillips T. G., 2006, ApJ, 650, 592
Kovács A. et al., 2010, ApJ, 717, 29
Krumholz M. R., Thompson T. A., 2007, ApJ, 669, 289
Leitherer C. et al., 1999, ApJS, 123, 3
Lo Faro B., Monaco P., Vanzella E., Fontanot F., Silva L., Cristiani S., 2009, MNRAS, 399, 827
Lonsdale C. J., Farrah D., Smith H. E., 2006, in Mason J. W., ed., Ultraluminous Infrared Galaxies. Springer-Verlag, Berlin, p. 285
Lucy L. B., 1977, AJ, 82, 1013
Lupu R. E. et al., 2010, preprint (arXiv:1009.5983)
Madau P., Ferguson H. C., Dickinson M. E., Giavalisco M., Steidel C. C., Fruchter A., 1996, MNRAS, 283, 1388
Magdis G. E. et al., 2010, MNRAS, 409, 22
Magnelli B. et al., 2010, A&A, 518, L28
Magnelli B., Elbaz D., Chary R. R., Dickinson M., Le Borgne D., Frayer D. T., Willmer C. N. A., 2011, A&A, 528, A35
Magnelli B. et al., 2012, A&A, 539, A155
Menéndez-Delmestre K. et al., 2007, ApJ, 655, L65
Menéndez-Delmestre K. et al., 2009, ApJ, 699, 667
Michałowski M. J., Hjorth J., Watson D., 2010a, A&A, 514, A67
Michałowski M. J., Watson D., Hjorth J., 2010b, ApJ, 712, 942
Michałowski M. J., Dunlop J. S., Cirasuolo M., Hjorth J., Hayward C. C., Watson D., 2012, A&A, 541, A85
Mihos J. C., Hernquist L., 1996, ApJ, 464, 641
Narayanan D., Cox T. J., Shirley Y., Davé R., Hernquist L., Walker C. K., 2008a, ApJ, 684, 996
Narayanan D. et al., 2008b, ApJS, 174, 13
Narayanan D., Cox T. J., Hayward C. C., Younger J. D., Hernquist L., 2009, MNRAS, 400, 1919
Narayanan D., Hayward C. C., Cox T. J., Hernquist L., Jonsson P., Younger J. D., Groves B., 2010a, MNRAS, 401, 1613
Narayanan D. et al., 2010b, MNRAS, 407, 1701
Narayanan D., Cox T. J., Hayward C. C., Hernquist L., 2011a, MNRAS, 412, 287
Narayanan D., Krumholz M., Ostriker E. C., Hernquist L., 2011b, MNRAS, 418, 664
Narayanan D., Krumholz M. R., Ostriker E. C., Hernquist L., 2012, 421, 3127
Neri R. et al., 2003, ApJ, 597, L113
Noeske K. G. et al., 2007a, ApJ, 660, L47
Noeske K. G. et al., 2007b, ApJ, 660, L43
Nordon R. et al., 2012, ApJ, 745, 182
Papadopoulos P. P., Isaak K., van der Werf P., 2010a, ApJ, 711, 757
Papadopoulos P. P., van der Werf P., Isaak K., Xilouris E. M., 2010b, ApJ, 715, 775
Papadopoulos P. P., van der Werf P., Xilouris E., Isaak K. G., Gao Y., 2012, ApJ, 751, 10
Poglitsch A. et al., 2010, A&A, 518, L2
Pope A. et al., 2006, MNRAS, 370, 1185
Pope A. et al., 2008, ApJ, 675, 1171
Reddy N. A., Erb D. K., Pettini M., Steidel C. C., Shapley A. E., 2010, ApJ, 712, 1070
Ricciardelli E., Trujillo I., Buitrago F., Conselice C. J., 2010, MNRAS, 406, 230
Riechers D. A. et al., 2011a, ApJ, 733, L11
Riechers D. A., Hodge J., Walter F., Carilli C. L., Bertoldi F., 2011b, ApJ, 739, L31
Robertson B. E., Bullock J. S., 2008, ApJ, 685, L27
Robertson B. E., Kravtsov A. V., 2008, ApJ, 680, 1083
Robertson B., Hernquist L., Cox T. J., Matteo T. D., Hopkins P. F., Martini P., Springel V., 2006, ApJ, 641, 90
Rodighiero G. et al., 2010, A&A, 518, L25
Rodighiero G. et al., 2011, ApJ, 739, L40
Rothberg B., Fischer J., 2010, ApJ, 712, 318
Sajina A., Scott D., Dennefeld M., Dole H., Lacy M., Lagache G., 2006, MNRAS, 369, 939
Sanders D. B., Mirabel I. F., 1996, ARA&A, 34, 749
Schmidt M., 1959, ApJ, 129, 243
Shakura N. I., Sunyaev R. A., 1973, A&A, 24, 337
Shetty R., Kauffmann J., Schnee S., Goodman A. A., 2009a, ApJ, 696, 676

Shetty R., Kauffmann J., Schnee S., Goodman A. A., Ercolano B., 2009b, ApJ, 696, 2234
Shetty R., Glover S. C., Dullemond C. P., Klessen R. S., 2011a, MNRAS, 412, 1686
Shetty R., Glover S. C., Dullemond C. P., Ostriker E. C., Harris A. I., Klessen R. S., 2011b, MNRAS, 415, 3253
Sijacki D., Vogelsberger M., Kereš D., Springel V., Hernquist L., 2011, preprint (arXiv:1109.3468)
Smail I., Ivison R. J., Blain A. W., 1997, ApJ, 490, L5
Smail I., Chapman S. C., Blain A. W., Ivison R. J., 2004, ApJ, 616, 71
Snyder G. F., Cox T. J., Hayward C. C., Hernquist L., Jonsson P., 2011, ApJ, 741, 77
Solomon P. M., Downes D., Radford S. J. E., Barrett J. W., 1997, ApJ, 478, 144
Springel V., 2005, MNRAS, 364, 1105
Springel V., 2010a, ARA&A, 48, 391
Springel V., 2010b, MNRAS, 401, 791
Springel V., Hernquist L., 2002, MNRAS, 333, 649
Springel V., Hernquist L., 2003, MNRAS, 339, 289
Springel V., Hernquist L., 2005, ApJ, 622, L9
Springel V., Yoshida N., White S. D. M., 2001, New Astron., 6, 79
Springel V., Di Matteo T., Hernquist L., 2005, MNRAS, 361, 776
Steidel C. C., Giavalisco M., Pettini M., Dickinson M., Adelberger K. L., 1996, ApJ, 462, L17
Swinbank A. M., Smail I., Chapman S. C., Blain A. W., Ivison R. J., Keel W. C., 2004, ApJ, 617, 64
Swinbank A. M. et al., 2008, MNRAS, 391, 420
Tacconi L. J. et al., 2006, ApJ, 640, 228
Tacconi L. J. et al., 2008, ApJ, 680, 246
Tacconi L. J. et al., 2010, Nat, 463, 781
Targett T. A., Dunlop J. S., McLure R. J., Best P. N., Cirasuolo M., Almaini O., 2011, MNRAS, 412, 295
Torrey P., Vogelsberger M., Sijacki D., Springel V., Hernquist L., 2011, preprint (arXiv:1110.5635)
Valiante E., Lutz D., Sturm E., Genzel R., Tacconi L. J., Lehnert M. D., Baker A. J., 2007, ApJ, 660, 1060
Veilleux S., Kim D.-C., Sanders D. B., 2002, ApJS, 143, 315
Vogelsberger M., Sijacki D., Kereš D., Springel V., Hernquist L., 2011, preprint (arXiv:1109.1281)
Wang B., Heckman T. M., 1996, ApJ, 457, 645
Wang W.-H., Cowie L. L., Barger A. J., Williams J. P., 2011, ApJ, 726, L18
Weingartner J. C., Draine B. T., 2001, ApJ, 548, 296
Wilson G. W. et al., 2008, MNRAS, 386, 807
Wuyts S. et al., 2009, ApJ, 700, 799
Wuyts S., Cox T. J., Hayward C. C., Franx M., Hernquist L., Hopkins P. F., Jonsson P., van Dokkum P. G., 2010, ApJ, 722, 1666
Wuyts S. et al., 2011a, ApJ, 738, 106
Wuyts S. et al., 2011b, ApJ, 742, 96
Younger J. D. et al., 2008, ApJ, 688, 59
Younger J. D. et al., 2009a, ApJ, 704, 803
Younger J. D., Hayward C. C., Narayanan D., Cox T. J., Hernquist L., Jonsson P., 2009b, MNRAS, 396, L66
Younger J. D. et al., 2009c, MNRAS, 394, 1685
Younger J. D. et al., 2010, MNRAS, 407, 1268
Yun M. S. et al., 2012, MNRAS, 420, 957

This paper has been typeset from a TEX/LATEX file prepared by the author.