

LETTER

The Formation of Submillimetre-Bright Galaxies from Gas Infall over a Billion Years

Desika Narayanan^{1,13}, Matthew Turk², Robert Feldmann^{3,12}, Thomas Robitaille⁴, Philip Hopkins⁵, Robert Thompson^{6,7}, Christopher Hayward^{5,8}, David Ball^{7,9}, Claude-André Faucher-Giguère¹⁰, Dušan Kereš¹¹

Submillimetre-luminous galaxies at high-redshift are the most luminous, heavily star-forming galaxies in the Universe¹, and are characterised by prodigious emission in the far-infrared at 850 microns ($S_{850} \geq 5$ mJy). They reside in halos $\sim 10^{13} M_{\odot}$ ², have low gas fractions compared to main sequence disks at a comparable redshift³, trace complex environments^{4,5}, and are not easily observable at optical wavelengths⁶. Their physical origin remains unclear. Simulations have been able to form galaxies with the requisite luminosities, but have otherwise been unable to simultaneously match the stellar masses, star formation rates, gas fractions and environments⁷⁻¹⁰. Here we report a cosmological hydrodynamic galaxy formation simulation that is able to form a submillimetre galaxy which simultaneously satisfies the broad range of observed physical constraints. We find that groups of galaxies residing in massive dark matter halos have rising star formation histories that peak at collective rates $\sim 500 - 10^3 M_{\odot} \text{yr}^{-1}$ at $z = 2 - 3$, by which time the interstellar medium is sufficiently enriched with metals that the region may be observed as a submillimetre-selected system. The intense star formation rates are fueled in part by a reservoir gas supply enabled by stellar feedback at earlier times, not through major mergers. With a duty cycle of nearly a gigayear, our simulations show that the submillimetre-luminous phase of high- z galaxies is a drawn-out one that is associated with significant mass buildup in early Universe proto-clusters, and that many submillimetre-luminous galaxies are actually composed of numerous unresolved components (for which there is some observational evidence¹¹).

We have conducted our cosmological hydrodynamic zoom galaxy formation simulations utilising the new hydrodynamic code GIZMO¹², and include a model for the impact of radiative and thermal pressure from stars on the multiphase interstellar medium. This feedback both regulates the star formation rate, and shapes the structure in the interstellar medium. Informed by clustering measurements of observed SMGs², we focus on a massive ($M_{\text{DM}} \sim 10^{13} M_{\odot}$ at $z = 2$) halo with baryonic particle mass $M_{\text{bary}} \sim 10^5 M_{\odot}$ as the host of our “main galaxy”, and run the simulation to $z = 2$. The only condition of the tracked galaxy pre-selected to match the physical properties of observed SMGs is the chosen halo mass. We combine this with a new

dust radiation transport package, POWDERDAY, that simulates the traverse of stellar photons through the dusty ISM of the galaxy, allowing us to robustly translate our hydrodynamic simulation into observable measures. We simulate the radiative transfer from a 200 kpc region around the main galaxy. This simulation represents the first cosmological model of a galaxy this massive to be explicitly coupled with dust radiative transfer calculations. The details of both codes and the simulation setup are fully described in the Methods section.

We define two distinct regions in the simulations. The “submillimetre emission region” is the 200 kpc region surrounding the central galaxy in the halo of interest. This is the region where all of the modeled 850 μm emission comes from, and is what relates most directly to observations. The “submillimetre galaxy” refers to the central galaxy in the halo. Physical quantities from the submillimetre galaxy are most applicable to high-resolution observations, as well as placing these models in the context of other theoretical galaxy formation models. As we will show, the submillimetre emission from the region is generally dominated by the central submillimetre galaxy, though the contribution from lower mass galaxies is often non-negligible.

We track the submillimetre properties of the galaxies within the region from $z \sim 6$. The star formation rates of galaxies in the region rise from this redshift toward later times $z \approx 2$, owing to accretion of gas from the intergalactic medium (Figure 1). As stars form, stellar feedback-driven galactic winds generate outflows and fountains allowing recycled gas to be available for star formation at later times (ED Fig 1). This phenomena shapes a star formation history that is still rising at $z \sim 2$, in contrast to galaxy formation models with more traditional implementations of subresolution feedback, which peak at $z \sim 3 - 6$ for galaxies of this mass¹³⁻¹⁵. Mergers and global instabilities drive short-term variability in the global SFR, while outflows and infall driven by the feedback model can impact features in the star formation history with a somewhat cyclical ‘saw-tooth’ pattern.

At its earliest stages ($z \sim 4-6$), the integrated SFR from the galaxies in the region varies from $\sim 100 - 300 M_{\odot} \text{yr}^{-1}$, with a significant stellar mass ($0.5 - 1 \times 10^{11} M_{\odot}$) in place, comparable to some high-redshift detections¹⁶. Feedback from massive stars enrich the interstellar medium with metals, and the dust content simultaneously rises. By $z \approx 3$, the combination of gas accumulation with substantial metal enrichment drives a factor of ~ 50 increase in the dust mass, with masses approaching $\sim 1 \times 10^9 M_{\odot}$. Radiation from the delayed peak in the star formation rate interacting with this substantive dust reservoir drives the observed 850 μm flux density to detectable values of > 5 mJy. The galaxies associated with the main halo enter a long-lived submillimetre-luminous phase, with a duty cycle of ~ 0.75 Gyr. While our main model is only run to $z = 2$ owing to computational restrictions for models of this resolution, tests with lower-resolution models reveal that at later times ($z < 1.5$), a declining star formation rate due to inefficient accretion as well as exhausted gas supply drives a drop

¹Haverford College, 370 W Lancaster Ave, Haverford, PA 19041, USA

²National Center for Supercomputing Applications, University of Illinois, 1205 W. Clark Street, Urbana-Champaign, IL, 61820, USA

³Department of Astronomy and Theoretical Astrophysics Center, University of California, Berkeley, CA 94720 USA

⁴Max Planck Institute for Astronomy, Konigstuhl 17, D-69117, Heidelberg, Germany

⁵TAPIR, California Institute of Technology, MC 350-17, Pasadena, CA 91125, USA

⁶University of the Western Cape, Bellville, Cape Town, South Africa, 7535

⁷Steward Observatory, University of Arizona, 933 N Cherry Ave, Tucson, AZ 85721

⁸Harvard-Smithsonian Center for Astrophysics, 60 Garden St. Cambridge, MA 02138

⁹Whitman College, 345 Boyer Ave, Walla Walla, WA, 99362, USA

¹⁰CIERA, Northwestern University, 2145 Sheridan Road, Evanston, IL, 60208

¹¹CASS, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093

¹²Hubble Fellow

¹³Corresponding Author

in the submillimetre flux density (for more details, see the Methods section). The star formation history of galaxies residing in ($z = 2$) $M_{\text{DM}} \approx 10^{13} M_{\odot}$ halos as controlled by the underlying stellar feedback provides a physical explanation for the peak in the observed SMG redshift distribution at $z = 2 - 3$ ¹⁷.

During the submillimetre-luminous phase, the emitting region is almost always occupied by multiple detectable galaxies. In Figure 2, we present gas surface density projections of six arbitrarily chosen snapshots during the evolution of the submillimetre-luminous phase ($z = 2 - 3$). The panels are 250 kpc on a side; for reference, the full width at half maximum (FWHM) of the Submillimetre Common-user Bolometer Array (SCUBA) on the James Clerk Maxwell Telescope (JCMT), the first instrument to detect SMGs, is ~ 125 kpc at $z \approx 2$. Multiple clumps of gas falling into the central are nearly always present. The observed flux density from the region is typically dominated by the central, with (on average) $\sim 30\%$ arising from emission from subhalos (ED Fig 2). The submillimetre flux density of the central galaxy rises dramatically between $z \sim 2 - 3$, reaching a peak value of ~ 20 mJy. Owing to contributions from subhalos surrounding the central, the flux from the overall 200 kpc region can exceed this, peaking at ~ 30 mJy. Similarly extreme systems have recently been detected with the Herschel Space Observatory and South Pole Telescope^{4,18,19}.

While the central is being bombarded by subhalos over a range of mass ratios during the submillimetre-luminous phase, major galaxy mergers akin to local prototypical analogues such as Arp 220 or NGC 6240 do not drive the onset of the long-lived submillimetre-luminous phase in the central galaxy. In Figure 1, we highlight when the galaxy undergoes a major merger with mass ratio $\geq 1 : 3$. While major mergers are common at early times (and indeed drive some short-lived bursts in star formation), the bulk of the submillimetre-luminous phase at later times ($z \approx 2 - 3$) occurs nearly a gigayear after the last major merger. The ratio of the SFR to its integral over cosmic time (the specific SFR) of the overall emitting region is generally on the main sequence of galaxy formation at $z \sim 2$ (defined as the main locus of points on the SFR- M_* relation), though the central can have values comparable both to main sequence galaxies between $z \sim 2 - 3$, as well as outliers. One consequence of a model in which SMGs typically lie on the main sequence of star formation is that the gas surface densities show a broad range, from $\sim 100 - 10^4 M_{\odot} \text{pc}^{-2}$ (ED Fig 3), as well as diverse gas spatial extents (Figure 3). This is manifested observationally in the broad swath occupied by SMGs in the Kennicutt-Schmidt star formation relation¹. The gaseous spatial extent and surface densities is to be contrasted, however, with local merger-driven ultraluminous infrared galaxies (ULIRGs), which exhibit typical full width at half maximum radii of $\sim 100 - 500 \text{ pc}^{20}$. Idealised galaxy merger simulations with initial conditions designed to form SMGs further underscore this contrast, as they also result in compact morphologies during final coalescence, and can be inefficient producers of submillimetre radiation owing to increased dust temperatures⁸.

The central submillimetre galaxy is amongst the most massive and highly star-forming of galaxies at this epoch. The stellar masses are diverse, and range from $\sim 1 - 5 \times 10^{11} M_{\odot}$, comparable to recent measurements of this population²¹, as well as constraints from abundance matching techniques²². The molecular gas fractions of the central galaxy ($f_{\text{gas}} \equiv M_{\text{H}_2} / (M_{\text{H}_2} + M_*)$) decline with stellar mass, and range from $\sim 40\%$ at lower stellar masses to $\lesssim 10\%$ at the highest masses. This is in agreement with observations²³, though is dependent on the conversion from carbon monoxide (^{12}CO) luminosity to H_2 gas mass. We note that these predictions are quantitatively different from those produced by previous cosmological efforts in this field, with some predicted gas fractions exceeding $f_{\text{gas}} = 0.75$ ^{7,9} and median stellar masses as low as $\sim 10^{10} M_{\odot}$ ⁷. We present the plots highlighting the gas fractions, and calculated SEDs of our model SMG in

the context of observations in ED Figs 4-5. The gas distributions within the central galaxy, which range from $\sim 1 - 8$ kpc, compare well with recent observed dust maps with the Atacama Large Millimetre Array²⁴.

The stellar masses, gas fractions and duty cycles are in agreement with previous lower-resolution cosmological efforts¹⁰, though the predicted SFR and luminosity from this model are substantially larger. The star formation rate of the group of galaxies in the region peaks at $\sim 1500 M_{\odot} \text{ yr}^{-1}$. Importantly, up to half of the total infrared luminosity can come from older stars with ages $t_{\text{age}} > 0.1 \text{ Gyr}$. Utilising standard conversions²⁵, the estimated star formation rate from the integrated infrared SED ($3 - 1100 \mu\text{m}$) can exceed $\sim 3000 M_{\odot} \text{ yr}^{-1}$ (ED Fig 6), and hence infrared-based star formation rate derivations of dusty galaxies at high- z may over-estimate the true SFR by a factor ~ 2 . Indeed, the contribution of satellite galaxies to the global SFR, alongside the contribution of old stars to the infrared luminosity may relieve some tensions between the inferred star formation rates from submillimetre galaxies and massive galaxies modeled in cosmological hydrodynamic simulations¹⁰.

The end-product of the central submillimetre galaxy at $z \sim 2$ is a galaxy with a stellar mass of $\sim 4 - 5 \times 10^{11} M_{\odot}$, distributed over a similarly compact region of $\sim 1 - 5$ kpc as the gas (Figure 3). This is similar in extent and mass to the $z \sim 2$ compact quiescent galaxy population, an observed population with mean half-light radius of $R_e \approx 1.5$ kpc, stellar mass $M_* > 10^{11} M_{\odot}$, and ages $t_{\text{age}} \sim 0.5 - 1 \text{ Gyr}$ ²⁶, suggesting a plausible connection between the galaxy populations. Indeed, a calculation of the stellar velocity dispersion along three orthogonal sightlines of the central during the submillimetre-luminous phase results in $\sigma_* \approx 600 - 700 \text{ km s}^{-1}$, comparable to measurements of high- z compact quiescents. A large sample of simulated SMGs will allow for a robust analysis of expected abundances of SMGs and compact quiescents at $z \sim 2$.

Our picture for SMG formation suggests that they are not transient events, but rather natural long-lived phases in the evolution of massive halos. The $\sim 0.75 \text{ Gyr}$ duty cycle combined with the comoving abundance²⁷ of dark matter halos of this mass results in an expected abundance of our model SMGs of $\sim 1.5 \times 10^{-5} h^3 \text{ Mpc}^{-3}$, comparable to the $\sim 10^{-5} h^3 \text{ Mpc}^{-3}$ observed for SMGs²⁸. While modeling the full number counts involves convolving the typical duty cycle as a function of halo mass with halo mass functions over a range of redshifts, the approximate abundances implied by this model are encouraging.

This model suggests that galaxies that form in halos of mass $M_{\text{DM}} \approx 10^{14}$ at $z = 0$ will represent typical SMGs near the peak of their redshift distribution. Lower mass models show that they do not achieve the requisite star formation rate and metal enrichment to generate submillimetre-luminous galaxies (see Methods section). More extreme SMGs being detected between $z = 5 - 6$ ^{29,30} may form in even more massive (and rare) halos than those considered here.

1. Casey, C. M., Narayanan, D. & Cooray, A. Dusty star-forming galaxies at high redshift. *Physics Reports* **541**, 45–161 (2014).
2. Hickox, R. C. *et al.* The LABOCA survey of the Extended Chandra Deep Field-South: clustering of submillimetre galaxies. *Mon. Not. R. Astron. Soc.* **421**, 284–295 (2012).
3. Geach, J. E. *et al.* On the Evolution of the Molecular Gas Fraction of Star-Forming Galaxies. *Astrophys. J. Let.* **730**, L19 (2011).
4. Fu, H. *et al.* The rapid assembly of an elliptical galaxy of 400 billion solar masses at a redshift of 2.3. *Nature* **498**, 338–341 (2013).
5. Daddi, E. *et al.* Two Bright Submillimetre Galaxies in a $z = 4.05$ Protocluster in Goods-North, and Accurate Radio-Infrared Photometric Redshifts. *Astrophys. J.* **694**, 1517–1538 (2009).
6. Swinbank, A. M. *et al.* The Rest-Frame Optical Spectra of SCUBA Galaxies. *Astrophys. J.* **617**, 64–80 (2004).
7. Baugh, C. M. *et al.* Can the faint submillimetre galaxies be explained in the Λ cold dark matter model? *Mon. Not. R. Astron. Soc.* **356**, 1191–1200 (2005).

8. Hayward, C. C. *et al.* Submillimetre galaxies in a hierarchical universe: number counts, redshift distribution and implications for the IMF. *Mon. Not. R. Astron. Soc.* **428**, 2529–2547 (2013).
9. Shimizu, I., Yoshida, N. & Okamoto, T. Submillimetre galaxies in cosmological hydrodynamic simulations: source number counts and the spatial clustering. *Mon. Not. R. Astron. Soc.* **427**, 2866–2875 (2012).
10. Davé, R. *et al.* The nature of submillimetre galaxies in cosmological hydrodynamic simulations. *Mon. Not. R. Astron. Soc.* **404**, 1355–1368 (2010).
11. Hodge, J. A. *et al.* An ALMA Survey of Submillimeter Galaxies in the Extended Chandra Deep Field South: Source Catalog and Multiplicity. *Astrophys. J.* **768**, 91 (2013).
12. Hopkins, P. F. GIZMO: A New Class of Accurate, Mesh-Free Hydrodynamic Simulation Methods. *arXiv/1409.7395* (2014).
13. Davé, R., Finlator, K. & Oppenheimer, B. D. An analytic model for the evolution of the stellar, gas and metal content of galaxies. *Mon. Not. R. Astron. Soc.* **421**, 98–107 (2012).
14. Hopkins, P. F. *et al.* Galaxies on FIRE (Feedback In Realistic Environments): stellar feedback explains cosmologically inefficient star formation. *Mon. Not. R. Astron. Soc.* **445**, 581–603 (2014).
15. Feldmann, R. & Mayer, L. The Argo Simulation: I. Quenching of Massive Galaxies at High Redshift as a Result of Cosmological Starvation. *arXiv/1404.3212* (2014).
16. Finkelstein, S. L. *et al.* A galaxy rapidly forming stars 700 million years after the Big Bang at redshift 7.51. *Nature* **502**, 524–527 (2013).
17. Weiß, A. *et al.* ALMA Redshifts of Millimeter-selected Galaxies from the SPT Survey: The Redshift Distribution of Dusty Star-forming Galaxies. *Astrophys. J.* **767**, 88 (2013).
18. Ivison, R. J. *et al.* Herschel-ATLAS: A Binary HyLIRG Pinpointing a Cluster of Starbursting Protoellipticals. *Astrophys. J.* **772**, 137 (2013).
19. Hezaveh, Y. D. *et al.* ALMA Observations of SPT-discovered, Strongly Lensed, Dusty, Star-forming Galaxies. *Astrophys. J.* **767**, 132 (2013).
20. Downes, D. & Solomon, P. M. Rotating Nuclear Rings and Extreme Starbursts in Ultraluminous Galaxies. *Astrophys. J.* **507**, 615–654 (1998).
21. Michałowski, M. J. *et al.* The stellar masses and specific star-formation rates of submillimetre galaxies. *Astron. Astrophys.* **541**, A85 (2012).
22. Behroozi, P. S., Wechsler, R. H. & Conroy, C. The Average Star Formation Histories of Galaxies in Dark Matter Halos from $z=0-8$. *arXiv/1207.6105* (2012).
23. Tacconi, L. J. *et al.* Phibss: Molecular Gas Content and Scaling Relations in $z \sim 1-3$ Massive, Main-sequence Star-forming Galaxies. *Astrophys. J.* **768**, 74 (2013).
24. Simpson, J. M. *et al.* The SCUBA-2 Cosmology Legacy Survey: ALMA Resolves the Rest-frame Far-infrared Emission of Sub-millimeter Galaxies. *Astrophys. J.* **799**, 81 (2015).
25. Kennicutt, R. C. & Evans, N. J. Star Formation in the Milky Way and Nearby Galaxies. *Annu. Rev. Astron. Astrophys.* **50**, 531–608 (2012).
26. van Dokkum, P. G. *et al.* Confirmation of the Remarkable Compactness of Massive Quiescent Galaxies at $z \sim 2.3$: Early-Type Galaxies Did not Form in a Simple Monolithic Collapse. *Astrophys. J. Lett.* **677**, L5–L8 (2008).
27. Murray, S. G., Power, C. & Robotham, A. S. G. HMFcalc: An online tool for calculating dark matter halo mass functions. *Astronomy and Computing* **3**, 23–34 (2013).
28. Chapman, S. C., Blain, A. W., Smail, I. & Ivison, R. J. A Redshift Survey of the Submillimeter Galaxy Population. *Astrophys. J.* **622**, 772–796 (2005).
29. Riechers, D. A. *et al.* A dust-obscured massive maximum-starburst galaxy at a redshift of 6.34. *Nature* **496**, 329–333 (2013).
30. Vieira, J. D. *et al.* Dusty starburst galaxies in the early Universe as revealed by gravitational lensing. *Nature* **495**, 344–347 (2013).

Acknowledgements The authors thank Michał J. Michałowski for providing observational data. Partial support for DN was provided by NSF AST-1009452, AST-1442650, NASA HST AR-13906.001, and a Cottrell College Science Award. PFH, CCH, MT and RT were funded by the Gordon and Betty Moore Foundation (GBMF4561 and Grant #776). PFH acknowledges the Alfred P. Sloan Foundation for support. CAFG was supported by NASA awards PF3-140106, NNX15AB22G, and NSF AST-1412836. DK was supported by NSF AST-1412153. RF was supported by NASA HF-51304.01-A. The simulations here were run on Stampede at TACC through NSF XSEDE allocations #TG-AST120025, TG-AST130039, and TG-AST140023, NASA Pleiades, and the Haverford College cluster.

Author Contributions D.N. wrote the text, and led the radiative transfer simulations and analysis. D.N., M.T., T.R. and R.T. wrote the POWDERDAY software. R.T., C.C.H. and D.B. contributed to simulation analysis, and R.F., P.H., C.A.F.G and D.K. performed the cosmological simulations

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to D.N. (dnarayan@haverford.edu).

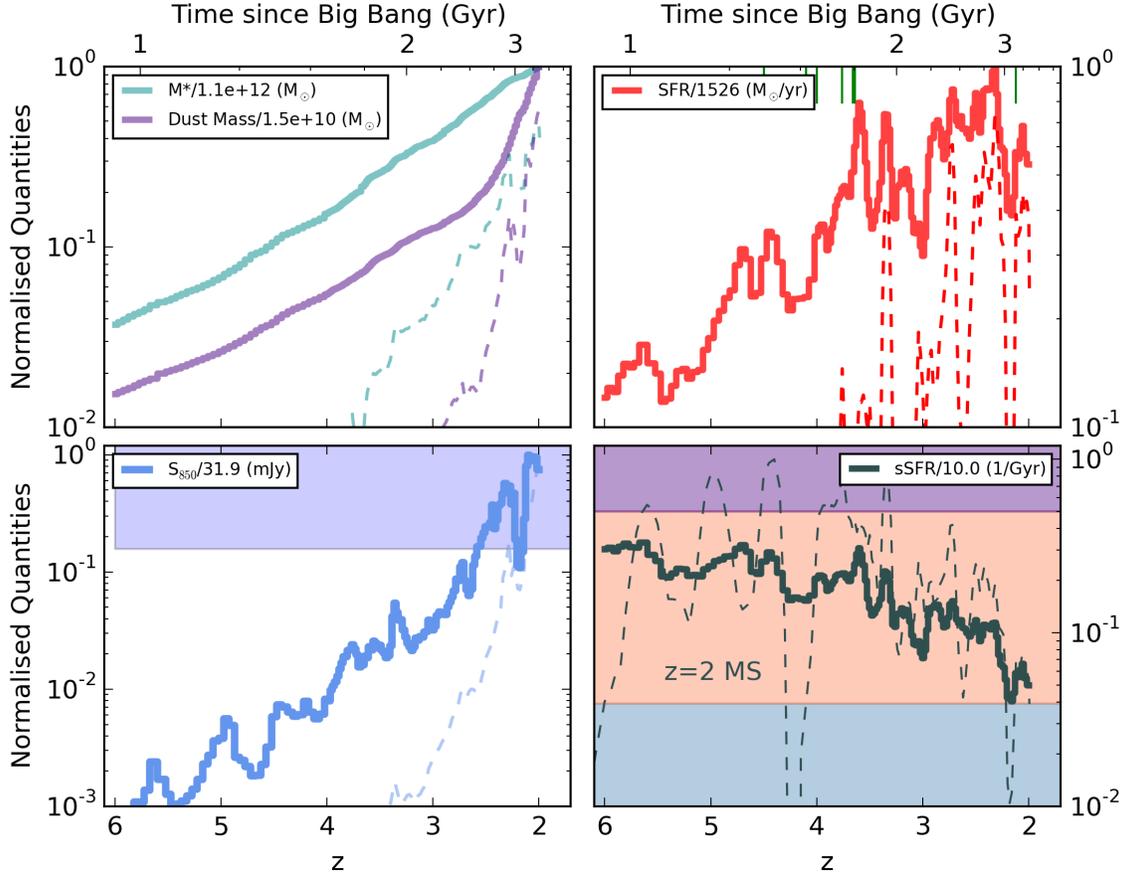


Figure 1: **Evolution of physical and observable properties of submillimetre emission region and central galaxy.** The 200 kpc submillimetre emission region are shown with thick solid lines, while the central galaxy's properties are given by thin dashed lines in each panel. Stellar and dust mass are in the top left, SFR at the top right; predicted observed $850 \mu\text{m}$ flux density at the bottom left; and specific SFR (M_*/SFR) at the bottom right. The SFR is averaged on 50 Myr timescales, and includes a correctional factor 0.7 for mass loss. Locations of major galaxy mergers ($> 1 : 3$) are noted by green vertical ticks on the top axis of the top right panel. The blue shaded region in the $850 \mu\text{m}$ curve shows when the galaxy would be detectable as an SMG with SCUBA ($S_{850} > 5 \text{ mJy}$). The yellow and purple shaded regions in the bottom right show the rough ranges for the $z = 2$ Main Sequence and Starburst regime. The grey region denotes below the Main Sequence.

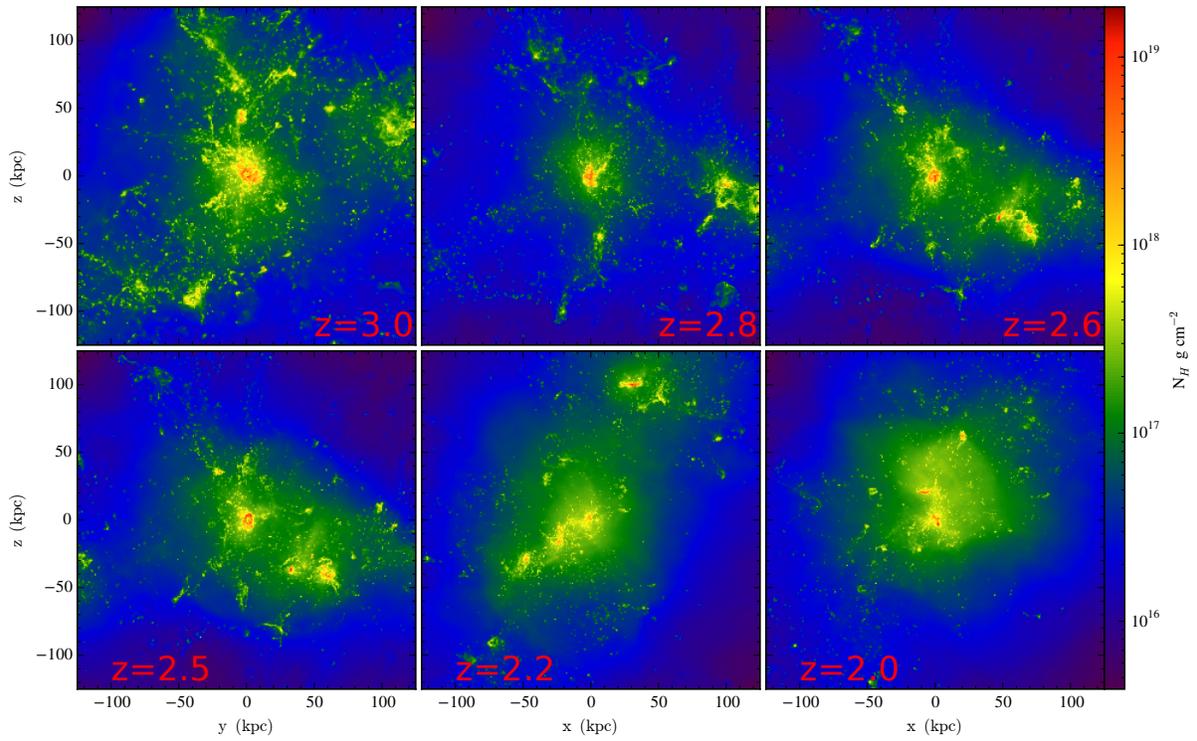


Figure 2: **Surface density projection maps of 250 kpc region around central submillimetre galaxy between redshifts $z \approx 2 - 3$.** The submillimetre-emission region probed in surveys typically encompasses a central galaxy in a massive halo that is undergoing a protracted bombardment phase by numerous sub-halos. Some of the brightest SMGs arise from numerous galaxies within the beam in a rich environment (bottom right panel).

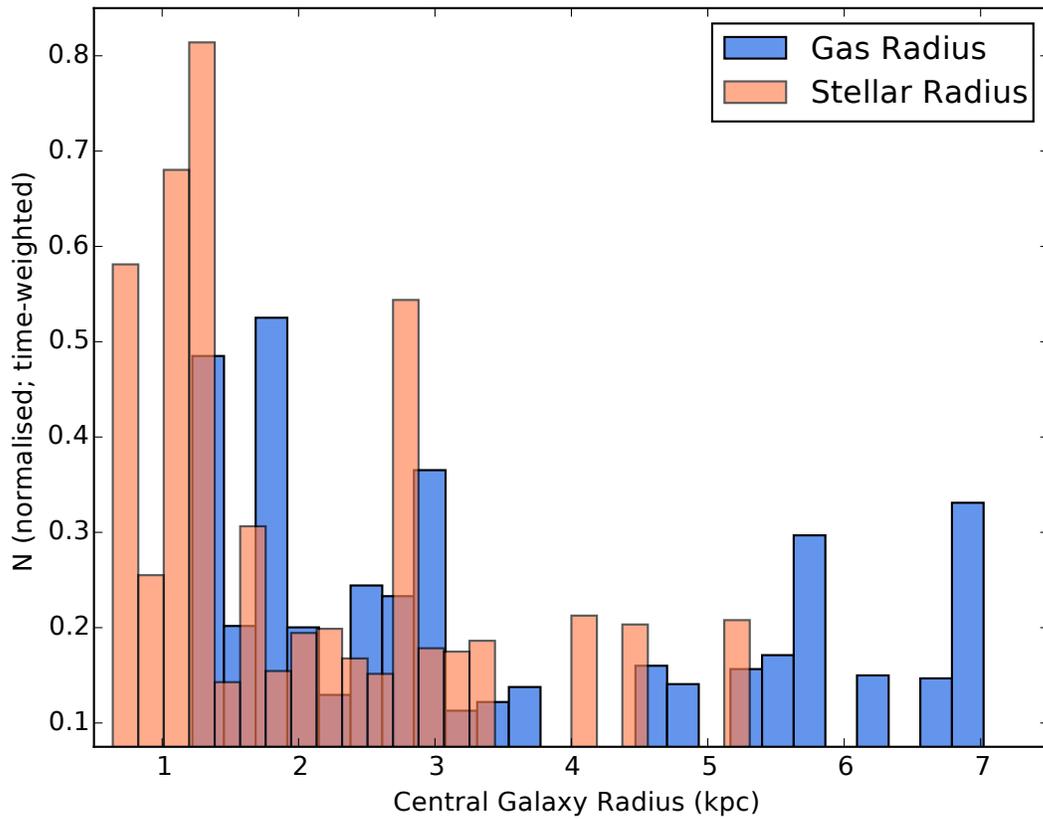


Figure 3: **Gas and stellar radius distribution for central submillimetre galaxy.** The orange histogram denotes the half mass radius of the stars, while the blue shows the gas. The galaxy gas is more distributed in the central than the (sub-kiloparsec) extent expected from major mergers, though still sufficiently compact that it will remain unresolved even with \sim arcsecond resolution.

METHODS

1 Cosmological Hydrodynamic Zoom Simulations

We utilise a newly-developed version of TreeSPH that employs a pressure-entropy formulation of smoothed particle hydrodynamics (SPH)³¹ that obviates many of the potential discrepancies noted between grid-based codes, traditional SPH codes, and moving-mesh algorithms^{32–34}. In particular, we employ the hydrodynamic code GIZMO¹² in P-SPH mode which conserves momentum, energy, angular momentum and entropy, and includes newly developed algorithms to treat the artificial viscosity, entropy diffusion and time-stepping^{31,35}. The gravity solver is a modified version of the GADGET-3 solver³⁶, and an updated softening kernel to better represent the potential of the SPH smoothing kernel is included³⁷.

The simulations are fully cosmological zoom-in calculations of the evolution of individual galaxies. A 144 Mpc³ cosmological volume was simulated at low resolution down to redshift $z = 0$ with dark matter only. The halo of interest was identified, and re-simulated at much higher resolution with baryons included. The initial conditions were generated with the MUSIC code³⁸. We simulate four zoom galaxies – one is our main galaxy, and the other three are at varying resolutions and masses for the purposes of testing. The main galaxy of interest to this study resides in a dark matter halo mass of $M_{\text{DM}} = 3 \times 10^{13} M_{\odot}$ at $z = 2$. The initial baryonic particle masses in the high-resolution region were $2.7 \times 10^5 M_{\odot}$, and the minimum baryonic/stars/dark matter force softening lengths were 9/21/142 proper pc at $z = 2$. The physical properties of all of the modeled galaxies are presented in ED Table 1 in the Extended Data.

The baryonic physics implemented into GIZMO are developed based on extensive tests studying idealised simulations of both isolated disks and galaxy mergers^{39–44}. The gas cools utilising an updated cooling curve to standard⁴⁵ implementations in SPH codes which includes both atomic and molecular line emission⁴⁶. The modeled interstellar medium is multiphase. The neutral ISM is broken into an atomic and molecular component following algorithms that scale the molecular fraction with column density and gas phase metallicity^{47,48}. Star formation occurs in molecular gas above a threshold density (here, this is set to $n_{\text{thresh}} = 10 \text{ cm}^{-3}$). Star formation is further restricted to gas that is locally self-gravitating, where:

$$\alpha \equiv \beta' \frac{|\nabla \cdot \mathbf{v}|^2 + |\nabla \times \mathbf{v}|^2}{G \rho} < 1 \quad (1)$$

This follows from studies⁴⁹ that show that the predicted spatial distribution of star formation in galaxies is more realistic when utilising a gas self-gravitating criterion compared to a variety of other algorithms (including a fixed density threshold, a pure molecular-gas law, a temperature threshold, a Jeans criterion, a cooling-time criterion and a converging flow criterion). The star formation rate follows a volumetric relation:

$$\dot{\rho}_* = \rho_{\text{mol}}/t_{\text{ff}} \quad (2)$$

That is, stars are allowed to form with 100% efficiency per free fall time. The star formation is subsequently self-regulated by stellar feedback, resulting in a time-averaged efficiency on galaxy scales of ϵ_{ff} of $\sim 0.005 - 0.1$ ³⁹.

Once stars have formed, they can impact the ISM via various feedback mechanisms. Assuming a Kroupa⁵⁰ stellar initial mass function, and utilising STARBURST99⁵¹ for luminosity, mass-return and supernova rate calculations as a function of stellar age and metallicity, we include the following forms of stellar feedback:

1. *Radiation Momentum Deposition*: At each timestep, the gas near young stars is impacted by a momentum flux given by

$$P_{\text{rad}} \approx (1 - \exp(-\tau_{\text{UV/optical}})) (1 + \tau_{\text{IR}}) L_{\text{incident}}/c \quad (3)$$

Where τ_{IR} is calculated directly from the simulation, as $\tau_{\text{IR}} = \Sigma_{\text{gas}} \kappa_{\text{IR}}$, and $\kappa_{\text{IR}} = 5(Z/Z_{\odot}) \text{ g}^{-1} \text{ cm}^2$.

2. *Supernovae and Stellar Winds*: We utilise tabulated Type-I and Type-II supernovae rates^{51,52}; if a supernova occurs during a timestep, thermal energy and radial momentum are injected within a smoothing length of the star. Gas and metal return is included as well. Stellar winds are similarly included with energy, wind momentum, mass and metals deposited within a smoothing length.
3. *Photoheating of HII regions*: The production rate of ionising radiation from stars determines the extent of HII regions (allowing for overlapping regions). These regions are heated to 10^4 K if the gas is below that threshold.

We utilise the second and third models in the Table to test the convergence properties of our simulations. One model is run with the same mass baryonic resolution as our main model (standard resolution; SR), and one a factor of ~ 8 higher resolution (high resolution; HR). In ED Fig 7, we show the modeled duty cycle above a given flux density as a function of flux density for these two models. We see that the shortest lived ($\lesssim 200 \text{ Myr}$) emission spikes present in the standard resolution model may not be converged in the highest resolution model. Notably, emission with longer duty cycles are either converged, or *underpredicted* in our standard resolution model, suggesting that the relatively long-lived submillimetre-luminous phase is robust.

We show the $M_* - z$ relation for the central in ED Fig 8 as compared to observational constraints²². The central galaxy has a stellar mass a factor ~ 2 greater than the observed median stellar mass for comparable mass halos at this epoch. The model galaxy may represent an outlier in the $M_* - z$ relation at these redshifts. Indeed, the thickness of the observational constraints shows the uncertainty, not range in possible values. Alternatively, it is possible that the inclusion of feedback from an active galactic nucleus (AGN) could impact the stellar mass buildup in the galaxy, though the level to which black hole growth can impact star formation near the submillimetre-luminous phase is unclear. Some models have shown that AGN can grow efficiently in the absence of major mergers^{53,54}, while other models and observations suggest that mergers may be necessary to grow massive holes^{55–57}. The last major merger before the submillimetre-luminous phase is $\sim 1 \text{ Gyr}$ prior. Tests with our low resolution model (m13m14) show that without AGN feedback, residual star formation drives a factor ~ 2 increase in stellar mass at late times ($z \approx 0 - 1$). Finally, we note that a higher mass resolution model could potentially also result in decreased final stellar masses. In our convergence tests, the final M_* (at $z = 2$) of the HR run is $\sim 60\%$ that of the SR run.

2 Dust Radiative Transfer Calculations

To calculate the inferred observational properties of our simulated galaxies, we have developed a dust radiative transfer package, POWDERDAY. In short, POWDERDAY takes hydrodynamic simulations of galaxies in evolution, projects the gas properties onto an adaptive mesh, and calculates the radiative transfer from the stellar sources through the dusty interstellar medium until an equilibrium dust temperature is achieved.

In detail, we identify galaxies utilising SKID to locate bound groups of baryonic particles^{58,59}, and track their progenitors back in time^{60,61}. Galaxies and halos are required to contain at least 64 particles each in order to be identified. We cut out a 200 kpc (side length) region around the galaxy of interest, and subdivide the domain into an adaptive grid with an octree memory structure. Formally, we begin with one cell encompassing the entire $8 \times 10^6 \text{ kpc}^3$ radiative transfer region. The cells then recursively subdivide into octs until there are a threshold maximum number of gas particles in

the cell (we employ $n_{\text{subdivide,thresh}} = 64$, though experiments with $n_{\text{subdivide,thresh}} = 32$ show converged results). The physical properties of the gas particles are projected onto the octree using a spline smoothing kernel⁶².

The spectral energy distribution of stars are calculated on the fly with the Flexible Stellar Population Synthesis code, FSPS^{63,64} through PYTHON-FSPS, a set of PYTHON hooks for FSPS (<https://github.com/dfm/python-fsps>). The SEDs are calculated as simple stellar populations with ages and metallicities determined by the hydrodynamic simulation, and assuming a Kroupa IMF.

The radiative transfer happens in a Monte Carlo fashion utilising the three-dimensional dust radiative transfer solver, HYPERION⁶⁵. The code uses an iterative methodology to determine the radiative equilibrium temperature⁶⁶, and we determine convergence when the energy absorbed by 99% of the cells has changed by less than 1% between iterations. We assume a dust grain-size distribution comparable to that of the Milky Way⁶⁷, with $R \equiv A_v/E(B-V) = 3.15$. The dust emissivities are updated to include an approximation for polycyclic aromatic hydrocarbons (PAHs) alongside thermal emission⁶⁸. We assume a constant dust to metals ratio of 0.4, motivated by both Milky Way and extragalactic observational constraints^{69–71}.

The underlying HYPERION code has passed the standard benchmarks for codes of this type⁷², and we have found that POWDERDAY compares well against other publicly available dust radiative transfer codes^{73–75} in test starburst SPH galaxy merger simulations.

3 Parameter Choices

In ED Fig 9, we present a number of tests of our parameter choices for the radiative transfer calculations. We show the predicted $850\mu\text{m}$ light curve from our lowest resolution model (m13m14) utilising both fiducial parameters, as well as three parameter choice variations.

We first ask whether our chosen radiative transfer grid size affects our principle results. Our fiducial model is a 200 kpc (on a side) box cut out of the global cosmological simulation centred on the halo of interest. This size was chosen to reflect a rough average of the typical (sub)millimetre beam sizes typically used to detect SMGs. For example, assuming Planck 2013 cosmological parameters⁷⁶, the Submillimetre Common-Use Bolometer Array (SCUBA) on the James Clerk Maxwell Telescope (JCMT) has a $15''$ full width at half maximum (FWHM) beam at $850\mu\text{m}$. At $z = 2$ this corresponds to ~ 128 kpc. At the same redshift, the beam of AzTEC and LABOCA at 1 mm on the JCMT corresponds to ~ 163 kpc ($19''$); The South Pole Telescope (SPT) has a beamsize of 540 kpc at 1.4 mm ($63''$); and Herschel’s SPIRE instrument ranges from 154 – 308 kpc ($250 - 500\mu\text{m}$; $18 - 36''$).

Because a few notable beam sizes (of particular relevance, the SCUBA beam) are smaller than our assumed box size of 200 kpc, we have run an additional model with box length 100 kpc (all other parameters exactly the same). We highlight the resultant $850\mu\text{m}$ light curve from this model in the top right panel of ED Fig 9. When comparing to our fiducial model, it is apparent that our results are robust to the highest resolution beams that have been used for SMG surveys at single dish facilities to date.

We additionally investigate whether our inclusion of PAHs in our model makes any difference to the calculated submillimetre-wave flux density of our model galaxy. This is presented in the bottom left panel of ED Fig 9. Again, we note minimal impact on the submillimetre SED of our model.

Finally, we ensure that our results are converged with the number of photons emitted. We fiducially run 10^7 photons per grid (roughly 100 per cell). In the bottom-right, we show the results from a run with 10^8 photons per grid, and show that the results are robust against this parameter choice.

4 Relation to other Models

Historically, the methods used, and physical models for SMG formation in numerical simulations are quite varied. Here, we summarise these methods and results, and place our own model into this context. Broadly, there are three classes of SMG formation models: cosmological semi-analytic models (SAMs), idealised non-cosmological simulations, and cosmological hydrodynamic models. The present model falls into the latter category. Our presented model is the first self-consistent cosmological simulation with baryons and bona fide radiative transfer to form a submillimetre galaxy with physical properties comparable to those observed.

The initial forays into this field were typically with SAMs. This owes to the fact that SAMs are computationally inexpensive, and allow for a large search in physical parameter space relatively easily. SAMs either utilise analytic halo merger trees, or directly simulate them, and then employ analytic prescriptions to describe the central galaxies. The Durham SAM^{7,77} couples galaxies formed in a semi-analytic model with dust radiative transfer. These simulations model galaxies that have axisymmetric geometries that consist of a disc and a bulge. Young stellar populations are assumed to still be enveloped in their birth clouds, and thus experience additional attenuation. This model suggests that SMGs typically owe to $\sim 22\%$ major mergers, the remainder as minor mergers, and that the stellar initial mass function is flat during the starburst. The typical duty cycle for the submillimetre-luminous phase is ~ 100 Myr (a factor ~ 7.5 lower than found in our work), galaxies extremely gas rich ($f_{\text{gas}} \sim 75\%$), and stellar masses a factor ~ 10 lower than our model ($M_* \sim 2 \times 10^{10} M_\odot$). While the stellar masses of SMGs are debated^{21,78,79}, the gas fractions appear to be uniformly lower in observations^{4,18,23,80,81}, and a flat stellar initial mass function likely ruled out by CO dynamical mass measurements⁸².

As an alternative to cosmological simulations, a number of works have explored SMG formation in idealised simulations^{8,83–86}. These works evolve hydrodynamic models of idealised discs and mergers over a range of merger mass ratios, and combine these with dust radiative transfer simulations⁷⁴. These models infer halo masses and stellar masses for SMGs comparable to those modeled here. This said, in the idealised galaxy models, $\sim 30 - 70\%$ of the SMGs (flux dependent) owe to merger-driven starbursts, substantially higher than what is found for our model. Some works⁸ have noted that binary mergers that cause SMGs may break up into multiples at high-resolution owing to the contribution to the total flux of individual inspiralling discs. Because idealised simulations are non-cosmological in nature, comparing the multiplicity inferred from these to our models is difficult: the major merger multiplicity can only be 2 when considering galaxies at the same redshift. On the other hand ED Fig 2 suggests potentially larger multiplicity can be observed for physically associated clumps.

To fully capture the cosmic environment of SMGs in formation, as well as their baryonic structure and morphology, cosmological hydrodynamic simulations are likely the best tool. Thusfar, cosmological hydrodynamic simulations used to simulate SMGs have not employed direct radiative transfer models¹⁰. As such, inferring when a galaxy is an SMG in cosmological simulations has necessitated the use of parameterised emission models, such as assumed grey-body emission laws⁹, or star formation rate thresholds¹⁰. The physical properties for SMGs derived from the most extensive of these works¹⁰ (i.e. M_* , M_{DM} , and f_{gas}) are similar to the model presented here, though with roughly a factor ~ 3 difference in SFR.

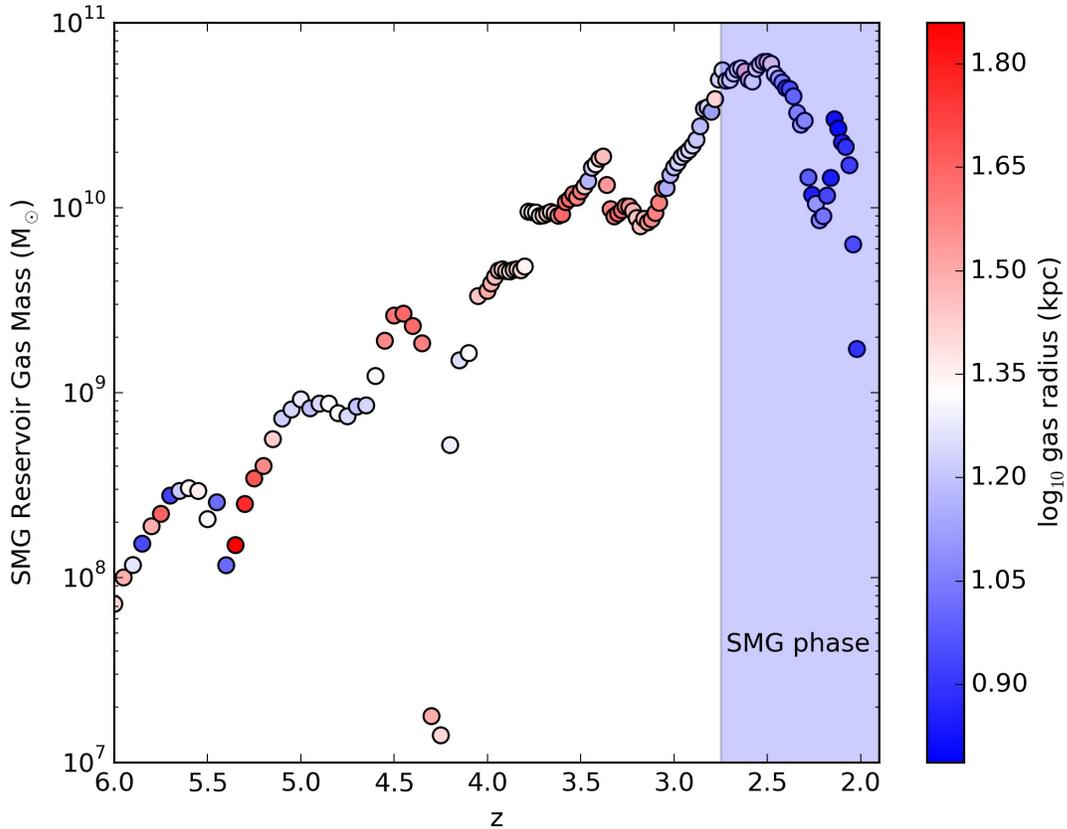
5 Code Availability

We have made POWDERDAY available at <https://bitbucket.org/desika/powderday>, and GIZMO available at <https://bitbucket.org/phopkins/gizmo>

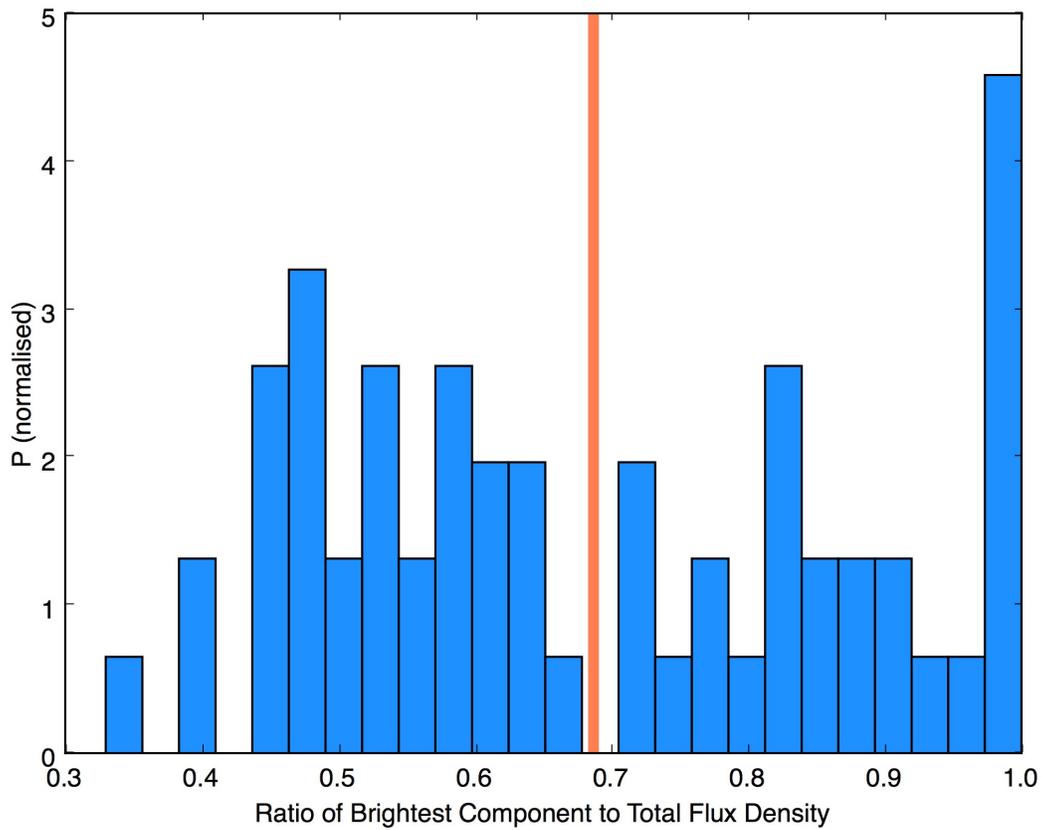
31. Hopkins, P. F. A general class of Lagrangian smoothed particle hydrodynamics methods and implications for fluid mixing problems. *Mon. Not. R. Astron. Soc.* **428**, 2840–2856 (2013).
32. Agertz, O. *et al.* Fundamental differences between SPH and grid methods. *Mon. Not. R. Astron. Soc.* **380**, 963–978 (2007).
33. Sijacki, D., Vogelsberger, M., Kereš, D., Springel, V. & Hernquist, L. Moving mesh cosmology: the hydrodynamics of galaxy formation. *Mon. Not. R. Astron. Soc.* **424**, 2999–3027 (2012).
34. Hayward, C. C., Torrey, P., Springel, V., Hernquist, L. & Vogelsberger, M. Galaxy mergers on a moving mesh: a comparison with smoothed-particle hydrodynamics. *arXiv/1309.2942* (2013).
35. Faucher-Giguere, C.-A. *et al.* Neutral hydrogen in galaxy halos at the peak of the cosmic star formation history. *arXiv/1409.1919* (2014).
36. Springel, V., Di Matteo, T. & Hernquist, L. Modelling feedback from stars and black holes in galaxy mergers. *Mon. Not. R. Astron. Soc.* **361**, 776–794 (2005).
37. Barnes, J. E. Gravitational softening as a smoothing operation. *Mon. Not. R. Astron. Soc.* **425**, 1104–1120 (2012).
38. Hahn, O. & Abel, T. Multi-scale initial conditions for cosmological simulations. *Mon. Not. R. Astron. Soc.* **415**, 2101–2121 (2011).
39. Hopkins, P. F., Quataert, E. & Murray, N. Self-regulated star formation in galaxies via momentum input from massive stars. *Mon. Not. R. Astron. Soc.* **417**, 950–973 (2011).
40. Hopkins, P. F., Quataert, E. & Murray, N. The structure of the interstellar medium of star-forming galaxies. *Mon. Not. R. Astron. Soc.* **421**, 3488–3521 (2012).
41. Hopkins, P. F., Narayanan, D., Murray, N. & Quataert, E. Dense molecular gas: a sensitive probe of stellar feedback models. *Mon. Not. R. Astron. Soc.* **433**, 69–77 (2013).
42. Hopkins, P. F. *et al.* Star formation in galaxy mergers with realistic models of stellar feedback and the interstellar medium. *Mon. Not. R. Astron. Soc.* **430**, 1901–1927 (2013).
43. Hopkins, P. F. *et al.* Resolving the generation of starburst winds in Galaxy mergers. *Mon. Not. R. Astron. Soc.* **433**, 78–97 (2013).
44. Narayanan, D. & Hopkins, P. F. Why is the Milky Way X-factor constant? *Mon. Not. R. Astron. Soc.* (2013).
45. Katz, N., Weinberg, D. H. & Hernquist, L. Cosmological Simulations with TreeSPH. *Astrophys. J. Supp.* **105**, 19–+ (1996).
46. Ferland, G. J. *et al.* The 2013 Release of Cloudy. *RMXAA* **49**, 137–163 (2013).
47. Krumholz, M. R., McKee, C. F. & Tumlinson, J. The Atomic-to-Molecular Transition in Galaxies. I. An Analytic Approximation for Photodissociation Fronts in Finite Clouds. *Astrophys. J.* **689**, 865–882 (2008).
48. Krumholz, M. R. & Gnedin, N. Y. A Comparison of Methods for Determining the Molecular Content of Model Galaxies. *Astrophys. J.* **729**, 36–+ (2011).
49. Hopkins, P. F., Narayanan, D. & Murray, N. The meaning and consequences of star formation criteria in galaxy models with resolved stellar feedback. *Mon. Not. R. Astron. Soc.* **432**, 2647–2653 (2013).
50. Kroupa, P. The Initial Mass Function of Stars: Evidence for Uniformity in Variable Systems. *Science* **295**, 82–91 (2002).
51. Leitherer, C. *et al.* Starburst99: Synthesis Models for Galaxies with Active Star Formation. *Astrophys. J. Supp.* **123**, 3–40 (1999).
52. Mannucci, F., Della Valle, M. & Panagia, N. Two populations of progenitors for Type Ia supernovae? *Mon. Not. R. Astron. Soc.* **370**, 773–783 (2006).
53. Bellovary, J. *et al.* The Relative Role of Galaxy Mergers and Cosmic Flows in Feeding Black Holes. *Astrophys. J.* **779**, 136 (2013).
54. Anglés-Alcázar, D. *et al.* Torque-limited Growth of Massive Black Holes in Galaxies across Cosmic Time. *Astrophys. J.* **800**, 127 (2015).
55. Hopkins, P. F., Kocevski, D. D. & Bundy, K. Do we expect most AGN to live in discs? *Mon. Not. R. Astron. Soc.* **445**, 823–834 (2014).
56. Kocevski, D. D. *et al.* CANDELS: Constraining the AGN-Merger Connection with Host Morphologies at $z \sim 2$. *Astrophys. J.* **744**, 148 (2012).
57. Treister, E., Schawinski, K., Urry, C. M. & Simmons, B. D. Major Galaxy Mergers Only Trigger the Most Luminous Active Galactic Nuclei. *Astrophys. J. Lett.* **758**, L39 (2012).
58. Governato, F. *et al.* The Local Group as a test of cosmological models. *New Astron. Rev.* **2**, 91–106 (1997).
59. Stadel, J. G. *Cosmological N-body simulations and their analysis*. Ph.D. thesis, UNIVERSITY OF WASHINGTON (2001).
60. Thompson, R. pyGadgetReader: GADGET snapshot reader for python. Astrophysics Source Code Library (2014). 1411.001.
61. Thompson, R. SPHGR: Smoothed-Particle Hydrodynamics Galaxy Reduction. Astrophysics Source Code Library (2015). 1502.012.
62. Turk, M. J. *et al.* yt: A Multi-code Analysis Toolkit for Astrophysical Simulation Data. *Astrophys. J. Supp.* **192**, 9 (2011).
63. Conroy, C., Gunn, J. E. & White, M. The Propagation of Uncertainties in Stellar Population Synthesis Modeling. I. The Relevance of Uncertain Aspects of Stellar Evolution and the Initial Mass Function to the Derived Physical Properties of Galaxies. *Astrophys. J.* **699**, 486–506 (2009).
64. Conroy, C. & Gunn, J. E. The Propagation of Uncertainties in Stellar Population Synthesis Modeling. III. Model Calibration, Comparison, and Evaluation. *Astrophys. J.* **712**, 833–857 (2010).
65. Robitaille, T. P. HYPERION: an open-source parallelized three-dimensional dust continuum radiative transfer code. *Astron. Astrophys.* **536**, A79 (2011).
66. Lucy, L. B. Computing radiative equilibria with Monte Carlo techniques. *Astron. Astrophys.* **344**, 282–288 (1999).
67. Weingartner, J. C. & Draine, B. T. Dust Grain-Size Distributions and Extinction in the Milky Way, Large Magellanic Cloud, and Small Magellanic Cloud. *Astrophys. J.* **548**, 296–309 (2001).
68. Robitaille, T. P. *et al.* A self-consistent model of Galactic stellar and dust infrared emission and the abundance of polycyclic aromatic hydrocarbons. *Astron. Astrophys.* **545**, A39 (2012).
69. Dwek, E. The Evolution of the Elemental Abundances in the Gas and Dust Phases of the Galaxy. *Astrophys. J.* **501**, 643–+ (1998).
70. Vladilo, G. Dust and Elemental Abundances in Damped Ly alpha Absorbers. *Astrophys. J.* **493**, 583–+ (1998).
71. Watson, D. The Galactic dust-to-metals ratio and metallicity using gamma-ray bursts. *Astron. Astrophys.* **533**, A16 (2011).
72. Pascucci, I. *et al.* The 2D continuum radiative transfer problem. Benchmark results for disk configurations. *Astron. Astrophys.* **417**, 793–805 (2004).
73. Jonsson, P. SUNRISE: polychromatic dust radiative transfer in arbitrary geometries. *Mon. Not. R. Astron. Soc.* **372**, 2–20 (2006).
74. Jonsson, P., Groves, B. A. & Cox, T. J. High-resolution panchromatic spectral models of galaxies including photoionization and dust. *Mon. Not. R. Astron. Soc.* 186–+ (2010).
75. Torrey, P. *et al.* Synthetic galaxy images and spectra from the Illustris simulation. *Mon. Not. R. Astron. Soc.* **447**, 2753–2771 (2015).
76. Planck Collaboration *et al.* Planck 2013 results. XVI. Cosmological parameters. *Astron. Astrophys.* **571**, A16 (2014).
77. González, J. E., Lacey, C. G., Baugh, C. M. & Frenk, C. S. The role of submillimetre galaxies in hierarchical galaxy formation. *Mon. Not. R. Astron. Soc.* **413**, 749–762 (2011).
78. Hainline, L. J. *et al.* The Stellar Mass Content of Submillimetre-selected Galaxies. *Astrophys. J.* **740**, 96 (2011).
79. Michałowski, M. J. *et al.* Determining the stellar masses of submillimetre galaxies: the critical importance of star formation histories. *Astron. Astrophys.* **571**, A75 (2014).
80. Bothwell, M. S. *et al.* A survey of molecular gas in luminous submillimetre galaxies. *Mon. Not. R. Astron. Soc.* **429**, 3047–3067 (2013).
81. Narayanan, D., Bothwell, M. & Davé, R. Galaxy gas fractions at high redshift: the tension between observations and cosmological simulations. *Mon. Not. R. Astron. Soc.* **426**, 1178–1184 (2012).
82. Tacconi, L. J. *et al.* Submillimetre Galaxies at $z \sim 2$: Evidence for Major Mergers and Constraints on Lifetimes, IMF, and CO-H₂ Conversion Factor. *Astrophys. J.* **680**, 246–262 (2008).
83. Chakrabarti, S., Fenner, Y., Cox, T. J., Hernquist, L. & Whitney, B. A. An Evolutionary Model For Submillimetre Galaxies. *Astrophys. J.* **688**, 972–989 (2008).
84. Narayanan, D., Cox, T. J., Hayward, C. C., Younger, J. D. & Hernquist, L. The star-forming molecular gas in high-redshift Submillimetre Galaxies. *Mon. Not. R. Astron. Soc.* **400**, 1919–1935 (2009).
85. Narayanan, D. *et al.* The formation of high-redshift submillimetre galaxies. *Mon. Not. R. Astron. Soc.* **401**, 1613–1619 (2010).
86. Hayward, C. C. *et al.* What Does a Submillimetre Galaxy Selection Actually Select? The Dependence of Submillimetre Flux Density on Star Formation Rate and Dust Mass. *Astrophys. J.* **743**, 159 (2011).
87. Narayanan, D., Krumholz, M. R., Ostriker, E. C. & Hernquist, L. A general model for the CO-H₂ conversion factor in galaxies with applications to the star formation law. *Mon. Not. R. Astron. Soc.* **421**, 3127–3146 (2012).

Table Extended Data 1: **Summary of Model Galaxies.** M_* and M_{halo} refer to the stellar and halo mass at $z = 2$; ϵ_b and ϵ_{DM} refer to the minimum force softening lengths for baryons and dark matter particles.

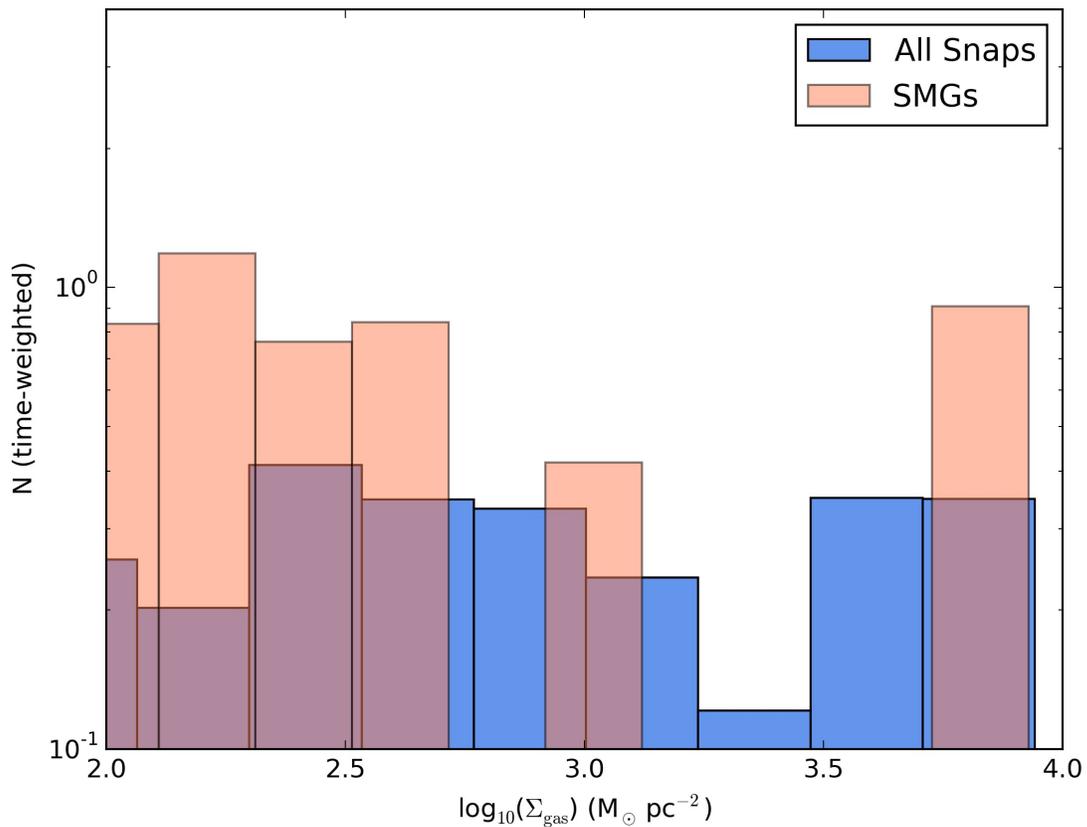
Model Name	Model Purpose	M_* ($z=2$) M_\odot	M_{halo} ($z=2$) M_\odot	m_b M_\odot	m_{DM} M_\odot	ϵ_b pc	ϵ_{DM} pc	Final Redshift
B100	Main Model	4×10^{11}	3×10^{13}	2.7×10^5	1.3×10^6	9	142	2
TL37 SR	Resolution Test	8×10^{10}	7×10^{12}	2.7×10^5	1.3×10^6	9	142	2
TL37 HR	Resolution Test	5×10^{10}	7×10^{12}	3.3×10^4	1.7×10^5	9	142	2
m13m14	RT Parameter Survey	3×10^{11}	7×10^{12}	4.4×10^6	2.3×10^7	70	700	0.2



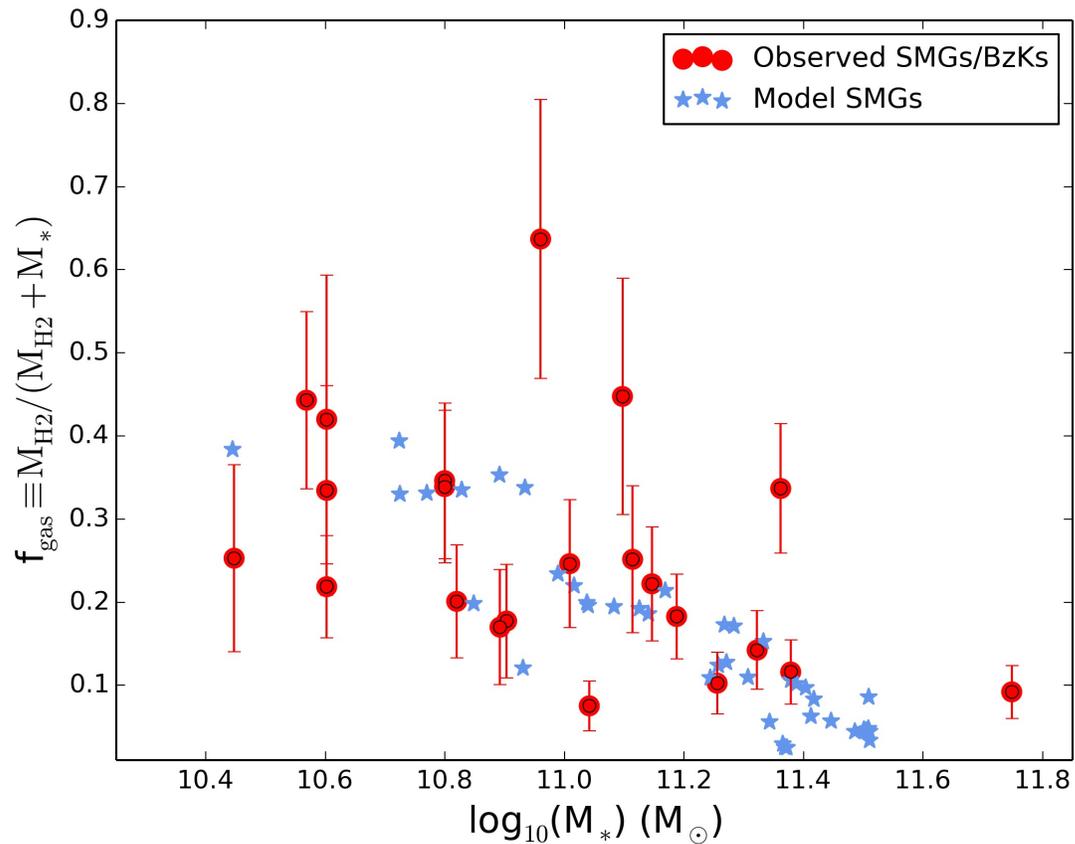
Extended Data Figure 1: Mass of gas in central galaxy that will be consumed during SMG starburst as a function of z . The gas mass consumed during the starburst is calculated by tracking the evolution of gas particles that turn into stars during the SMG phase ($z \approx 2 - 2.7$), and is measured for the central galaxy itself (i.e. gas ejected into the halo is not included). The colours denote the median scale height from the galaxy centre of mass. The SMG gas reservoir follows a cycle of being pushed outward followed by re-accretion.



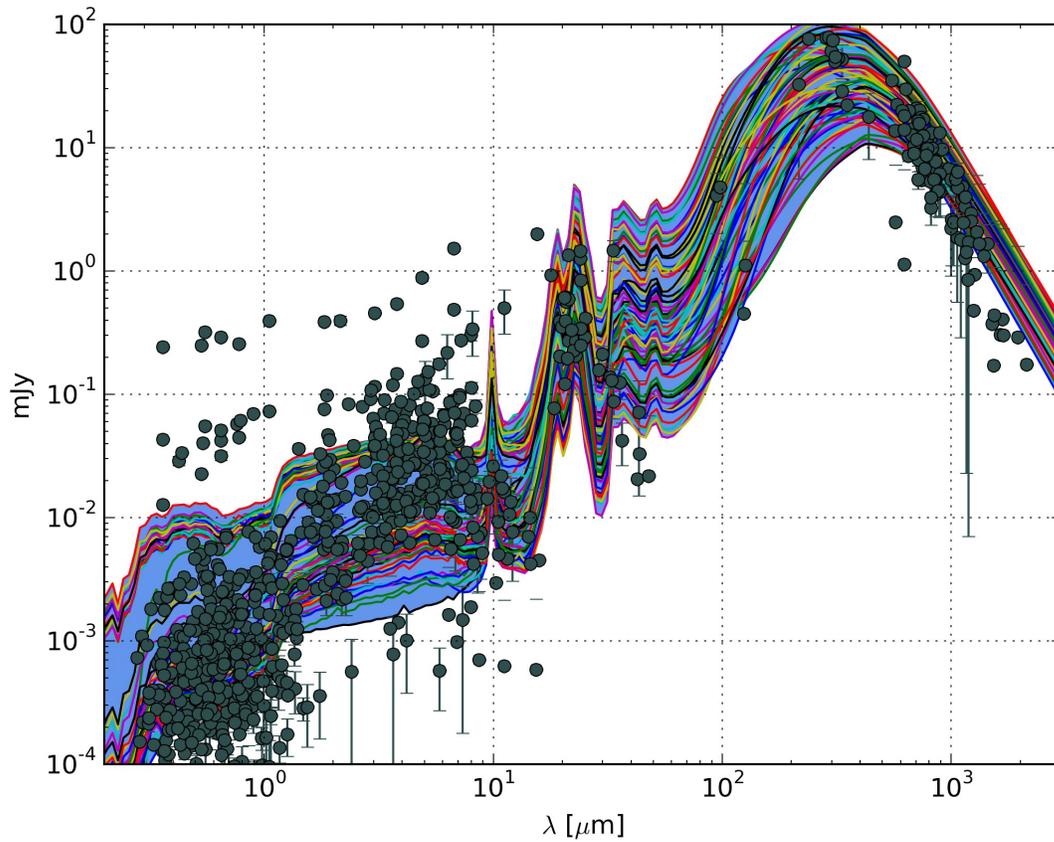
Extended Data Figure 2: Predicted contribution of submillimetre-luminous region components to total flux density. Submillimetre-luminous regions often break up into multiples. Shown is a histogram of the ratio of the brightest component to the total flux density from the region, with the average denoted with the vertical line. The region is generally dominated by one component, though smaller subhalos can contribute on average $\sim 30\%$ of the observed flux density.



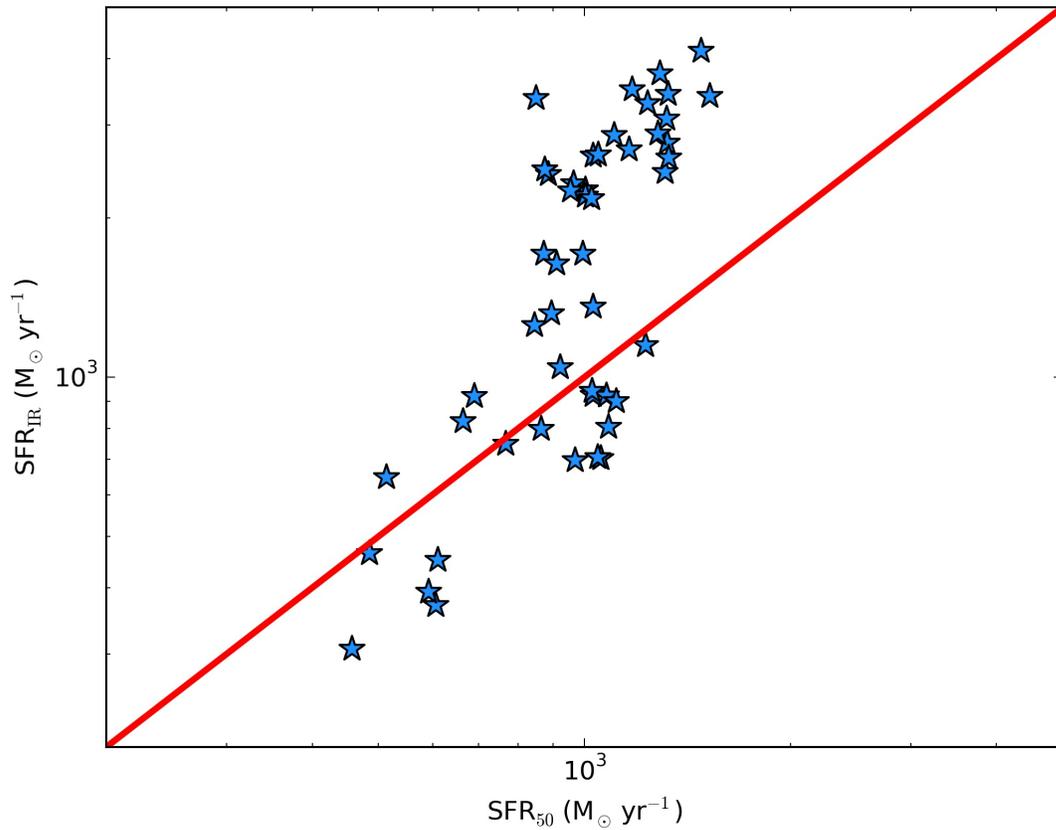
Extended Data Figure 3: Gas surface density for central submillimetre galaxy. The blue histograms show the time-weighted distribution of surface densities during all phases, whilst the orange show the same for the submillimetre-luminous phase. We predict that the submillimetre-luminous phases do not have dramatically different surface density distributions compared to the non-submillimetre-luminous phases. This feature may be tentatively observed^{1,87}.



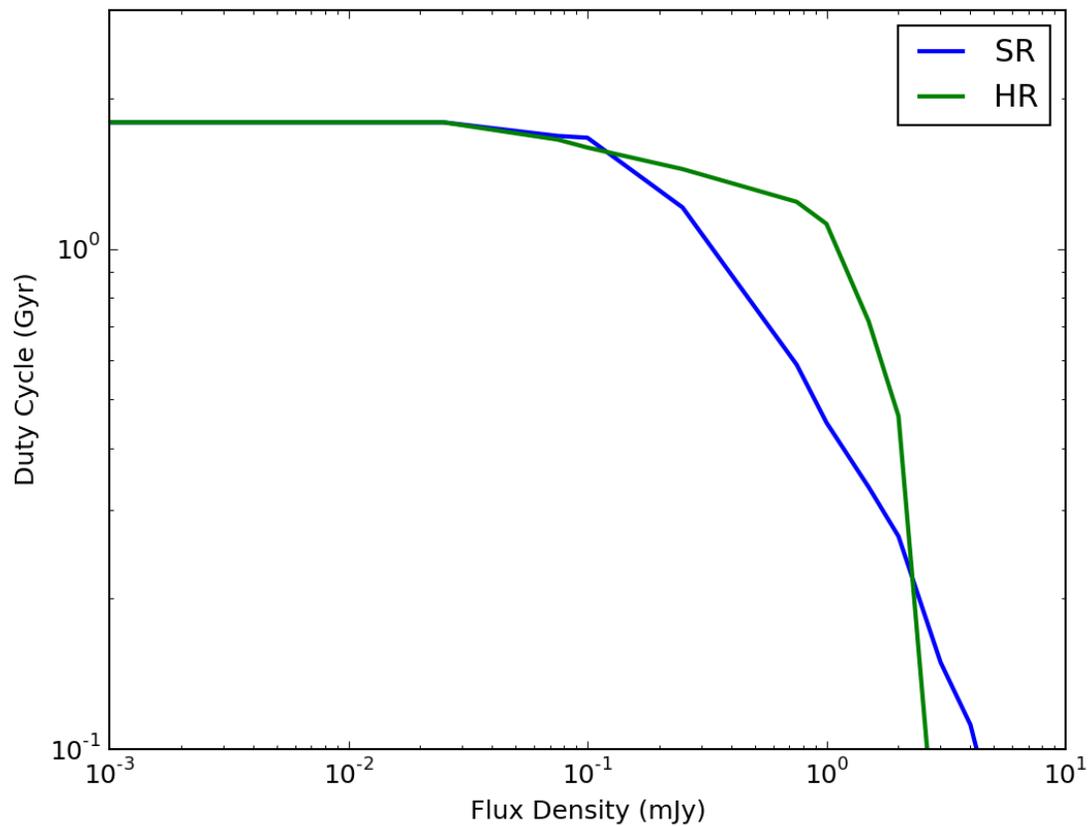
Extended Data Figure 4: Molecular gas fraction as a function of galaxy stellar mass. Blue stars show individual snapshots of central submillimetre galaxy whilst red circles with error bars (1σ) show observations with direct CO ($J=1-0$) measurements (to avoid complications in converting from higher-lying CO rotational lines to the ground state for a mass conversion). Both observations and our model show a declining molecular gas fraction with increasing galaxy mass, with a typical range of $f_{\text{gas}} = 0.1 - 0.4$ for galaxies of SMG mass.



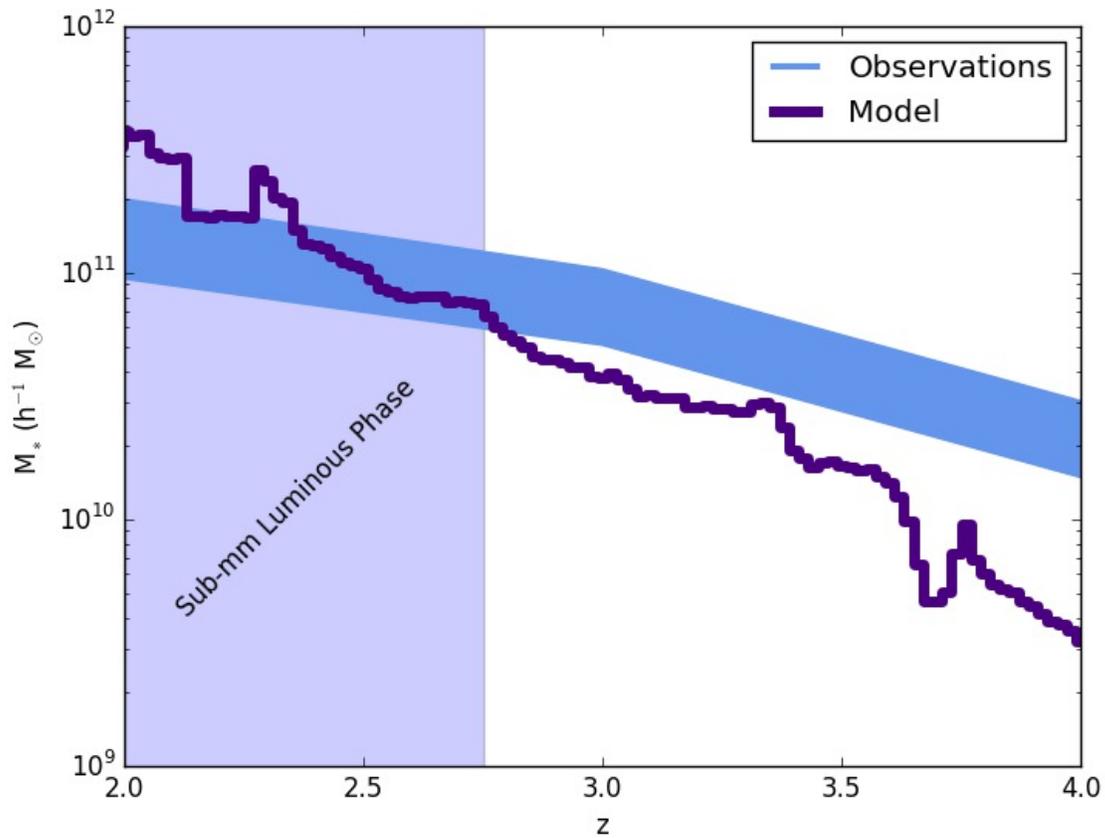
Extended Data Figure 5: Predicted spectral energy distribution (SED) for central submillimetre galaxy. The blue shaded region shows the range of SEDs for all snapshots that satisfy the fiducial $F_{850\mu\text{m}} > 5$ mJy submillimetre galaxy selection criteria, whilst the grey points with error bars (1σ) are a compilation of observed data. The individual coloured lines show the SEDs for individual submillimetre luminous snapshots. The data and models are redshifted to a common redshift $z = 2$. The model and data compare well, and the model suggests a diverse range of SMG SEDs.



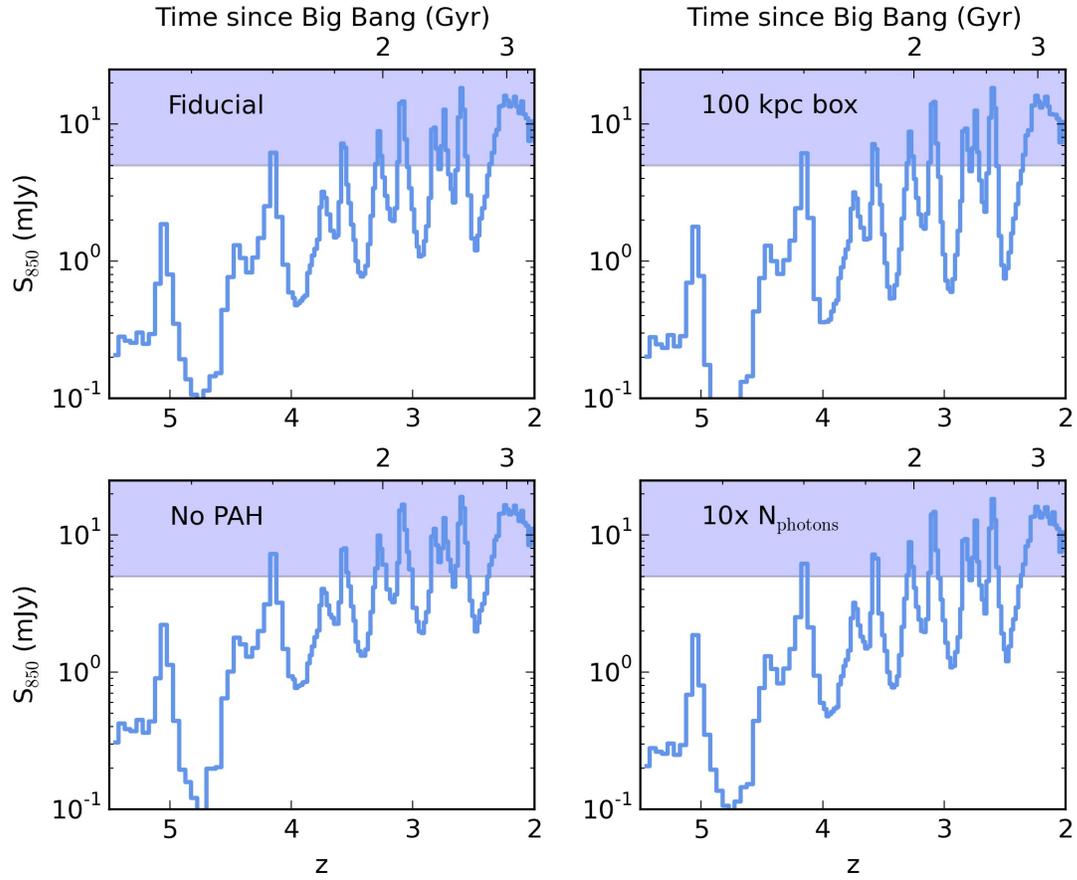
Extended Data Figure 6: Overestimate of SFR of High- z SMGs. The ordinate denotes the SFR as determined from the infrared SED²⁵, while the abscissa shows the SFR averaged over the last 50Myr in the simulations. Up to an SFR of $\sim 800 M_{\odot} \text{ yr}^{-1}$ the two correspond well. At higher SFRs, however, there is a dramatic departure owing to substantial contribution to the infrared luminosity by older stars.



Extended Data Figure 7: Resolution tests for hydrodynamic zoom simulations. Lines show $850\mu\text{m}$ duty cycle above a given flux density as a function of flux density for our resolution test models presented in the Methods section. SR denotes our standard resolution (the resolution of our main model) whilst HR is one level higher refinement.



Extended Data Figure 8: Stellar Mass - Redshift Relation for Model Galaxy. Purple line shows model whilst blue filled region shows observational constraints from an abundance matching assumption²². The model and observations are in reasonable agreement, especially during the submillimetre-luminous phase (purple shaded region). At late times, the stellar mass of the galaxy is a factor ~ 2 higher than the median observed galaxy.



Extended Data Figure 9: Tests of parameter choices for radiative transfer calculations. The simulated galaxy for these tests is our lowest resolution cosmological simulation (m13m14). Each panel shows the 850 μm flux density lightcurve of the tested model, with time noted on the abscissa (redshift on the bottom, time since the Big Bang on top). In all panels, the purple shaded region denotes $S_{850} \geq 5$ mJy, the canonical selection criteria for SMGs. *Top Left:* Our fiducial set of parameters; *Top Right:* Simulation with a 100 kpc (on a side) emission region instead of 200 kpc; *Bottom Left:* Simulation with our model for PAHs turned off; *Bottom Right:* Fiducial simulation run with ten times the number of photons.