The ROMULUS cosmological simulations: a physical approach to the formation, dynamics and accretion models of SMBHs

M. Tremmel,[1]* M. Karcher,[2] F. Governato,[1] M. Volonteri,[3] T. R. Quinn,[1] A. Pontzen,[4] L. Anderson[1] and J. Bellovary[5]

[1]*Astronomy Department, University of Washington, Box 351580, Seattle, WA, 98195, USA*
[2]*Statistics Department, University of Washington Seattle, WA, 98195, USA*
[3]*Sorbonne Universitès, UPMC Univ Paris 6 et CNRS, UMR 7095, Institut d'Astrophysique de Paris, 98 bis bd Arago, 75014 Paris, France*
[4] *Department of Physics and Astronomy, University College London, 132 Hampstead Road, London NW1 2PS, United Kingdom*
[5] *Department of Physics, Queensborough Community College, 222-05, 56th Avenue, Bayside, NY 11364, USA*

ABSTRACT

We present a novel implementation of supermassive black hole (SMBH) formation, dynamics and accretion in the massively parallel tree+SPH code, CHANGA. This approach improves the modelling of SMBHs in fully cosmological simulations, allowing for a more detailed analysis of SMBH-galaxy co-evolution throughout cosmic time. Our scheme includes novel, physically motivated models for SMBH formation, dynamics and sinking timescales within galaxies and SMBH accretion of rotationally supported gas. The sub-grid parameters that regulate star formation (SF) and feedback from SMBHs and SNe are optimized against a comprehensive set of $z = 0$ galaxy scaling relations using a novel, multidimensional parameter search. We have incorporated our new SMBH implementation and parameter optimization into a new set of high-resolution, large-scale cosmological simulations called ROMULUS. We present initial results from our flagship simulation, ROMULUS25, showing that our SMBH model results in SF efficiency, SMBH masses and global SF and SMBH accretion histories at high redshift that are consistent with observations. We discuss the importance of SMBH physics in shaping the evolution of massive galaxies and show how SMBH feedback is much more effective at regulating SF compared to SNe feedback in this regime. Further, we show how each aspect of our SMBH model impacts this evolution compared to more common approaches. Finally, we present a science application of this scheme studying the properties and time evolution of an example dual active galactic nucleus system, highlighting how our approach allows simulations to better study galaxy interactions and SMBH mergers in the context of galaxy-BH co-evolution.

Key words: methods: numerical – galaxies: general – quasars: supermassive black holes.

1 INTRODUCTION

Supermassive black holes (SMBHs) are ubiquitous in galaxies across a wide range of masses. SMBHs are observed not only in massive galaxies (e.g. Gehren et al. 1984; Kormendy & Richstone 1995; Kormendy & Ho 2013) but also in small, bulge-less disc galaxies (Filippenko & Ho 2003; Shields et al. 2008) as well as dwarfs (Reines et al. 2011; Reines & Deller 2012; Reines, Greene & Geha 2013; Moran et al. 2014). Accreting SMBHs lead to extremely energetic events throughout cosmic time, including luminous $z > 6$ quasars powered by SMBHs with masses as high as 10^9 M$_\odot$ (Fan et al. 2001; Mortlock et al. 2011).

Despite their importance to galaxy evolution theory, understanding how these black holes form, the mechanisms that regulate their growth, and in what ways they affect their host galaxies are still open areas of study. Empirical scaling relations between SMBH mass and the stellar mass and velocity dispersion of their host galaxies are indicative of co-eval growth (Häring & Rix 2004; Gültekin & et al. 2009; Volonteri & Bellovary 2012; Kormendy & Ho 2013; Schramm & Silverman 2013). While there have been attempts to quantify the evolution of this relationship (e.g. Alexander et al. 2008; Bennert et al. 2011; Bongiorno et al. 2014; Sun et al. 2015), these high redshift observations can be highly biased (Lauer et al. 2007) and cannot effectively probe lower mass galaxies and black holes. There is also evidence that the relationships break down at low redshift for lower mass, star-forming galaxies (Reines & Volonteri 2015). Thus, understanding the genesis of these empirical relations and

** E-mail: m.tremmel6@gmail.com

their mass dependency requires predictions from simulations that accurately follow SMBH growth in low-mass ($M_{vir} < 10^{10}$ M$_\odot$) haloes at both high and low redshifts.

Previous works have been fundamental in showing how energy from SMBH feedback is necessary to shape the bright end of the galaxy luminosity function, quench the formation of bulges in field galaxies and support a close causal connection between early rapid growth, galaxy mergers and QSO and active galactic nucleus (AGN) activity (Di Matteo, Springel & Hernquist 2005; Teyssier et al. 2011; Schaye et al. 2015; Sijacki et al. 2015; Bonoli et al. 2016; Volonteri et al. 2016a). With spatial resolutions of the order of 10–100s of pc, physical processes involved in SMBH accretion, feedback, and dynamics are necessarily implemented in cosmological simulations via sub-grid prescriptions, under the broad assumption that conditions at the smallest resolved scale drive the SMBH evolution at much smaller scales.

However, the simplifications inherent in these sub-grid models can hinder our understanding of the connection between galaxy and SMBH evolution and growth, in particular our ability to predict the merging rate of binary SMBHs and how the inflow of gas on to the host galaxy (Bellovary et al. 2013) feeds their growth. For example, a common treatment of SMBH dynamics is to assume that they are always stable at the centre of their host galaxies, often obtained by shifting an SMBH towards the nearest potential minimum (Sijacki et al. 2015), a process we refer to as 'advection'. This approach fails to capture the Gyr timescale of sinking orbits for black holes during satellite accretions or galaxy mergers (Governato, Colpi & Maraschi 1994; Taffoni et al. 2003; Tremmel et al. 2015), resulting in an unrealistic coupling of SMBH and galaxy mergers, as well as artificially high accretion rates during these perturbing events. Furthermore, SMBH accretion is commonly calculated via a boosted Bondi-Hoyle prescription (e.g. Booth & Schaye 2009), but the assumptions of this approach break down for gas supported by rotation rather than internal pressure (Hopkins & Quataert 2010). Finally, SMBH 'seeds' have often been placed based on the host halo mass, irrespective of the local gas properties (Bonoli et al. 2016). This approach leads to a protracted epoch of SMBH formation and an occupation probability artificially connected to the observed population of active SMBHs, rather than to physically motivated models of SMBH formation, which predict seeds that form at very high redshift (Begelman, Volonteri & Rees 2006; Volonteri 2012).

The main goal of this paper is to present a set of novel implementations of SMBH physics, improving on the way SMBH formation, dynamics, and accretion are handled with sub-grid models in cosmological simulations. Specifically we:

(i) Connect SMBH seed formation to dense, very low metallicity gas that allows us to predict the SMBH population in both high-mass galaxies and dwarf galaxies.

(ii) Incorporate the sub-grid model for dynamical friction (DF) presented in Tremmel et al. (2015) so that SMBHs experience realistic dynamical evolution, allowing us to predict SMBH dynamics, the frequency and mass ratio distribution of SMBH mergers and SMBH growth during galaxy interactions.

(iii) Introduce a new sub-grid model for SMBH accretion that naturally accounts for the angular momentum support of nearby gas at resolved scales. This creates a more physical picture of how, when and where SMBHs grow compared to the more common Bondi-Hoyle prescription, while avoiding the additional assumptions and free parameters required by other current methods (e.g. Rosas-Guevara et al. 2015; Anglés-Alcázar et al. 2017).

We also outline a novel approach to constrain 'free' parameters within simulations, specifically those that govern star formation (SF) and stellar feedback as well as SMBH growth and feedback. Due to the computational cost, simulation studies have often relied on rerunning a small number of large volumes while only changing one parameter at a time (e.g. Schaye et al. 2015), only rarely running grids of simplified models (Governato et al. 2007). In this work we use a quantitative and efficient strategy, based on a large number of 'zoomed-in' cosmological simulations, to decide the optimal combination of sub-grid SF- and SMBH-related parameters for a given set of physical modules and resolution. This general strategy is not specific to our simulations and can be easily applied to any set of free parameters that govern any relevant physical processes.

In Section 2, we describe the simulations and in Section 3, we discuss the sub-grid parameter optimization technique. We describe the sub-grid models for SF and feedback in Section 4 and our novel approach to SMBH physics in Section 5. In Section 6, we present results from our flagship 25 Mpc volume, and in Section 7, we discuss the role of SMBH feedback in limiting SF compared to supernovae (SN) feedback alone, and in Section 8 we show how more common implementations of SMBH physics result in appreciably different galaxies compared to our implementation. Finally, in Section 9 we present an example that illustrates how our SMBH implementation allows us to study dual AGN in unprecedented detail and in Section 10 we summarize our results. In Appendices A and B we discuss the rationale behind our sub-grid parameter optimization approach and explain our model for post-processing dust absorption.

2 THE ROMULUS SIMULATIONS

2.1 CHANGA

The simulations are run using the new Tree + SPH code CHANGA (Menon et al. 2015), which includes standard physics modules previously used in GASOLINE (Wadsley, Stadel & Quinn 2004; Stinson et al. 2006; Wadsley, Veeravalli & Couchman 2008; Shen, Wadsley & Stinson 2010) such as a cosmic UV background, SF, 'blastwave' SN feedback and low-temperature metal cooling. The 'blastwave' implementation of SN feedback is a well-tested approach that has been shown to reliably reproduce observable properties of galaxies, including cored dark matter profiles in dwarf galaxies (Governato et al. 2010). This is distinct from 'super bubbles' (Keller et al. 2014), a newer approach to SN feedback that will be implemented in future simulations. The SPH implementation includes thermal diffusion (Shen et al. 2010) and eliminates artificial gas surface tension through the use of a geometric mean density in the SPH force expression (Ritchie & Thomas 2001; Governato et al. 2015; Menon et al. 2015). This update accurately simulates shearing flows with Kelvin–Helmholtz instabilities. Our flagship simulation, ROMULUS25, used up to 100 000 cores with good scaling. CHANGA (Menon et al. 2015) is part of the AGORA (Kim et al. 2014) code comparison collaboration.

2.2 Simulation properties

In addition to our flagship 25 Mpc per side uniform, periodic volume simulation (ROMULUS25), we are currently running a set of three zoom-in cluster simulations (ROMULUSC) comprising haloes of mass 10^{14}–10^{15} M$_\odot$ as well as a 50 Mpc per side uniform volume (ROMULUS50). Both ROMULUS25 and ROMULUSC will be run to $z = 0$ and ROMULUS50 will be run to $z = 2$. See Table 1 for a list of simulations and parameters. We use ROMULUS25 for the analysis in this

Table 1. Physics implementations in different simulations presented in this paper.

Name	Box size (Mpc)	Accretion[a]	SMBH dynamics[b]	ϵ_{SN}^{c}	Run to $z=$
ROMULUS8	8	Bondi+AM	Dyn. Frict.	0.75	0.5
ROMULUS25	25	Bondi+AM	Dyn. Frict.	0.75	0
ROMULUSC	N/A	Bondi+AM	Dyn. Frict.	0.75	0
ROMULUS50	50	Bondi+AM	Dyn. Frict.	0.75	2
Advect	8	Bondi+AM	Advection	0.75	0.5
Bondi	8	Bondi	Dyn. Frict.	0.75	0.5
highSN	8	N/A	N/A	2.0	0.5

a Bondi+AM denotes the implementation described in this work.
b 'Advection' denotes the method utilized in Sijacki et al. (2007) and 'Dyn. Frict' is that from Tremmel et al. (2015).
c How much energy per SN is coupled to gas (in units of 10^{51}ergs). All the runs have identical particle mass, force resolution and numerical parameters.

paper because it provides a large, uniform sample of low redshift galaxies. Smaller 8 Mpc uniform volume simulations are also used for our comparative studies (see Sections 7 and 8).

The simulations are run assuming a ΛCDM cosmology following the most recent results from Planck ($\Omega_0 = 0.3086$, $\Lambda = 0.6914$, $h = 0.67$, $\sigma_8 = 0.77$; Planck Collaboration XIII 2016) and at the same resolution, with a Plummer equivalent force softening of 250 pc. Unlike many similar cosmological runs, the dark matter distribution is *oversampled*, such that we simulate 3.375 times more dark matter particles than gas particles, resulting in a dark matter particle mass of 3.39×10^5 $M_\odot$ and gas particle mass of 2.12×10^5 $M_\odot$. This is an important shift from the standard approach of simulating the same number of gas and dark matter particles, as it allows us to decrease numerical noise and allow for more accurate black hole dynamics (Tremmel et al. 2015). Our mass resolution is better than recent large-volume simulations (Sijacki et al. 2015; Volonteri et al. 2016a) and our force resolution is comparable to the highest resolution runs of the EAGLES series (Schaye et al. 2015). Spline force softening converges to a Newtonian force at scales twice the gravitational softening, ϵ_g.

In order to showcase the results of our model and compare with other common SMBH implementations, we also run a series of 8 Mpc per side uniform volume simulations with different realizations of SMBH physics, as well as a simulation with no SMBHs and an enhanced SN feedback efficiency. These smaller simulations (e.g. ROMULUS8) have the same cosmology and resolution as our main ROMULUS data set. ROMULUS25, along with the 8 Mpc runs, makes up the data used in the analysis presented in this paper. For a complete list of simulations, see Table 1.

2.3 Halo and galaxy extraction

For all simulations referred to in this work, we use the Amiga Halo Finder (Knollmann & Knebe 2009) to extract individual haloes. We calculate galaxy properties based on all of the particles within a given halo. However, for a better 'apples-to-apples' comparison with observational results, we utilize the corrections from Munshi et al. (2013) to account for the mass of stars missed in observations and the baryonic effects on halo mass not accounted for in dark matter only (DMO) simulations. These corrections have been calibrated for haloes with virial mass $10^8 - 10^{12}$ $M_\odot$ and are shown to be roughly constant across this range. Specifically, $M_{\star,\,\mathrm{obs}} = 0.6 M_{\star,\,\mathrm{sim}}$ and $M_{\mathrm{vir,\,sim}} = 0.8 M_{\mathrm{vir,DMO}}$. We apply these corrections to haloes with M_{vir} as large as 10^{13} $M_\odot$. In these haloes, such corrections are

particularly necessary, as ~40 per cent of stars exist far from halo centre, either in an extended stellar halo or in satellite galaxies, and would not be included in observational estimates for stellar mass.

3 SUB-GRID PARAMETERS OPTIMIZATION

In CHANGA, SF and SMBH physics are regulated through a series of sub-grid prescriptions that parametrize unresolved physics into several free parameters. In order to set these parameters to their optimal values, we employ a quantitative optimization technique to map out the suitability of the parameter space and near-converge on the 'best' parameters. The idea of this approach is similar to that of Bower et al. (2010), but tailored specifically for more complicated simulations where only a few galaxies can be run with 10 s of different parameter combinations. A summary of the procedure is the following (see Appendix A for a more detailed description):

(i) We simulate a large number of sets of 4 'zoomed-in' galaxies (Governato et al. 2007, 2009) at the same resolution as ROMULUS, with halo masses ranging from $10^{10.5}$ to 10^{12} $M_\odot$, with dozens of different sub-grid parameter realizations.

(ii) We compare the properties of the resulting galaxies to local empirical scaling relations, grading each parameter set accordingly based on the logarithmic distance of each galaxy from the relation. The score of each parameter realization is then the sum of the distance (in log space) of each halo from each empirical relation.

(iii) The procedure is repeated, each time sampling in more detail around the best graded models until a best set of parameters is converged upon. The Kriging algorithm (see Appendix A) is used to efficiently explore parameter space and determine convergence.

A first set of simulations was run with only SF physics and with higher weight placed on reproducing the observed properties of lower mass galaxies, where the effect of SMBH physics should be less important. The parameters searched were the local SF efficiency, the density threshold for SF and the fraction of SN energy coupled to the surrounding gas (see Section 4). Once the best SF parameters were identified (with the SN efficiency being the most important overall), a second set of galaxies was run including SMBHs physics, leaving the SF parameters unchanged but varying (1) the SMBH accretion and (2) energy coupling efficiencies (see Section 5.3). For results from the SF parameter search, we point the reader to Appendix A and Anderson et al. (2017).

The $z = 0$ relations used to grade the galaxy sets were: (1) The stellar mass–halo mass relation, (2) the HI gas fraction as a function of stellar mass,[1] (3) the galaxy specific angular momentum versus stellar mass and (4) the SMBH mass versus stellar mass (SMBHs only). The first two scaling relations (Cannon et al. 2011; Haynes et al. 2011; Moster, Naab & White 2013) allow us to respectively constrain the SF efficiency over the whole Hubble time, and the low redshift gas depletion time (i.e the recent SF rates). Our simulations follow the HI abundance of gas so M_{HI} is derived explicitly from the total gas content of each halo. The relationship between stellar mass, angular momentum and morphology (Obreschkow & Glazebrook 2014) is a useful proxy of galaxy sizes as well as the removal of low angular momentum gas through feedback processes. The M_{BH}–$M_\star$ relation (Schramm & Silverman 2013) is a final test specific for SMBH physics. These four scaling relations control several fundamental aspects of galaxy formation connected to the regulation of SF, angular momentum evolution and the growth of SMBHs.

---

[1] ALFALFA data from private correspondence with Jessica Rosenberg.

Taken together, they provide useful, low-z constraints to our model without *unconsciously* biasing our effort to reproduce one specific scaling relation. For the sake of simplicity and to avoid biasing the analysis, we use just the raw logarithmic distance from each relation to determine the plausibility of each parameter set, implementing no weighting between different relations. However, we do exclude the dwarf galaxy from the morphological and SMBH relations, as explained in Appendix A.

When applied to setting three SF parameters, the technique was able to converge with little user input after 27 realizations (a total of 80 simulations; see Appendix A). For two SMBH parameters, we were able to find a suitable parameter set after 12 realizations (a total of 48 simulations; see Section 5.5).

This 'zoomed-in' approach to parameter optimization allows us to efficiently explore the parameter space without having to simulate as many parameter realizations as would be required for a standard random-walk Markov-chain. It presents several advantages over shutting off or including individual physics modules (Genel et al. 2014) or to running a small cosmological volume multiple times (Schaye et al. 2015, 2010), the main issue being that running large simulations, particularly those at high resolution, is computationally expensive and will result in only a very limited parameter space exploration. Using this approach, the non-linear effect of changing more than one parameter at a time can now be followed and the search for best parameters can cover the mass range of the final, large-scale simulation (which tend to have more massive haloes than small test volumes). Finally the set of zoomed-in runs provides a useful post-main run framework to understand significant deviations from observed properties of galaxies or SMBHs should they emerge from the production runs.

4 SF PHYSICS

As in our standard implementation (Stinson et al. 2006) for runs at this resolution, SF is regulated by:

(i) the normalization of the SF efficiency, $c_\star$, used to calculate the probability of creating a star particle from gas with dynamical time t_{dyn} and characteristic SF time, Δt, assumed to be 10^6 yr

$$p = \frac{m_{gas}}{m_{star}}(1 - e^{-c_\star \Delta t / t_{dyn}}),\qquad(1)$$

(ii) The fraction of SNe energy that is coupled to the ISM
(iii) the minimum density ($n_\star$) and maximum temperature ($T_\star$) thresholds beyond which cold gas is allowed to form stars.

The final values adopted for these three sub-grid parameters are:

(i) SF efficiency $c_\star = 0.15$
(ii) Gas temperature threshold, $T_\star = 10^4$ K
(iii) Gas density threshold, $n_\star = 0.2$ m$_p$/cc)
(iv) SNe energy coupling efficiency, ϵ_{SN}, of 75 per cent

SN feedback adopts the 'blastwave' implementation (Stinson et al. 2006). Gas cooling is regulated by metal abundance as in Guedes et al. (2011) and SPH hydrodynamics and thermal and metal diffusion are described in Shen et al. (2010) and Governato et al. (2015). Our simulations do not include H$_2$ cooling as their resolution is not sufficient to model individual star-forming regions. We use a Kroupa IMF (Kroupa 2001), with the associated metal yields.

It is important to note that without SMBH feedback, parameters that work the best for dwarf galaxies based on our grading criteria

(see Section 3) are different from those that work best for higher mass galaxies. The parameters used here represent those that grade the highest when dwarf galaxy results are more heavily weighted. The idea is to start with an SF model that performs very well at low masses and allow SMBH physics to create better results for high-mass galaxies.

One 8 Mpc cosmological simulation was also run, with $\epsilon_{SN} = 2$ (see Table 1). Note that $\epsilon_{SN} = 2$ can be justified by implying a top heavy IMF or contribution from 'early feedback' (Governato et al. 2015). During our parameter search, we found that this run produced galaxies in 10^{12} M$_\odot$ haloes that better matched observed relations, at the expense of dwarf galaxies. However, strong SNe feedback alone still results in too much SF at late times (see Section 7). We find the inclusion of SMBH feedback as described below is necessary to reproduce the 'bend' in the $M_\star$–M_{halo} relation at high halo masses while maintaining realistic dwarf galaxy properties.

5 MODELING BLACK HOLE PHYSICS

5.1 Seed formation

Unlike SMBH seeding methods *directly* tied to halo mass thresholds that are often utilized in other large cosmological volume simulations (e.g. Di Matteo et al. 2003; Schaye et al. 2015; Sijacki et al. 2015), our approach allows for a more realistic seeding at high redshift without any a priori assumptions regarding halo occupation fraction.

Instead, SMBH seed formation is connected to the physical state of the gas by converting a gas particle already selected to form a star (see Section 4) into an SMBH seed instead if it has:

(i) Low-mass fraction of metals ($Z < 3 \times 10^{-4}$).
(ii) Density 15 times that of the SF threshold (3 m_p/cc).
(iii) Temperature between 9500 and 10 000 K.

These criteria ensure that black holes form only from gas that (a) is collapsing quickly (i.e. faster than the SF timescale as it has not been turned into a star already) while (b) cooling relatively slowly, approximating formation cites predicted for SMBH seed formation (Begelman et al. 2006; Volonteri 2012).

The criteria above were not chosen via an extensive parameter search. Rather, they were empirically derived via analysis of star-forming gas particles in high redshift volume simulations. The model limits SMBH growth to the highest density peaks in the early Universe with high Jeans mass. This is a marked improvement over stochastic formation from star-forming gas, resulting in seed formation that occurs in environments that are different than the average unenriched star-forming region, as seen in higher resolution tests of SMBH formation sites (Agarwal et al. 2014; Habouzit, Volonteri & Dubois 2017). Because we are following conditions of gas at resolved scales (i.e. hundreds of pc), these criteria are designed to capture the regions where SMBH seeds should exist and then be able to grow quickly to large masses, regardless of the specifics of the true formation mechanism at unresolved scales.

The metallicity threshold of 3×10^{-4} was chosen to select gas that had seen very little chemical evolution. We found that choosing more strict (lower) metallicity criteria or colder gas, biases SMBH formation *away* from the densest regions of the early Universe, an undesired outcome due to the finite resolution of our runs, that we specifically decided to avoid. SF will often form stars nearly simultaneously with SMBH particles. As stars form and massive stars give off stellar winds and SNe explode, metal rich gas

permeates throughout the halo and beyond, effectively shutting down any potential seed formation within the parent halo as well as nearby haloes. Metal diffusion in SPH codes is explicitly regulated by a diffusion equation; here we follow the implementation in Shen et al. (2010) with coefficients for both metal and thermal diffusion both set to 0.03, which give realistic values for galaxy metallicity gradients in high-resolution dwarfs (Brooks et al., in preparation). Once formed, the SMBH seed mass is set to 10^6 $M_\odot$. To attain this mass, the newly formed SMBH accretes as much mass as it needs from surrounding gas particles (total mass is then explicitly conserved), representing rapid, unresolved growth. The initial mass, while somewhat higher than most theoretical estimates (Johnson et al. 2012; Volonteri 2012), is motivated by the fact that much of the early growth on to SMBH seeds, or the exotic objects that may precede them, can exceed 0.1 $M_\odot$ yr^{-1} and be governed by the environment and physical processes well below the resolution limit of our simulations (e.g. Hosokawa et al. 2013; Schleicher et al. 2013). In reality, SMBH seeds would likely attain a spectrum of masses early on, but since such processes are unresolved, we cannot differentiate between where a larger SMBH seed should grow. This mass is also sufficiently large compared to our DM and gas particle masses that the dynamics of all SMBHs will be well resolved (Tremmel et al. 2015). We verified that even with this initial mass, SMBH seeds that exist in unfavorable environments (i.e. dwarf galaxies) naturally have limited growth, with about 50 per cent of SMBH seeds having less than 10 per cent mass growth over a Hubble time. We also verify that SMBHs that grow to more than 10^7 $M_\odot$ have grown enough through accretion to be insensitive to initial conditions.

In much of our future analysis, including that presented in Section 6, we take growth occurring in SMBHs with mass less than 110 per cent of their initial mass as still undergoing their initial growth phases, a process that has not yet been observed and the physics of which is highly uncertain. Therefore, we will exclude these systems from analysis where appropriate.

Fig. 1 plots the distribution of formation times of all SMBH seeds formed using the above approach within the ROMULUS25 volume. As a comparison, we plot the distribution of seed formation times we would have using a halo mass threshold of 7×10^{10} $M_\odot$. This is meant to approximate a more common seeding mechanism utilized in other large cosmological simulations. Specifically, our threshold approximates this from the Illustris Simulation (Genel et al. 2014; Sijacki et al. 2015). Our approach forms SMBHs much earlier, closer to what would be expected in SMBH seed formation scenarios (Volonteri 2010, 2012; Habouzit et al. 2017). We note that similar halo threshold techniques that have lower threshold masses will form seeds earlier, though they will still have a more substantial tail towards low redshift formation times. In ROMULUS25 SMBH, seeds still form out to low redshift in some rare cases within small, unenriched haloes. These SMBHs constitute a very small fraction (<1 per cent) of the overall SMBH population in the simulation.

5.2 Black hole mergers

SMBHs are allowed to merge based on the same criteria as Bellovary et al. (2011). Once SMBHs become closer than two softening lengths in relative distance, they merge if they have low enough relative velocities such that they would be considered gravitationally bound to one another, i.e. $\frac{1}{2}\Delta \mathbf{v} < \Delta \mathbf{a} \cdot \Delta \mathbf{r}$, where $\Delta \mathbf{v}$, $\Delta \mathbf{a}$ and $\Delta \mathbf{r}$ are the relative velocity, acceleration and distance vectors between two SMBH particles.
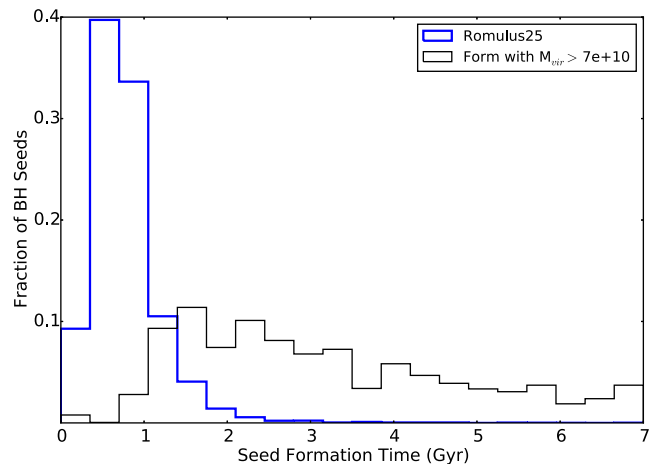
Figure 1. Seed formation times. The distribution of black hole seed formation times using our approach applied to a 25 Mpc run (ROMULUS25; blue line) compared to the seed formation if we applied a threshold halo mass criterion similar to other common approaches to seed formation in large simulations of this type (Di Matteo et al. 2003; Schaye et al. 2015; Sijacki et al. 2015). Using our scheme, black hole seeds form much earlier, the vast majority forming within the first Gyr of the simulation, similar to the expected formation epoch for SMBHs (Volonteri 2012). We compare this to the halo threshold scheme, meant to approximate that used in Sijacki et al. (2015), where haloes are seeded once a halo reaches a critical mass of $7 \times 10^{10} M_\odot$. Using this, black holes are seeded at much later times, even in the most massive haloes, which would cause the earliest periods of SMBH growth to be missed.

5.3 Black hole dynamics

DF, the force exerted by the gravitational wake caused by a massive object moving in an extended medium (Chandrasekhar 1943; Binney & Tremaine 2008), causes the orbits of SMBHs to decay towards the centre of massive galaxies (Governato et al. 1994; Kazantzidis et al. 2005). However, this effect is difficult to resolve in cosmological simulations due to numerical noise and limited gravitational force resolution. Our implementation includes a subgrid approach for modelling unresolved DF that has been shown to produce realistically sinking SMBHs (Tremmel et al. 2015). This allows us to follow the dynamics of SMBHs without assuming that they should always be stable at the centres of galaxies. As described in detail in Tremmel et al. (2015), our approach assumes that within ϵ_g from the black hole the velocity distribution is isotropic, giving Chandrasekhar's DF formula (Chandrasekhar 1943) for a BH of mass M and surrounding particle mass m_a with velocity distribution $f(\mathbf{v})$.

$$\mathbf{a}_{DF} = -4\pi G^2 M m_a \ln\Lambda \frac{\mathbf{v}_{BH}}{v_{BH}^3} \int_0^{v_{BH}} dv_a v_a^2 f(v_a). \quad (2)$$

The velocities of the BH and surrounding particles (v_{BH} and v_a, respectively) are both taken relative to the local centre of mass velocity within the smoothing kernel and $\ln\Lambda$ is the Coulomb logarithm. This equation can be further simplified by substituting the integral for $\rho(< v_{BH})$, which is the density of particles moving slower than the black hole.

$$\mathbf{a}_{DF} = -4\pi G^2 M \rho(< v_{BH}) \ln\Lambda \frac{\mathbf{v}_{BH}}{v_{BH}^3}. \quad (3)$$

Taking $\ln\Lambda \sim \ln(\frac{b_{max}}{b_{min}})$, we set $b_{max} = \epsilon_g$ to avoid double counting frictional forces that are already occurring on larger scales, which are well resolved due to the high-mass and spatial resolution of our

simulations. We take the minimum impact parameter, b_{min} to be the $90°$ deflection radius, with a lower limit set to the Schwarzschild Radius, R_{Sch}. The calculation is done using 64 collision-less particles (i.e. dark matter and star particles) closest to the black hole, with velocities taken relative to the COM velocity of all 64 particles.

A common technique in cosmological simulations is to reposition or push the SMBH along its local potential gradient (e.g. Di Matteo et al. 2005; Sijacki et al. 2007, 2015). However, these techniques (broadly referred to as 'advection' from here on) fail to properly reflect what is often a significant characteristic timescale for sinking SMBHs (see Tremmel et al. 2015, and references therein). During galaxy mergers, 'advection' techniques will result in a nearly immediate SMBH merger. It also prevents SMBHs from becoming perturbed away from galactic centre, which can affect SMBH growth during galaxy interactions and mergers.

With our approach instead, we are able to resolve the dynamics of SMBHs during and after galaxy mergers down to sub-kpc scales. The merger rates of SMBHs will be realistically decoupled from galaxy mergers. This will result in realistic SMBH growth and new predictions for gravitational wave observations. Our approach will also naturally produce dual and offset AGN down to sub-kpc distances, allowing us to study and understand these transient events in a broader evolutionary context (see Section 9).

5.4 Accretion and feedback

Black holes are allowed to grow by accreting mass from nearby gas particles. Energy from accretion is then isotropically imparted to the 32 nearest gas particles, distributing the energy among them according to the smoothing kernel. To ensure that the feedback energy is realistically dissipated, gas particles that receive energy from an SMBH are not allowed to cool for a time equal to the timestep of the SMBH (typically 10^3–10^4 yr), which is meant to represent the continuous transfer of energy during each SMBH timestep. This is a similar technique that is used in the Blastwave supernova feedback prescription, though here we utilize a different cooling shutoff time meant to approximate the continuous accretion and subsequent feedback that should occur during a timestep. The amount of energy coupled to surrounding gas particles is given by

$$E = \epsilon_r \epsilon_f \dot{M} c^2 dt, \tag{4}$$

where the radiative efficiency, ϵ_r, is assumed to be 10 per cent and the efficiency that energy couples to gas and ϵ_f is set to 2 per cent (see below for discussion on free parameter calibration). The accretion rate is assumed to be constant throughout one black hole timestep, dt.

The underlying assumption of these approaches is that the state of the gas at the smallest resolved scales drives the evolution of the unresolved physics on timescales relevant to the simulation.

The accretion rate, $\dot{M}$, is estimated via a modified Bondi–Hoyle prescription applied to the smoothed properties of the 32 nearest gas particles. The initial derivation of our approach is exactly the same as Bondi accretion. If we define some accretion radius, R, relative to the SMBH beyond which gas is bound to the black hole, and assume that mass continuity is roughly upheld on long timescales, the accretion rate on to the SMBH should be similar to the rate of mass flowing through a spherical surface of that radius:

$$\dot{M} \sim \pi R^2 \rho v. \tag{5}$$

Here v is the characteristic velocity of gas through the surface and ρ is the density of the ambient gas. In Bondi–Hoyle accretion, the calculation of the accretion radius, R, balances the SMBH's

gravitational potential and both the internal and bulk kinetic energies of the gas. In order to avoid underestimating the accretion rate due to resolution effects when calculating the density and temperature of nearby gas, we apply a density-dependent boost factor to this accretion rate, following the prescription of Booth & Schaye (2009), where the standard Bondi rate is multiplied by a density-dependent factor, $\left(\frac{n_{gas}}{n_*}\right)^\beta$, where β is a free parameter and n_* is the SF density threshold.

However, even with a well-motivated (but often poorly constrained) density boost, Bondi–Hoyle accretion is unable to account for angular momentum support, which often dominates the dynamics of cold gas at resolved scales, as in the discs of star-forming galaxies (Hopkins & Quataert 2010, 2011). Past efforts have focused on sub-grid models for angular momentum transport on sub-galactic scales (Anglés-Alcázar et al. 2017) or within the SMBH's accretion torus (Rosas-Guevara et al. 2015).

To take advantage of the improved spatial resolution of modern simulations, we implement an accretion algorithm that accounts for the angular momentum of gas *at resolved scales*. Our approach avoids any additional assumptions of sub-grid physics or free parameters beyond those required by the conventional Bondi–Hoyle prescription. Namely that the accretion rate, averaged over timescales relevant to the simulation, is a direct consequence of mass flux across the accretion radius, defined as the radial distance at which the gravitational potential of the SMBH balances the internal and bulk energetics of the gas as measured at the smallest resolved scales of the simulation.

In the reference frame of rotating gas, angular momentum provides an effectively lower gravitational potential such that $U_{eff}(\mathbf{r}) \sim -\frac{GM}{r} + \frac{j(r)^2}{2r^2}$, where $j(r)$ is the angular momentum per unit mass of the gas at distance r from the SMBH. We can replace $j(r)^2/r^2$ with v_θ^2, the rotational velocity of the surrounding gas. It is important to note that v_θ is distinct from the bulk velocity, which we will refer to as v_{bulk}, in the Bondi–Hoyle formula, which accounts for a flow of gas, not a coherent rotational motion.

If the dominant motion of the gas is rotational rather than a bulk flow, we can use the effective potential above and solve for R, ignoring order unity terms, such that the effective potential balances with the thermal energy of the gas, i.e $U_{eff} \sim c_s^2$. By definition the tangential motion must not contribute to the mass flux through our area. Returning to the simple equation for $\dot{M}$ above, we get the following relation:

$$\dot{M} \sim \frac{\pi (GM)^2 \rho c_s}{\left(v_\theta^2 + c_s^2\right)^2}. \tag{6}$$

Note that we do not assume v_θ is constant on unresolved scales, only that its value should inform the radius, R, at which the gravity of the SMBH dominates the gas dynamics. This is similar to the original Bondi–Hoyle formalism, where the energetics of gas far from the black hole are used to approximate the accretion radius. In this case, v_θ encapsulates the amount of angular momentum support the gas has on the smallest resolved scales, translating to a smaller accretion radius and therefore lower accretion rate.

To avoid uncertainties in particle dynamics below the force softening scale, we calculate the specific angular momentum, j, relative to a target black hole for gas particles that are between 3 and 4 softening lengths away (with our spline kernel softening, Newtonian forces are followed exactly at $2 \epsilon_g$). We then calculate the tangential velocity that gas one softening length, ϵ_g, away from the SMBH would have if the angular momentum on the larger scales was conserved, $v_\theta(\epsilon_g) \sim j/\epsilon_g$. The smallest relative velocity of the 32 gas
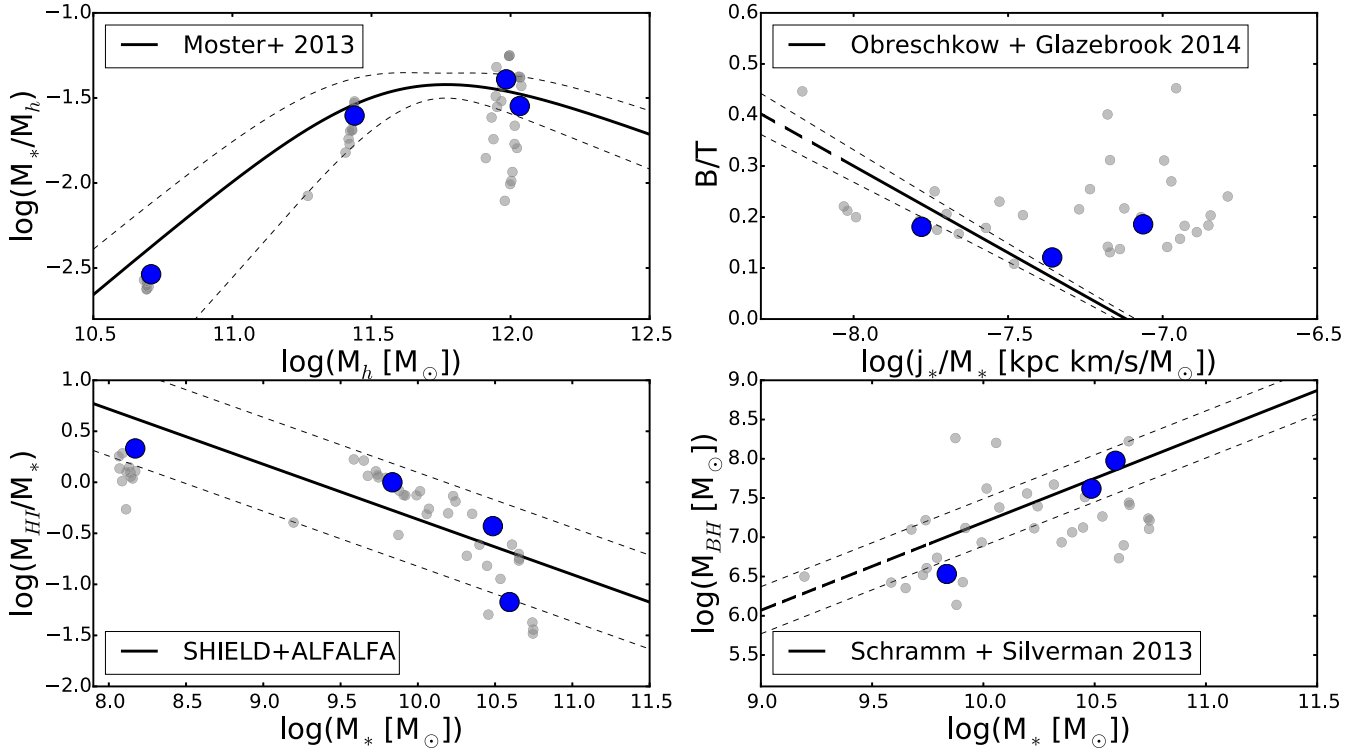
Figure 2. SMBH parameter optimization. Results from the search for optimal free parameters related to SMBH accretion and feedback. 12 realizations of accretion boost factor (β) and feedback efficiency (ϵ_f) for SMBHs were run, each with four zoomed in runs of galaxies. All of the models are shown in light grey points and the best-fitting model (the one that best matches overall to the four relations shown) is in blue. Each model is compared to different empirical relations governing SF efficiency (upper left; Moster et al. 2013), angular momentum (upper right; Obreschkow & Glazebrook 2014), HI content (lower left, derived from SHIELD and ALFALFA data; see Cannon et al. 2011; Haynes et al. 2011) and black hole growth (lower left; Schramm & Silverman 2013). The thin dashed lines represent 1σ errors. The thick dashed lines represent where each relation has been extrapolated beyond observations. The blue points have the parameters, $\beta = 2$, $\epsilon_f = 0.02$, which are what we implement in the ROMULUS models as well as the other simulations listed in Table 1. Note that for the angular momentum and SMBH mass tests, the dwarf galaxy was excluded. The former is due to the fact that angular momentum decomposition is difficult for a galaxy of this size. The latter is because observed SMBH masses are uncertain for dwarf galaxies and in our simulations, including in these parameter search runs, not every dwarf galaxy forms a SMBH.

particles closest to the SMBH, which we take as a proxy for $v_{\rm bulk}$, is compared to v_θ. If $v_\theta > v_{\rm bulk}$, we use equation (6) to calculate $\dot{M}$. Otherwise, we use the normal Bondi rate. Both calculations include the density-dependent boost factor, resulting in:

$$\dot{M} = \alpha$$

$$\times \begin{cases} \dfrac{\pi (GM)^2 \rho}{\left(v_{\rm bulk}^2 + c_s^2\right)^{3/2}} & \text{if } v_{\rm bulk} > v_\theta \\[2em] \dfrac{\pi (GM)^2 \rho c_s}{\left(v_\theta^2 + c_s^2\right)^2} & \text{if } v_{\rm bulk} < v_\theta \end{cases} ; \alpha = \begin{cases} \left(\dfrac{n}{n_{\rm th,*}}\right)^\beta & \text{if } n \geq n_{\rm th,*} \\[2em] 1 & \text{if } n < n_{\rm th,\star} \end{cases}.$$

(7)

Unlike Rosas-Guevara et al. (2015), we do not implement a viscosity parameter in our accretion rate calculation. This was an explicit choice made to avoid the inclusion of an additional free parameter and is justified by the fact that we are not attempting to approximate the behaviour of an accretion torus, as in Rosas-Guevara et al. (2015), where viscous timescales can be more critical. Still, there is uncertainty in the normalization of equation (7) when $v_\theta >> c_s$, which will be explored in future work. It should also be noted that equation (7) is not continuous at $v_{\rm bulk} = v_\theta$. We find this effect is sub-dominant compared to variations in density and velocity inherent to discreet calculations. This is shown in prac-

tice in Fig. 14, where our approach produces a less bursty accretion history in MW-mass haloes compared to normal Bondi accretion.

For the density-dependent boost factor, we compare the local density to the SF density threshold, $n_\star$, meant to represent the limit beyond which the simulation fails to resolve the multiphase ISM. The exponent β is a free parameter that we take to have a value of 2 (see the next section). Equation (7) is then compared to the Eddington rate, $\dot{M}_{\rm edd}(M)$, given the SMBH's mass at time t such that $\dot{M}_{\rm BH,final}(t) = \min(\dot{M}(t), \dot{M}_{\rm edd}(M_{\rm BH}(t)))$.

5.5 Calibration of SMBH free parameters

Our model of SMBH accretion and feedback has two free parameters controlling the accretion rate (β) and the efficiency at which radiated energy is transferred to surrounding gas (ϵ_f). In a similar approach as for the SF parameters (see Section 3 and Appendix A), we run 48 zoom-in cosmological simulations, in identical sets of four galaxies ranging from dwarf to Milky Way masses over several choices of these two parameters. Each set of simulations was run using the same set of SF parameters, optimized in a separate parameter search without SMBH physics (see Section 3). This ensures that we start with a model that performs as well as possible before the inclusion of SMBH physics. Fig. 2 shows the results of this search graphically. We tested values for β between 1.5 and 3 and values for ϵ_f between 0.005 and 0.1. Our parameter space exploration was
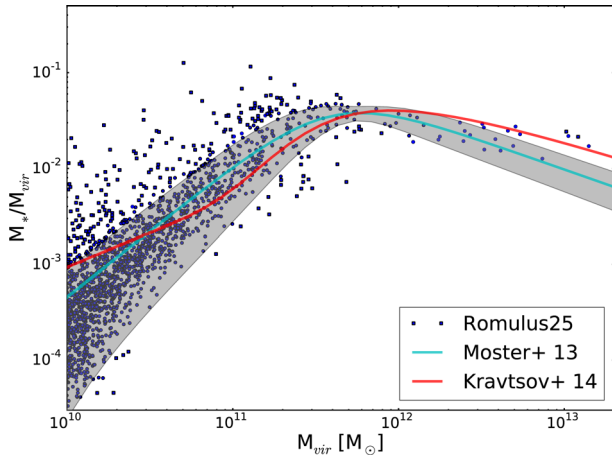
Figure 3. Stellar mass halo mass (SMHM) relation. Data from ROMULUS25 at $z = 0.25$ are shown in blue, plotted against two abundance matching relations from Moster et al. (2013) and Kravtsov, Vikhlinin & Meshscheryakov (2014). Any halo at least partially within the virial radius of a larger halo is not counted in this analysis in order to exclude satellites and interacting systems. The grey region shows the error in the Moster et al. (2013) relation, calculated from the errors reported for the best-fitting parameters. The stellar and virial masses for each halo are corrected to make them more directly comparable to observations following Munshi et al. (2013) (see Section 2.3). Our results match well with those from abundance matching. Of particular interest are the high-mass galaxies ($M_{\mathrm{vir}} > 10^{12}$ M$_\odot$), which indicate that SMBH feedback is correctly regulating their growth.

guided by the Kriging algorithm (see Appendix A). Each parameter set was graded in the same way as described in Section 3, each galaxy being compared to each scaling relation. Changing the parameters just for SMBH physics has enough of an effect to clearly isolate a 'best' set of parameters, i.e. the one in which the summed deviation of each galaxy from each scaling relation was the least. We find that the model that performs the best overall has $\beta = 2.0$ and $\epsilon_{\mathrm{f}} = 0.02$ and we adopted those values for all the production runs. While no explicit assumption has been made in the model for the mass scale at which SMBH feedback becomes important, the dwarf galaxy stellar mass and HI content exhibit neither minimal dependence on the SMBH model nor a dependence on the inclusion of SMBHs at all, as several iterations of the dwarf galaxy simulation never form a central SMBH.

6 FIRST RESULTS FROM ROMULUS25: THE BUILD-UP OF STARS AND BLACK HOLES

In this section we present initial results from our flagship ROMULUS25 uniform volume simulation, run to $z = 0$. It should be noted that such a small volume will miss some of the effects of large-scale structure and will not include the population of satellite galaxies in large haloes. We see this effect most strongly in regard to downsizing of both SF and SMBH accretion (see below). In future work, we will include the cluster simulations in our analysis as well. Within the scope of this paper, we find the ROMULUS25 simulation to be sufficient as a proof of concept that our method produces realistic galaxies and SMBHs at $z = 0$.

Fig. 3 shows the stellar mass halo mass (SMHM) relationship in ROMULUS25 at $z = 0$ after removing all satellite galaxies from the sample. Our results are consistent with results from Moster et al. (2013), which our model has been calibrated to reproduce, as well as Kravtsov et al. (2014) for haloes spanning more than three decades in mass. It should be noted that while these results are in part
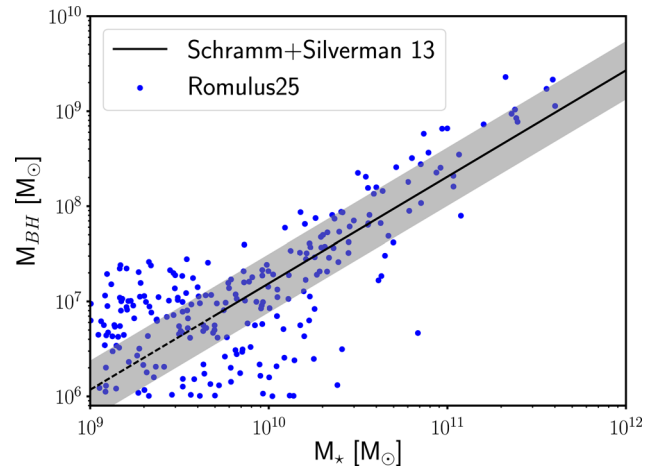
Figure 4. The SMBH Mass–Stellar Mass Relation. Each point plots the mass of the largest black hole in each galaxy against each galaxy's stellar mass, corrected by a factor of 0.6 from the total stellar mass in each halo (see Section 2.3). Also shown is the empirical relation from Schramm & Silverman (2013), where the grey region represents the 1σ scatter and the dashed part of the line is where the relation has been extrapolated past observations. The overall match to the data is good, particularly at higher masses. High-mass galaxies tend to exhibit less scatter and lie near the relation, though slightly biased towards higher mass SMBHs. Less massive systems show a broader scatter in black hole mass. The relation from Schramm & Silverman (2013) was derived from higher mass galaxies and there is evidence that smaller, star-forming galaxies lie on different relations (Reines & Volonteri 2015; Savorgnan et al. 2016).

due to our parameter calibration, the results for high-mass haloes ($M_{\mathrm{vir}} > 10^{12}$ M$_\odot$) have not been calibrated and can be considered predictions of our model. At $M_{\mathrm{vir}} > 10^{12.5}$, the ROMULUS25 haloes match better to the Kravtsov et al. (2014) results. As discussed in Section 2.3, we utilize the corrections from Munshi et al. (2013) for the stellar and virial masses to attain a more 'apples-to-apples' comparison. The correction, particularly when applied to larger group-size haloes, accounts for the mass that exists in extended stellar haloes and satellites.

Fig. 4 plots the mass of SMBHs in ROMULUS25 against the stellar masses of their host galaxies, again applying the correction from Munshi et al. (2013). Satellite galaxies have also been removed from this sample. This is another empirical relation that we had used to constrain our sub-grid model, so the fact that the simulation data match the relation from Schramm & Silverman (2013) is a success of our parameter search technique. High-mass galaxies show less scatter than low-mass galaxies, but are slightly biased to higher SMBH mass compared to the empirical relation. At low mass, we see a lot of scatter, both above and below the relation. While it is beyond the scope of this paper to examine in detail the nature of this scatter, it follows from recent observations that low-mass, star-forming galaxies have significantly more scatter in SMBH mass than higher mass galaxies, indicating that not all galaxies should lie on the same relation (Reines & Volonteri 2015; Savorgnan et al. 2016). The significant scatter above the relation could be explained by tidal stripping (Volonteri, Haardt & Gültekin 2008; Volonteri et al. 2016a; Barber et al. 2016), but we have removed satellite galaxies, making this connection less obvious. Likely it is due to stochastic SMBH growth in smaller galaxies. We will explore this further in future work.

The parameter search was meant to ensure that stars and SMBHs form and grow in the correct places. This is achieved in ROMULUS out to mass scales beyond those that the parameter search probed.
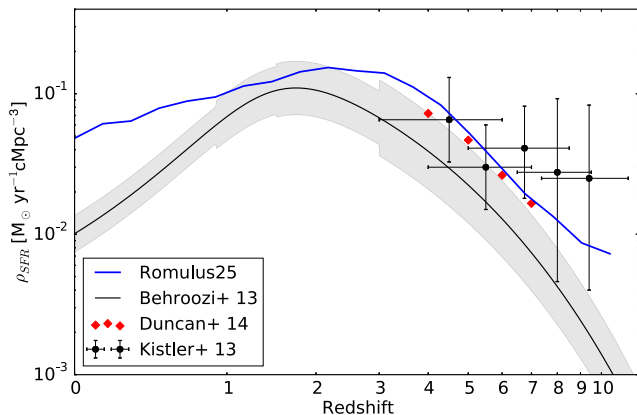
Figure 5. Cosmic star formation history. The solid blue line shows the total cosmic SF history in Romulus25 plotted against a fit to observation data from (Behroozi, Wechsler & Conroy 2013) as well as recent high redshift observations (Kistler, Yuksel & Hopkins 2013; Duncan et al. 2014). The grey region represents the spread in observational data for different redshift bins, as reported by Behroozi et al. (2013). Romulus25 accurately reproduces the evolution of the cosmic SF rate density at high redshift, reaching a maximum at $z = 2$ and declining towards lower redshift. The overproduction of stars at low redshift, which is in stark contrast with observations, is due to only a handful of high SFR systems, a result of our relatively small volume. A 25 Mpc volume lacks larger systems that would better sample the effect of cosmic downsizing at late times. At $z > 5$ a significant portion (50–90 per cent) of SF in Romulus25 occurs galaxies with stellar masses less than 10^8 M$_\odot$, a regime where the observed luminosity function is not well constrained (Anderson et al. 2017).

Of particular interest are the high-mass haloes ($M_{vir} > 10^{12}$ M$_\odot$) that were not explicitly constrained with our parameter search and represent the regime in which feedback from SMBHs dominates stellar feedback in regulating SF (Croton et al. 2006; Keller, Wadsley & Couchman 2016). The fact that these haloes produce galaxies with stellar masses very similar to abundance matching results as well as SMBH masses that are consistent with empirical scaling relations is a very promising result. How and when the growth of stars and SMBHs occurs in Romulus25 is also a testable prediction of the model.

Fig. 5 shows the cosmic SF history in Romulus25, which matches nicely with observations at high ($z > 2$) redshift, reaching a maximum just before $z = 2$ and then dropping off accordingly towards $z = 0$. At high redshift ($z > 5$), we find the bulk of SF is occurring in small galaxies ($M_\star < 10^8$ M$_\odot$), likely missed by high redshift observations (Anderson et al. 2017). This explains why ROMULSU25 lies above the derived SF history from Behroozi et al. (2013) but is more similar to estimates using more recent data that are more sensitive to lower mass galaxies. At low redshift ($z < 2$) Romulus25 lies far above the observed SF rates. This overproduction of stars at low redshift is due to only a handful of high SFR systems, a result of our relatively small volume that does not properly sample the higher density environments needed to recover the behaviour of cosmic downsizing at late times.

Fig. 6 plots the cumulative mass density accumulated in luminous SMBH accretion events across time. We only include data from SMBHs with mass greater than 110 per cent of their initial seed mass (see Section 5.1). We verify that excluding these systems does not substantially change our results. The cuts in luminosity are meant to not only show the contribution of different varieties of active SMBHs, but also ensure that we only sample the portion of the luminosity function that can be accurately constrained by
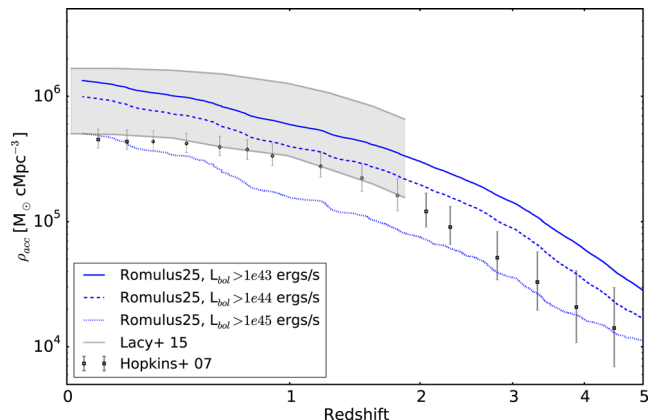

Figure 6. SMBH accretion history. The cumulative mass density accumulated in luminous SMBH accretion events in Romulus25 across cosmic time. SMBH growth is faster at high redshift and slows down at later times. Higher luminosity systems ($L_{bol} > 10^{44}$ ergs s^{-1}) account for 50–80 per cent of the accreted mass density at all times. The late time evolution at $z < 1$ is driven by a small number ($\sim 1 - 5$) of systems with $L_{bol} > 10^{45}$ ergs s^{-1}. These results are consistent with the integration of AGN luminosity functions out to high redshift (Lacy et al. 2015, shown as the grey region for a range of different values of radiative efficiency). Also shown are the results from Hopkins, Richards & Hernquist (2007), which are the result of different assumptions regarding absorption and bolometric corrections.

observations. We verify that for each luminosity cut, the contribution from low Eddington ratio SMBHs, where accretion is thought to become radiatively inefficient ($\lambda_{fedd} < 0.01$) is negligible. At early times, the black hole population grows more rapidly, slowing as it gets to lower redshifts. At all times the overall growth is dominated ($\sim 50 - 80$ per cent) by the more luminous SMBHs ($L_{bol} > 10^{44}$ ergs s^{-1}). Below $z = 1$, a significant fraction of this growth is taking place in a small number (1–5) of very luminous SMBHs ($L_{bol} > 10^{45}$ ergs s^{-1}). This is similar to the effect we see with SF, where our small volume is unable to appropriately sample AGN downsizing.

The overall growth of the black hole population in Romulus25 is consistent with observations. The grey region is from Lacy et al. (2015) and is obtained from integrating the observed AGN luminosity function between $z = 0$ and $z = 5$, assuming a radiative efficiency, ϵ_r between 0.06 (upper limit) and 0.18 (lower limit) and the data points with error bars are from Hopkins et al. (2007). The data from Lacy et al. (2015) were obtained from *Spitzer* observations in the mid-infrared. This makes them less sensitive to absorption, which can significantly impact optical and X-ray observations across all redshifts (Treister et al. 2010; Lansbury et al. 2015; Buchner et al. 2015; Lacy et al. 2015). However, the data from Lacy et al. (2015) are poorly constrained at redshifts higher than ~ 2, which is why we limit this region to $z < 2$. The higher luminosity data from Romulus25 ($L_{bol} > 10^{44}$ ergs s^{-1}) fit well with both observational data sets shown. The divergence away from the Hopkins et al. (2007) data at $z < 1$ is due to a small number of bright SMBHs, a consequence of our relatively small volume. Bolometric luminosities less than 10^{44} ergs s^{-1} represent a regime in which the observed luminosity functions are poorly constrained, particularly at high redshift, and sensitive to assumptions regarding the redshift and luminosity dependencies of absorption and bolometric correction (Merloni 2016).

These initial results show that Romulus25 (1) produces galaxies with stellar and black hole masses that are consistent with observations at low redshift (Figs 3 and 4) and (2) produces high redshift SF and SMBH accretion histories that are consistent with
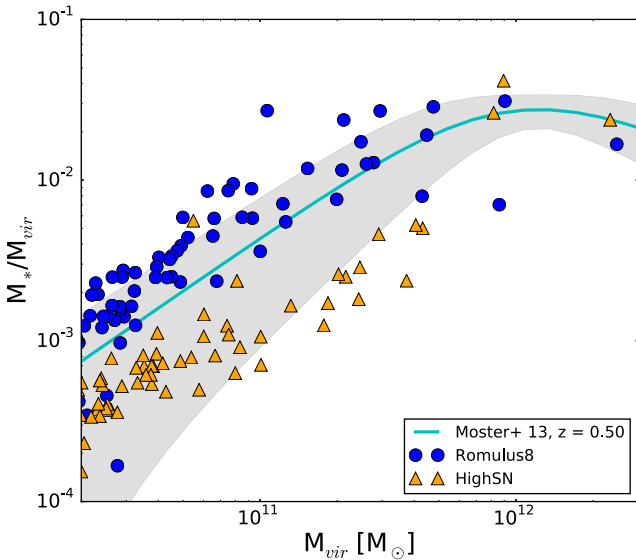
Figure 7. How feedback from SMBHs and SN affects SF efficiency. The SMHM relation for the ROMULUS8 (blue) and HighSN (orange) simulations. Increasing the efficiency of stellar feedback to produce stellar masses that match observations for higher mass galaxies (HighSN) causes an under-production of stars in low-mass systems. The high-mass galaxies match the observed relations well in the HighSN simulation, but this success is misleading, as the galaxies maintain significant SF through the end of the simulation (see Figs 8 and 11). The inclusion of black hole feedback combined with a lower stellar feedback efficiency (see Table 1) produces realistic stellar masses in haloes ranging from dwarfs to MW-mass.

observations, where differences arising at low redshift ($z < 2$) are due to our small volume not being able to properly capture downsizing for high-mass galaxies. These results show the strength of both our SMBH sub-grid model and our method for free parameter calibration. We leave the analysis of gas content and kinematics in ROMULUS for future work.

7 BLACK HOLE FEEDBACK COMPARED TO STELLAR FEEDBACK

In this section, we wish to explore the differences in SMBH and SN feedback mechanisms. It is often possible to tune parameters in order to reproduce observations of galaxies of a certain mass. During our parameter search (see Section 3 and Appendix A), we found that the models for SF and SN feedback without SMBHs that produced the most realistic galaxies in MW-mass haloes did not work well in reproducing realistic smaller galaxies. However, in this section we go beyond this to show that SMBH feedback is not only a crucial ingredient for reproducing scaling relations across all mass scales, it also has important consequences for reproducing the evolution of galaxies. In this case, we focus on MW-mass haloes ($M_{vir} \sim 10^{12}\ M_{\odot}$).

We compare two 8 Mpc uniform volume simulations, ROMULUS8 and HighSN (see Table 1), in order to gain insight into how the addition of extra feedback in the form of black holes compares to simply increasing the efficiency of SN feedback. The feedback efficiency in HighSN was chosen based off of the value we found to best reproduce scaling relations for galaxies in $10^{12}\ M_{\odot}$ haloes. The simulations are run to $z = 0.5$ to avoid some of the biases such a small volume will introduce into the evolution at later times.

Fig. 7 shows the SMHM relationship for the two simulations, plotted against the $z = 0.5$ best-fitting relationship from Moster
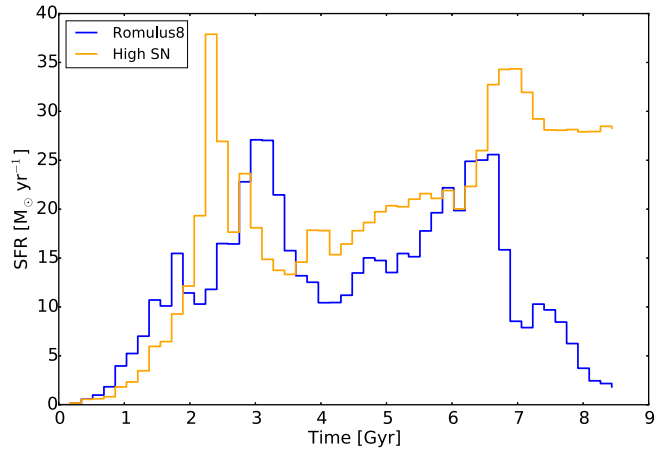

Figure 8. SMBHs and Galaxy quenching. The SF rate as a function of time for the most massive halo in the 8 Mpc volume, run with both the ROMULUS8 and HighSN models. The halo mass is consistent with being a Milky Way progenitor. The SF histories are similar up until about 2 Gyr prior to the end of the simulation. While the enhanced SN feedback is able to make stellar masses consistent with observations (see Fig. 7), the feedback from stars alone is unable to turn off SF at late times, which is expected for systems of this mass (Papovich et al. 2015). With lower SN feedback but the inclusion of black hole accretion and feedback (ROMULUS8), the galaxy is able both to attain a realistic stellar mass and to have SF quench before $z = 0.5$.

et al. (2013), applying the correction to stellar and halo masses from Munshi et al. (2013). The Romulus8 model fits the data well. The highSN model drastically underproduces stars in intermediate mass haloes. This is, of course, due to the fact that SN feedback is much more efficient in lower mass haloes that exhibit a shallower potential well (Governato et al. 2010; Brook et al. 2011). Such high efficiencies are necessary, however, to reproduce observed stellar masses in higher mass haloes without SMBH feedback. Because SMBH growth naturally depends on the host galaxy mass, SMBH feedback is able to preferentially limit the growth of higher mass galaxies, while not quenching the SF in low-mass haloes.

Fig. 8 shows the SF history of the most massive halo ($M_{vir}(z = 0) \sim 2 \times 10^{12} M_{\odot}$) in the volume for each simulation. While the final stellar masses are within realistic bounds in both simulations, the galaxy in ROMULUS8 has very low SF by the end of the simulation while the same galaxy in highSN fails to quench. The majority of galaxies (70–80 per cent) in this mass range should be quenched by $z = 0.5$ (Papovich et al. 2015).

Fig. 9 shows the colour evolution of the two most massive galaxies in the simulations run with SMBH physics (ROMULUS8) compared to that run only with enhanced SN feedback (HighSN). The galaxies show different colour evolution, with ROMULUS8 following much more closely the results from the CANDELS and ZFOURGE data (Papovich et al. 2015). Colours from stellar emission are calculated using tables generated from population synthesis models using http://stev.oapd.inaf.it/cgi-bin/cmd (Marigo et al. 2008; Girardi et al. 2010). Dust is accounted for using a simple approach based on metallicity and cold gas content of a galaxy (see Appendix B). In the highSN simulation, the colours of the galaxies remain dominated by dust at late times, never falling into the 'quenched' regime. The colour evolution also fails to follow the evolutionary path seen in the multi-epoch observations.

SMBH feedback, because it is more concentrated than SN feedback, is able to drive more powerful winds, which can disrupt inflowing material and lead to galaxy quenching (Volonteri et al.
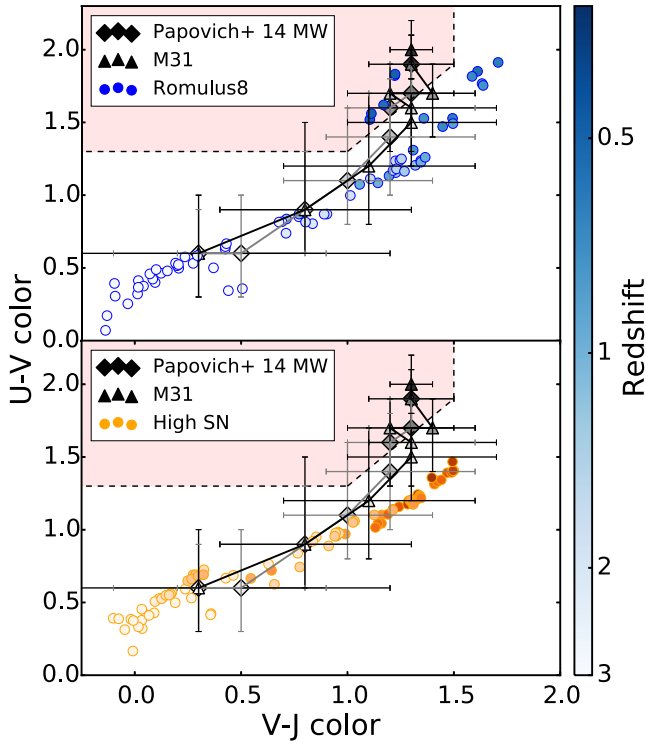
Figure 9. A colour–colour history of MW haloes in the Romulus8 and HighSN simulations with a simple prescription for the average dust attenuation (see Appendix B). Darker points represent lower redshifts. The observed data points (black) are from CANDELS and ZFOURGE, using abundance matching techniques to define Milky Way and M31 progenitors across cosmic time (Papovich et al. 2015). In Romulus8 (blue), the two Milky Way progenitors follow closely the average observed evolution, becoming quenched by $z = 0.5$. In the HighSN simulation (orange), the galaxy remains in the realm where colour is dominated by dust attenuation and ultimately fails to quench by $z = 0.5$. Without black hole feedback, Milky Way mass haloes remain very gaseous and dusty, with SF continuing at high levels.

2016b; Pontzen et al. 2017). Here we have shown that this effect is important for reproducing the observed evolution of MW-mass progenitor galaxies. One of the failures of simulations without SMBH feedback is the inability to quench galaxies in MW-mass haloes, something that our SMBH model is able to produce. Quenching galaxies in haloes of $\sim 10^{12}$ M$_\odot$ has generally been challenging for modern cosmological simulations (e.g. Bluck et al. 2016).

8 RESULTS FROM DIFFERENT BLACK HOLE PHYSICS IMPLEMENTATIONS

In this section we compare our implementation for SMBH dynamics and accretion (model Romulus8) against more common implementations found in large cosmological simulations (models Bondi and Advect). It is instructive to note that our parameter optimization was done using our SMBH implementation. While it may be possible to find a combination of parameters that create galaxies that fall on various empirical relations using these other models, the point of this section is to explore the effects that the additional physics our implementation includes have on galaxy evolution.

We are again using a smaller 8 Mpc uniform volume realizations of our main simulation suite. Fig. 10 shows the SMHM relationship of the three simulations. The high-mass end of the relationship is the only part noticeably affected by the different models, indicating
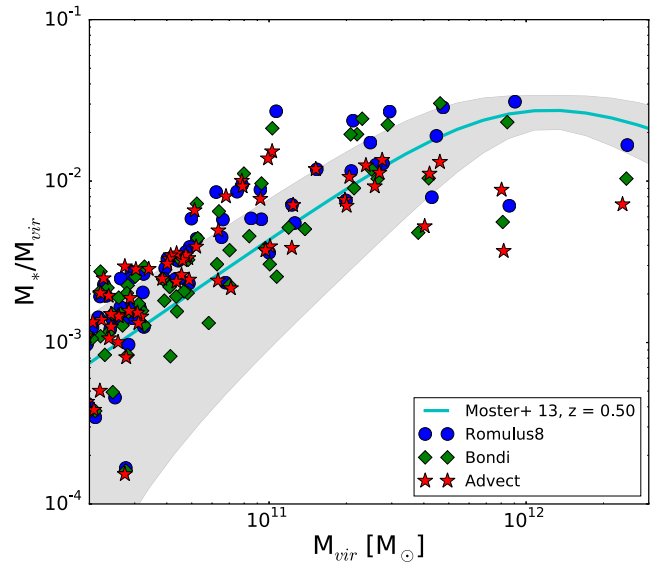

Figure 10. The effect of SMBH implementation on SF efficiency. The same as Fig. 7 but for the Romulus8, Advect, and Bondi simulations. The SMHM relation changes little between the simulations for low-mass haloes, but noticeable differences can be seen for haloes with virial masses above $\sim 2 \times 10^{11}$ M$_\odot$. SMBHs do exist in smaller haloes in this simulation (see Section 3.1) but, regardless of the SMBH physics implemented, small galaxies will not experience much black hole growth or feedback. For higher mass galaxies, artificial advection and Bondi accretion not limited by gas dynamics work to increase the effect of SMBHs on SF compared to our implementation utilized in the Romulus8 simulation.

that a lack of SMBH growth in low-mass galaxies is a natural consequence of the environment and not greatly affected by the choice of sub-grid SMBH physics. Both aspects of our implementation (described in section 3) work to soften the effect of SMBHs on their host galaxy, as both Advect and Bondi have lower stellar masses at a given halo mass.

Synthetic images of the stars in the central galaxy of the most massive halo in the volume are shown in Fig. 11, where a clear distinction between the three models can be seen. In Fig. 12, we plot the SF history of the most massive halo in the volume and the luminosity of the brightest black hole in that halo throughout time, averaged over 50 Myr intervals. While the SF histories are quite different between models, the accretion history of black holes in the halo are not strikingly different and at later times the Romulus8 model is the most active of the three.

The important difference in how black holes regulate the SF of their host galaxies occurs at high redshift. Fig. 13 plots the cumulative energy output of black holes within the central galaxy, tracking the halo backward in time along its main progenitor branch. The energies are reported relative to that in the Romulus8 model. We find that both Bondi and Advect experience more activity during the first several billion years of the simulation. The implications from this are that (1) early black hole activity can have important consequences for later galaxy evolution and (2) black hole dynamics and angular momentum limited accretion play an important role in determining accretion in the early Universe. It makes sense that the former is true, as the environment in which the black holes are active is different, namely the host halo is smaller, which would allow feedback from black holes to play a more drastic role in shaping the host galaxy. At early times, the black holes will exist in smaller galaxies that are undergoing more interactions, thus the
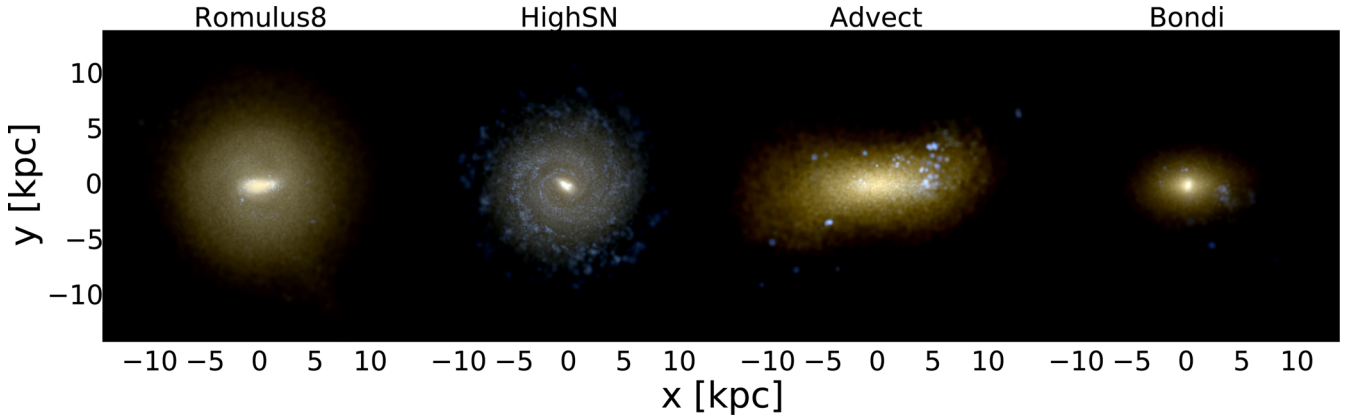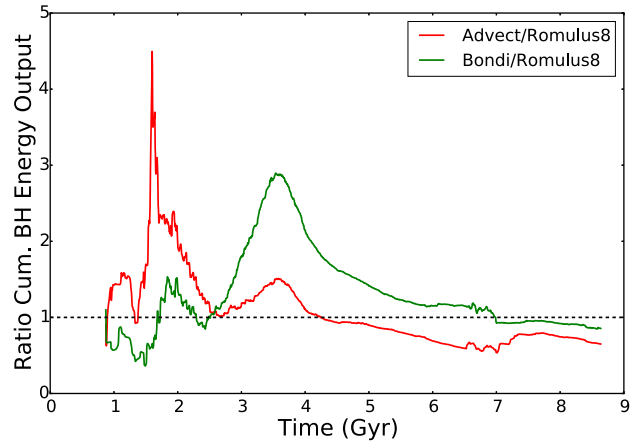
Figure 11. Mock images of stars in the largest galaxy from the ROMULUS8, HighSN, Bondi and Advect simulations at $z = 0.5$. The virial mass of the host halo is $\sim 2 \times 10^{12}$ M$_\odot$. On average, galaxies of this size should be quenched by this time (Papovich et al. 2015). Colours are based on the contribution of different bands within each pixel using U (blue), V (green) and J (red) assuming a Kroupa IMF, so young stars look blue and older stars look yellow. These images are indicative of the importance of physically motivated SMBH physics implementations on the evolution of large galaxies. It is clear that the inclusion of only SN feedback (HighSN) is not enough to quench the galaxy. SMBH feedback is able to quench in all cases, but the morphology and SF history (see Fig. 8) are noticeably affected by the details of the implementation.
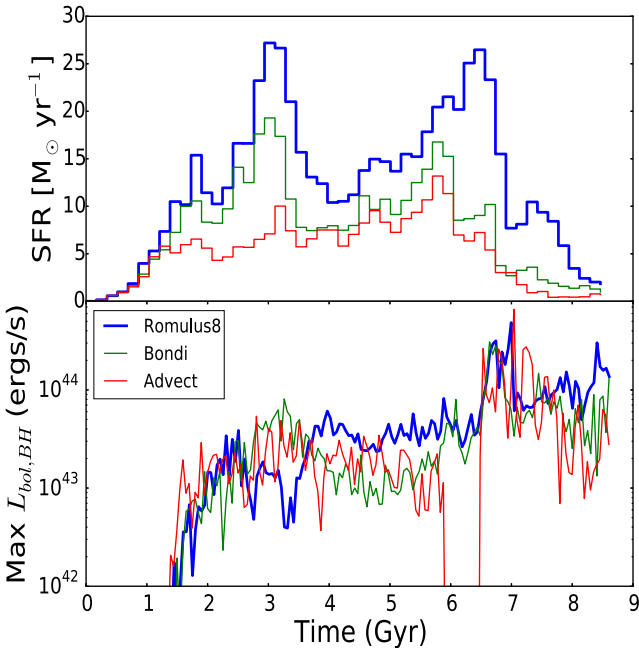
Figure 12. The effect of SMBH implementation on the SF history of massive galaxies (top). The SF history of the most massive halo in the 8 Mpc simulations, taken from the total stellar population of the galaxy at $z = 0.5$. A clear difference can be seen between ROMULUS8 (blue), Advect (red) and Bondi (green). (Bottom) For the same galaxy, the luminosity of the most luminous black hole across time within the galaxy's main progenitor branch. At later times ROMULUS8 has more active black holes. The values of luminosity are averaged over 50 Myr intervals. The strong dip in the red curve is due to the active black hole instantaneously transferring between two haloes during a major merger.

Figure 13. The cumulative energy output from SMBHs within the most massive halo in the 8 Mpc simulations. The Advect (red) and Bondi (green) models compared with ROMULUS8 across cosmic time. During the first 4 Gyr of the simulation, the ROMULUS8 halo experiences less feedback from SMBHs.

black holes are more likely to become perturbed away from the galaxy centre if they are allowed. In addition, at earlier times, when SF is climbing towards its peak, one would expect to see more cold, disc-dominated galaxies.

The black hole model affects not only where and when accretion takes place, but also how the accretion rate varies on smaller timescales. Fig. 14 plots the standard deviation of the accretion rate

for the most massive black hole in the most massive halo of the simulation, taken over intervals of 50 Myr and normalized by the average accretion rate throughout that time. For the entire simulation, both Advect and Bondi experience a significantly more bursty accretion history. So, while the smoothed accretion rate may look relatively similar between the models (see Fig. 12), there is a more bursty process occurring on smaller timescales. For Advect, the cause is numerical, as repositioning each timestep can cause the black hole to feel numerical noise, as the location of the potential minimum fluctuates (Wurster & Thacker 2013; Tremmel et al. 2015). In the Bondi simulation, the reason for such bursty accretion is that, without regulation, the accretion rate will rise quickly with gas density, which in turn will create a stronger feedback event that will drive gas back temporarily. The black hole then waits for the gas to relax again and the process continues. Including the gas dynamics in the accretion calculation softens this process because dense gas tends to also be in a disc, which will feel rotational support.

A future paper is planned to look in more detail of the relative effects of angular momentum limited accretion and stellar feedback
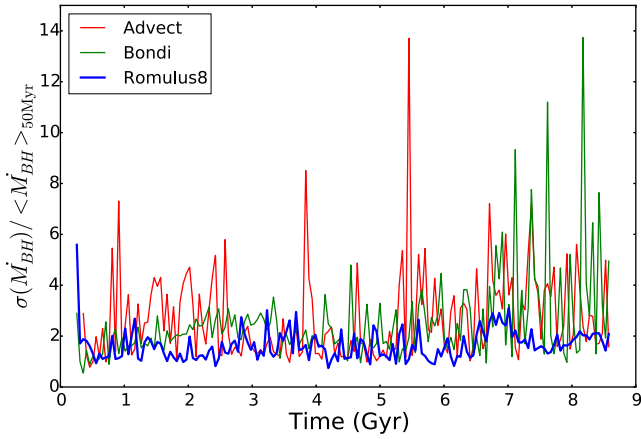
Figure 14. The 'burstiness' of SMBH accretion for the most massive black hole in the most massive halo in the three 8 Mpc simulations: ROMULUS8 (blue), Advect(red) and Bondi (green), defined to be the ratio of the standard deviation to the mean accretion rate over 50 Myr timescales. In both Advect and Bondi, we see that the black hole experiences a much more bursty accretion history.

on the evolution of Milky Way and sub-Milky Way mass galaxies. Within the scope of this paper, the important result is that both black hole dynamics and angular momentum regulated accretion have an appreciable effect on galaxy evolution for galaxies in higher mass haloes ($M_{\rm vir} > 10^{11.5}$).

9 APPLICATION: UNDERSTANDING DUAL AGN IN A LARGER CONTEXT

Dual AGN, systems with multiple active black holes, are beginning to be observed in the local Universe (Comerford et al. 2011, 2013, 2015) and represent an important regime for studies of SMBH-galaxy co-evolution, as they are a transient state possibly connected to a recent or ongoing galaxy merger. Being able to reproduce such systems in simulations is necessary to gain a theoretical understanding of their place in the broader context of SMBH-galaxy co-evolution. Some important work has already been done to that end (Van Wassenhove et al. 2012; Hirschmann et al. 2014; Steinborn et al. 2016) and the methods presented in this paper represent the logical next step.

Our approach to black hole physics is particularly well suited for realistically modelling dual AGN because we are able to accurately follow the dynamics of black holes within their host galaxies as they get perturbed away from centre or fall into a new host following a galaxy merger event. We are able to track the black hole orbits to an accuracy of the simulation's resolution limit (250 pc) and without making assumptions regarding the larger scale structure of the galaxy or halo in which the black hole resides. We are therefore not only able to create dual AGN down to a separation of <1 kpc, but we can follow the evolution of the system accurately throughout the parent system's evolution.

An example of dual AGN created using our approach, taken from the ROMULUS8 simulation during the last major merger of the most massive halo in the volume, is shown in Fig. 15 (the same halo used for analysis in the previous two sections; see Figs 8 and 11, and others in Sections 7 and 8). We show five snapshots in time of a
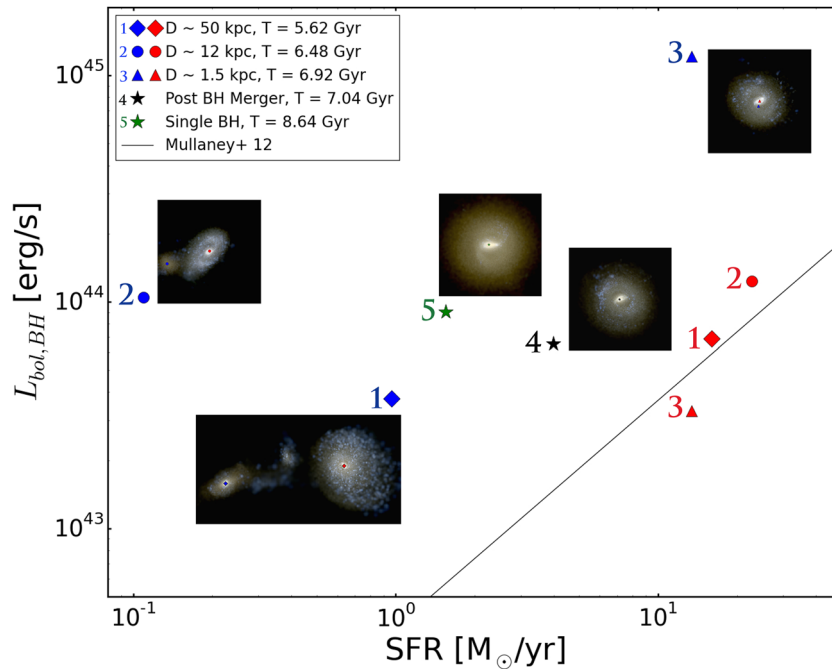
Figure 15. The evolution of dual AGN. The evolution a merging galaxy pair and resulting remnant galaxy in terms of SF rate and black hole luminosity. Thumbnails showing the stars of the galaxies are shown along with each data point set. The different colored points in each thumbnail represent the positions of the active black hole(s) at each time. The data points and thumbnails shown were chosen to encapsulate several important phases of evolution: (1) the beginning of the interaction, when the smaller galaxy has just entered the virial radius of the larger galaxy and is being stripped and environmentally quenched; (2) the end of the galaxy merger phase, where there are two distinct galaxies, but the smaller one has been completely stripped; (3) the remnant resulting from the galaxy merger, still with two separate, bright black holes; (4) just after the two black holes merge and (5) The merger remnant after it has been given time to relax, showing the galaxy quenching under the influence of a single, still very active black hole.

single galaxy merger that results in three instances of a dual AGN with separations of ∼50 , ∼12 , and ∼1.5 kpc. These are progenitor events leading up to a black hole merger and the quenching of the host galaxy, which by the end of the simulation, has halo and stellar masses similar to the Milky Way. By looking at each snapshot, we gain insight into how the simulation is evolving. The entire process takes less than 1.5 Gyr from the initial dual AGN event until the black hole merger. Two of these dual AGN events (snapshots 2 and 3 on the plot) look analogous to systems found by Comerford et al. (2015). When searching for these events, we defined 'active' to mean a bolometric luminosity of more than 10^{43} ergs s^{-1}.

To give the events more context, we plot the black hole luminosity as a function of SF rate for the merging galaxies in each snapshot (Fig. 15). The smaller galaxy is in the process of being stripped by the larger galaxy. The original baryonic masses of the galaxies before the merger was $M_1/M_2 \sim 1.2$ and the ratio of black hole masses was $M_{BH,1}/M_{BH,2} \sim 0.5$, where the less massive galaxy, denoted by 2, is the one that is being stripped and the one that hosts the more massive black hole.

The stripped galaxy is clearly in the process of being quenched by a combination of its environment and the active black hole within it. As the galaxies get closer, the black holes become more active. The SF rate of the larger galaxy remains roughly constant while the stripped galaxy is further quenched. Throughout the interaction, the black hole activity and SF rate of the more massive galaxy matches well with the relation derived from observations of $z = 1$–2 galaxies (Mullaney et al. 2012). The stripped galaxy always lies above the relation. After the two galaxies merge, the black hole originally in the stripped galaxy becomes even more active, with a luminosity much higher than expected given the SF rate in its new host. After the black holes merge, the central black hole remains very active and the galaxy moves further to the left on the plot as it quenches. This merger event, over the course of ∼1.5 Gyr and resulting in different instances of dual AGN, is the progenitor to a newly quenched galaxy. The heightened black hole activity corresponds with the quick decay of the SF rate over the next billion years.

In the example given here from the Romulus8 simulation, we show that the dual AGN event is a direct result of a major merger taking place between the galaxies. The black hole in the smaller galaxy becomes active as its galaxy quenches, with activity increasing as it moves closer to the more massive galaxy. In this case, having multiple black holes was indicative of a future black hole merger and would result in the quenching of what originally was a gas-rich, star-forming galaxy.

This is only one example, but it shows the level of detail with which we can approach the problem of dual AGN. It is also indicative that these systems are not necessarily very rare across cosmic time, as we were able to generate a relatively long-lived event in a volume of only 578 Mpc^3. In a future paper we will search both the 25 Mpc volume (Romulus25) and the cluster (RomulusC) for dual AGN events across cosmic time, giving us a much larger sample to look at and understand better the physical processes necessary to generate dual AGN.

10 SUMMARY

In this paper, we present a novel approach for modelling SMBH formation, dynamics and feedback that represents a marked improvement over currently common approaches utilized in most cosmological simulation to date. Our approach, combined with a new method of parameter optimization, has been applied to a new set of cosmological simulations called Romulus.

We presented the initial results from our flagship simulation, Romulus25, showing that our model reproduces the observed SMHM and M_{BH}–$M_\star$ relations for $z = 0$ galaxies. We also show that both the SF and SMBH accretion histories are consistent with observations at high redshift, though both suffer from our small volume's inability to capture cosmic downsizing. Using a set of smaller simulations, we also show how SMBH physics is a necessary component for quenching SF in massive galaxies and reproducing the observed evolution of MW-mass galaxies. We also show that our implementation gives appreciably different results for galaxies in massive (10^{12} M$_\odot$) haloes compared with more common approaches. This highlights not only the importance of including SMBH physics in cosmological simulations, but also that the details of the implementation are imprinted on the evolution of massive galaxies.

Finally, we present an illustrative example of how our implementation will not only result in realistic SMBH mergers, but also allow us to study the dual AGN that may often precede such events with unprecedented detail. This will be explored more thoroughly in future work, but represents an important proof of concept that our model will provide new data to put transient events such as dual AGN and SMBH mergers into a broader context.

The Romulus simulation suite, with resolution on par with the highest resolution cosmological simulations run to date, will provide a crucial data set with which to study the evolution of galaxies with halo mass $10^9 - 10^{13}$ M$_\odot$. The inclusion of RomulusC and Romulus50 will provide further insight into galaxy evolution in rarer, high-density regions not sampled by Romulus25 alone. The high resolution of these simulations is necessary not only to study the structure of galaxies, but also to properly follow the dynamics of SMBHs (Tremmel et al. 2015). The SMBH implementation we presented in this paper will allow SMBHs to form in the early Universe and exist in both large galaxies and dwarfs, while ensuring that they respond realistically to their changing environment. This is the first set of simulations of this size and resolution to simultaneously provide physically motivated sub-grid models for SMBH formation (Section 5.1) and dynamics (Section 5.2 and Tremmel et al. 2015) while also accounting for resolution effects (Booth & Schaye 2009) and dynamically supported gas (Section 5.3) when calculating SMBH accretion. Romulus represents a natural next step for cosmological simulations to provide more detailed insight into the evolving structure of galaxies, the co-evolution of galaxies and SMBHs, and transient events such as dual AGN and SMBH mergers.

Much of the analysis done in this work was done using the Pynbody package (Pontzen et al. 2013)

REFERENCES

Agarwal B., Dalla Vecchia C., Johnson J. L., Khochfar S., Paardekooper J.-P., 2014, MNRAS, 443, 648
Alexander D. M. et al., 2008, AJ, 135, 1968
Anderson L., Governato F., Karcher M., Quinn T., Wadsley J., 2017, MNRAS, 468, 4077
Anglés-Alcázar D., Davé R., Faucher-Giguère C.-A., Özel F., Hopkins P. F., 2017, MNRAS, 464, 2840
Barber C., Schaye J., Bower R. G., Crain R. A., Schaller M., Theuns T., 2016, MNRAS, 460, 1147
Begelman M. C., Volonteri M., Rees M. J., 2006, MNRAS, 370, 289
Behroozi P. S., Wechsler R. H., Conroy C., 2013, ApJ, 770, 57
Bellovary J., Volonteri M., Governato F., Shen S., Quinn T., Wadsley J., 2011, ApJ, 742, 13
Bellovary J., Brooks A., Volonteri M., Governato F., Quinn T., Wadsley J., 2013, ApJ, 779, 136
Bennert V. N., Auger M. W., Treu T., Woo J.-H., Malkan M. A., 2011, ApJ, 742, 107
Benson A. J., 2014, MNRAS, 444, 2599
Binney J., Tremaine S., 2008, Galactic Dynamics, 2nd edn. Princeton Univ. Press, Princeton, NJ
Bluck A. F. L. et al., 2016, MNRAS, 462, 2559
Bongiorno A. et al., 2014, MNRAS, 443, 2077
Bonoli S., Mayer L., Kazantzidis S., Madau P., Bellovary J., Governato F., 2016, MNRAS, 459, 2603
Booth C. M., Schaye J., 2009, MNRAS, 398, 53
Bower R. G., Vernon I., Goldstein M., Benson A. J., Lacey C. G., Baugh C. M., Cole S., Frenk C. S., 2010, MNRAS, 407, 2017
Brook C. B. et al., 2011, MNRAS, 415, 1051
Buchner J. et al., 2015, ApJ, 802, 89
Calzetti D., 2001, PASP, 113, 1449
Calzetti D., Armus L., Bohlin R. C., Kinney A. L., Koornneef J., Storchi-Bergmann T., 2000, ApJ, 533, 682
Cannon J. M. et al., 2011, ApJ, 739, L22
Chandrasekhar S., 1943, ApJ, 97, 255
Comerford J. M., Pooley D., Gerke B. F., Madejski G. M., 2011, ApJ, 737, L19
Comerford J. M., Schluns K., Greene J. E., Cool R. J., 2013, ApJ, 777, 64
Comerford J. M., Pooley D., Barrows R. S., Greene J. E., Zakamska N. L., Madejski G. M., Cooper M. C., 2015, ApJ, 806, 219
Croton D. J. et al., 2006, MNRAS, 365, 11
Di Matteo T., Croft R. A. C., Springel V., Hernquist L., 2003, ApJ, 593, 56
Di Matteo T., Springel V., Hernquist L., 2005, Nature, 433, 604
Draine B. T. et al., 2007, ApJ, 663, 866
Duncan K. et al., 2014, MNRAS, 444, 2960
Fan X. et al., 2001, AJ, 122, 2833
Filippenko A. V., Ho L. C., 2003, ApJ, 588, L13
Gehren T., Fried J., Wehinger P. A., Wyckoff S., 1984, ApJ, 278, 11
Genel S. et al., 2014, MNRAS, 445, 175
Girardi L. et al., 2010, ApJ, 724, 1030
Governato F., Colpi M., Maraschi L., 1994, MNRAS, 271, 317
Governato F., Willman B., Mayer L., Brooks A., Stinson G., Valenzuela O., Wadsley J., Quinn T., 2007, MNRAS, 374, 1479
Governato F. et al., 2009, MNRAS, 398, 312
Governato F. et al., 2010, Nature, 463, 203
Governato F. et al., 2015, MNRAS, 448, 792
Guedes J., Callegari S., Madau P., Mayer L., 2011, ApJ, 742, 76
Gültekin K. et al., 2009, ApJ, 698, 198
Habouzit M., Volonteri M., Dubois Y., 2017, MNRAS, 468, 3935
Häring N., Rix H.-W., 2004, ApJ, 604, L89
Haynes M. P. et al., 2011, AJ, 142, 170
Hirschmann M., Dolag K., Saro A., Bachmann L., Borgani S., Burkert A., 2014, MNRAS, 442, 2304
Hopkins P. F., Quataert E., 2010, MNRAS, 407, 1529
Hopkins P. F., Quataert E., 2011, MNRAS, 415, 1027
Hopkins P. F., Richards G. T., Hernquist L., 2007, ApJ, 654, 731
Hosokawa T., Yorke H. W., Inayoshi K., Omukai K., Yoshida N., 2013, ApJ, 778, 178
Johnson J. L., Whalen D. J., Fryer C. L., Li H., 2012, ApJ, 750, 66
Kazantzidis S. et al., 2005, ApJ, 623, L67
Keller B. W., Wadsley J., Benincasa S. M., Couchman H. M. P., 2014, MNRAS, 442, 3013
Keller B. W., Wadsley J., Couchman H. M. P., 2016, MNRAS, 463, 1431
Kim J.-h. et al., 2014, ApJS, 210, 14
Kistler M. D., Yuksel H., Hopkins A. M., 2013, preprint (arXiv:1305.1630)
Knollmann S. R., Knebe A., 2009, ApJS, 182, 608
Kormendy J., Ho L. C., 2013, ARA&A, 51, 511
Kormendy J., Richstone D., 1995, 33, 581
Kravtsov A., Vikhlinin A., Meshscheryakov A., 2014, preprint (arXiv:1401.7329)
Kroupa P., 2001, MNRAS, 322, 231
Lacy M., Ridgway S. E., Sajina A., Petric A. O., Gates E. L., Urrutia T., Storrie-Lombardi L. J., 2015, ApJ, 802, 102
Lansbury G. B. et al., 2015, ApJ, 809, 115
Lauer T. R., Tremaine S., Richstone D., Faber S. M., 2007, ApJ, 670, 249
MacKay D. J., 1998, NATO ASI Series F Computer and Systems Sciences, 168, 133
Marigo P., Girardi L., Bressan A., Groenewegen M. A. T., Silva L., Granato G. L., 2008, A&A, 482, 883
Menon H., Wesolowski L., Zheng G., Jetley P., Kale L., Quinn T., Governato F., 2015, Comp. Astrophys. Cosmol., 2, 1
Merloni A., 2016, in Haardt F., Gorini V., Moschella U., Treves A., Colpi M., eds, Lecture Notes in Physics, Springer Verlag, Berlin, Vol. 905, Lecture Notes in Physics, Springer Verlag, Berlin, p. 101
Monaco P., Theuns T., Taffoni G., Governato F., Quinn T., Stadel J., 2002, ApJ, 564, 8
Moran E. C., Shahinyan K., Sugarman H. R., Vélez D. O., Eracleous M., 2014, AJ, 148, 136
Mortlock D. J. et al., 2011, Nature, 474, 616
Moster B. P., Naab T., White S. D. M., 2013, MNRAS, 428, 3121
Mullaney J. R. et al., 2012, ApJ, 753, L30
Munshi F. et al., 2013, ApJ, 766, 56
Nozawa T., Kozasa T., Umeda H., Maeda K., Nomoto K., 2003, ApJ, 598, 785
Obreschkow D., Glazebrook K., 2014, ApJ, 784, 26
Papovich C. et al., 2015, ApJ, 803, 26
Planck Collaboration XIII, 2016, A&A, 594, A13
Pontzen A., Roškar R., Stinson G., Woods R., 2013, pynbody: N-Body/SPH analysis for python. Astrophysics Source Code Library, record ascl:1305.002
Pontzen A., Tremmel M., Roth N., Peiris H. V., Saintonge A., Volonteri M., Quinn T., Governato F., 2017, MNRAS, 465, 547
Reines A. E., Comastri A., 2016, PASA, 33, e054
Reines A. E., Deller A. T., 2012, ApJ, 750, L24
Reines A. E., Volonteri M., 2015, ApJ, 813, 82
Reines A. E., Sivakoff G. R., Johnson K. E., Brogan C. L., 2011, Nature, 470, 66
Reines A. E., Greene J. E., Geha M., 2013, ApJ, 775, 116
Ritchie B. W., Thomas P. A., 2001, MNRAS, 323, 743
Rosas-Guevara Y. M. et al., 2015, MNRAS, 454, 1038
Ruiz A. N. et al., 2015, ApJ, 801, 139
Savorgnan G. A. D., Graham A. W., Marconi A., Sani E., 2016, ApJ, 817, 21
Schaye J. et al., 2010, MNRAS, 402, 1536
Schaye J. et al., 2015, MNRAS, 446, 521
Schleicher D. R. G., Palla F., Ferrara A., Galli D., Latif M., 2013, A&A, 558, A59
Schramm M., Silverman J. D., 2013, ApJ, 767, 13
Shen S., Wadsley J., Stinson G., 2010, MNRAS, 407, 1581
Shields J. C., Walcher C. J., Böker T., Ho L. C., Rix H.-W., van der Marel R. P., 2008, in Bureau M., Athanassoula E., Barbuy B., eds, IAU Symposium, Vol. 245, IAU Symposium, pp. 259–260.
Shimizu I., Yoshida N., Okamoto T., 2011, MNRAS, 418, 2273

Sijacki D., Springel V., Di Matteo T., Hernquist L., 2007, MNRAS, 380, 877

Sijacki D., Vogelsberger M., Genel S., Springel V., Torrey P., Snyder G. F., Nelson D., Hernquist L., 2015, MNRAS, 452, 575

Somerville R. S., Hopkins P. F., Cox T. J., Robertson B. E., Hernquist L., 2008, MNRAS, 391, 481

Steinborn L. K., Dolag K., Comerford J. M., Hirschmann M., Remus R.-S., Teklu A. F., 2016, MNRAS, 458, 1013

Stinson G., Seth A., Katz N., Wadsley J., Governato F., Quinn T., 2006, MNRAS, 373, 1074

Sun M. et al., 2015, ApJ, 802, 14

Taffoni G., Mayer L., Colpi M., Governato F., 2003, MNRAS, 341, 434

Teyssier R., Moore B., Martizzi D., Dubois Y., Mayer L., 2011, MNRAS, 414, 195

Todini P., Ferrara A., 2001, MNRAS, 325, 726

Treister E., Urry C. M., Schawinski K., Cardamone C. N., Sanders D. B., 2010, ApJ, 722, L238

Tremmel M., Governato F., Volonteri M., Quinn T. R., 2015, MNRAS, 451, 1868

Van Wassenhove S., Volonteri M., Mayer L., Dotti M., Bellovary J., Callegari S., 2012, ApJ, 748, L7

Volonteri M., 2010, A&A Rev., 18, 279

Volonteri M., 2012, Science, 337, 544

Volonteri M., Bellovary J., 2012, Rep. Prog. Phys., 75, 124901

Volonteri M., Haardt F., Gültekin K., 2008, MNRAS, 384, 1387

Volonteri M., Dubois Y., Pichon C., Devriendt J., 2016a, MNRAS, 460, 2979

Volonteri M., Habouzit M., Pacucci F., Tremmel M., 2016b, in Kaviraj S., ed., IAU Symposium, Vol. 319, Galaxies at High Redshift and Their Evolution Over Cosmic Time, pp. 72–79.

Wadsley J. W., Stadel J., Quinn T., 2004, New Astronomy, 9, 137

Wadsley J. W., Veeravalli G., Couchman H. M. P., 2008, MNRAS, 387, 427

Wurster J., Thacker R. J., 2013, MNRAS, 431, 2513

APPENDIX A: QUANTITATIVE PARAMETER SEARCH FOR SF AND SMBHs PHYSICS

Large simulations require proportionally vast computational resources and face two main problems: limited force and mass resolution and the extensive need for sub-grid physics, as the modelling of physical processes happening below the resolved scales. Examples for such sub-grid physics parameters are the density at which SF should form, the fraction of energy from SNe and SMBHs that couples to the surrounding gas, and the speed at which metals diffuse in the intergalactic medium. Note that the same points hold even in simulations that claim no free parameters, for numerical parameters such as the precision of the step integration, the value of the force softening or the adopted IMF.

A common problem in simulations has been how to design an efficient strategy to *quantitatively* optimize, in a statistically controlled way these physical, but poorly constrained parameters, hence optimizing the results for the chosen physical model (Governato et al. 2007). A similar problem is faced by the so-called semi-analytical models (Monaco et al. 2002; Somerville et al. 2008). However, parameter searches for SAM are computationally cheaper and can be performed using different statistical approaches such as emulation (Bower et al. 2010), Monte Carlo Markov Chain (MCMC) (Benson 2014) or Particle Swarm Optimization (Ruiz et al. 2015).

As described in Section 3, in this work we have implemented a novel optimization technique to optimally choose sub-grid parameters associated with the implementations of (1) SF and SNe feedback and then (2) SMBHs accretion and feedback. To optimize the SF and SNe feedback parameters, we proceeded in the way described in Section 2.2. Here we describe in a more detail some of the choices we made and the so-called Kriging techniques (see be-

Table A1. EXAMPLE SET OF PARAMETER SPACE REALIZATIONS. The free parameters tested are the SN efficiency, ϵ_{SN}, the threshold density for SF, $n_\star$, and the SF efficiency, $c_\star$. Different sets of parameters chosen based on the Kriging technique until a 'best' set of parameters is converged upon (run 26 here). Note that these runs were done with lower DM mass resolution compared to ROMULUS (see the text).

| Run	$n_\star$	$c_\star$	ϵ_{SN}
1	1.000	0.2000	2.000
2	0.100	0.1000	1.000
3	0.100	0.1000	4.000
4	0.100	0.4000	1.000
5	0.100	0.4000	4.000
6	4.000	0.1000	1.000
7	4.000	0.1000	4.000
8	4.000	0.4000	1.000
9	4.000	0.4000	4.000
10	0.1	0.1	1.5
11	0.1	0.1	2.0
12	0.1	0.2	1.0
13	0.1	0.2	1.5
14	0.1	0.2	2.0
15	1.0	0.1	1.0
16	1.0	0.1	1.5
17	1.0	0.1	2.0
18	1.0	0.2	1.0
19	1.0	0.2	1.5
20	0.05	0.05	0.5
21	0.05	0.05	1.5
22	0.05	0.15	0.5
23	0.05	0.15	1.5
24	0.2	0.05	0.5
25	0.2	0.05	1.5
26*	0.2	0.15	0.5
27	0.2	0.15	1.5

low) that we used to map out the suitability of the parameter space explored. The Kriging algorithm penalizes parameter values that lead to simulations that deviate from the properties of real galaxies and then searches for parameter values that instead minimize this deviation. Runs are repeated with the same galaxies set, but with the updated parameters until the desired 'convergence' to the SF values listed in Section 4.

To summarize, our approach introduces a number of desirable qualities compared when only a limited number of experiments, as typical of numerical simulations, can be carried out. It presents several advantages over shutting off or including individual physics modules (Genel et al. 2014) or to running a small cosmological volume multiple times (Schaye et al. 2015, 2010). Namely, the non-linear effect of changing more than one parameter at a time can now be followed (Schaye et al. 2015) and the search for best parameters can cover a mass range similar to that of the final, large scale simulation (which tend to have more massive haloes than small test volumes).

(1) Minimal resources are wasted in 'bad' regions of parameter space.

(2) There is no need to wait for convergence, every simulation is useful immediately (unlike Markov chain Monte Carlo and many optimization techniques) and
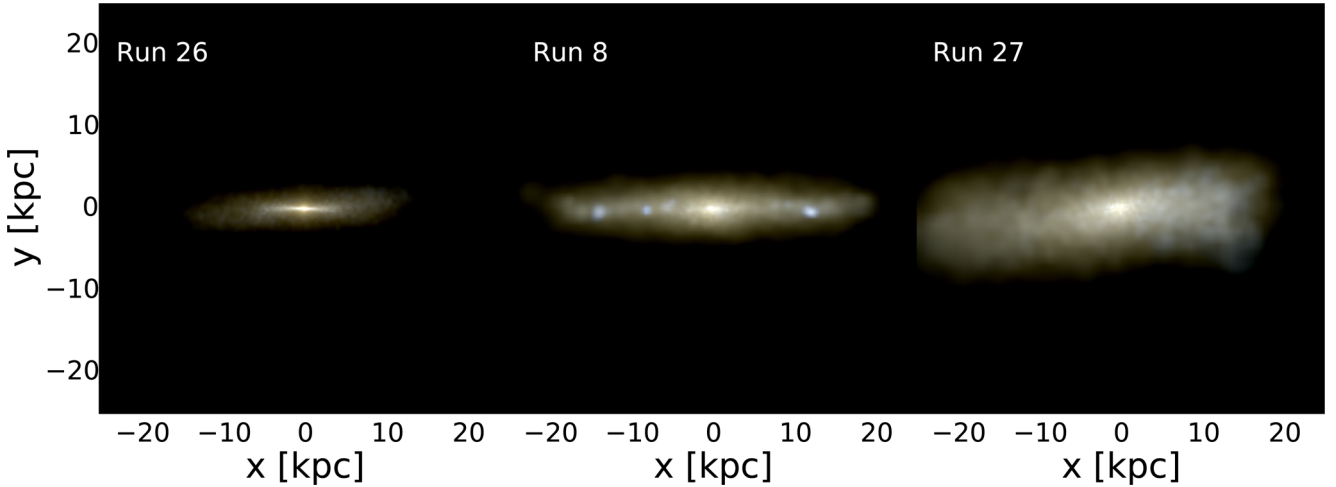
Figure A1. Kriging parameter search in practice. Here we show three realizations of our zoomed-in run of a $10^{11.5}$ M$_{\odot}$ halo at $z = 0$. From left to right we show the best parameter set, a poor set, and the worst set based on our grading criteria. See Table 1 for the parameters for each of these simulations. This illustrates that the parameters we chose and the way we varied them throughout our search has a clear effect on galaxy properties. In this case, run 26 has a clear thin disc, run 8 has a more diffuse disc and run 27 fails to form a thin disc at all. Our approach is able to thoroughly and efficiently search through the allowed parameter space and arrive at a set of parameters that results in realistic galaxies.

(3) Kriging is robust to changes in model choice and penalization/weighting methods as suitability values can easily be recalculated.

Fig. A1 illustrates the results of this process, showing images of the stars of a disc galaxy at $z = 0$ using the best, poor and worst SF parameters.

A1 Grading parameter realizations

Each parameter set realization is graded against a set of $z = 0$ empirical scaling relations that govern SF efficiency (Moster et al. 2013), the gas depletion time (Cannon et al. 2011; Haynes et al. 2011), galaxy size and angular momentum (Obreschkow & Glazebrook 2014), and SMBH growth (Schramm & Silverman 2013). The stellar mass fraction for our simulated galaxies is obtained following (Munshi et al. 2013), a procedure that includes the effects of a fixed aperture and the underweighting of older, redder stellar populations. The HI fractions are measured directly from the simulations, which track the HI content of each gas particle. To calculate the bulge to disc ratio, galaxies are decomposed into their different dynamical components based on the energy and angular momentum of each particle. Then the total angular momentum is calculated from every star particle not considered to be dynamically a part of the halo. Black hole masses are taken directly from the most massive black hole in each halo.

The SMHM relation constrains the SF efficiency over the whole Hubble time. SF efficiency also affects many other structural relations such as the $M_{\star} - V_{\text{peak}}$, and the stellar mass–metallicity relation. The $J_{\text{star}}/M_{\star}$ relation and the HI/stellar mass relation were included as good proxies of the effect of feedback processes on low redshift SF and the angular momentum distribution and size of a galaxy. Finally the M_{BH}-$M_{\star}$ is an important constraint on SMBHs processes, in particular the coeval growth of stars and SMBHs within galaxies. These grading choices are by no means unique, but allow us to be confident in the success of a given parameter set in creating galaxies that match what is observed in the local Universe, while still leaving room to make

predictions for the evolution of various galaxy properties over cosmic time.

A2 Finding the optimal parameters

In order to avoid a five-dimensional parameter space calculation, we first performed the full analysis, using the Kriging technique, on galaxies with no SMBH physics. This allowed us to converge upon the set of SF parameters that created the most realistic galaxies possible without the inclusion of SMBHs. A series of 27 parameter realizations (see Table 1) was run for sets of 3 haloes with $z = 0$ virial masses of $10^{10.5}$, $10^{11.5}$ and 10^{12} M$_{\odot}$. Each set was graded by summing up the logarithmic distance of each galaxy from each scaling relation, though the angular momentum of the dwarf galaxy was excluded due to the fact that the dynamical decomposition technique becomes unreliable at low masses. Each galaxy is weighted evenly in the final grade for each parameter realization. The best model converged upon by this approach is marked with a star in Table 1.

Once the SF parameters were chosen, another set of 12 simulations was run with SMBH physics to find the best parameters for accretion and feedback strength (see Section 5.4). The same general approach was used, though a more hands-on approach was used to dictate how we traversed the available parameter space (see below). Because SMBH physics is thought to preferentially affect more massive galaxies, we include a fourth halo, with virial mass 10^{12} M$_{\odot}$, in each set of simulations. When grading each parameter set, the average deviation of these two haloes is used instead of their individual deviations. Again, each galaxy is weighted evenly though the dwarf galaxy is again excluded from the SMBH relation due to the fact that the fraction of dwarfs hosting a central SMBH is not well known (see Volonteri (2010) for theoretical arguments and (Reines & Comastri 2016) for an observational review). Furthermore, as noted in Section 5.5, the inclusion of an SMBH does not have a significant impact on the scaling relationships.
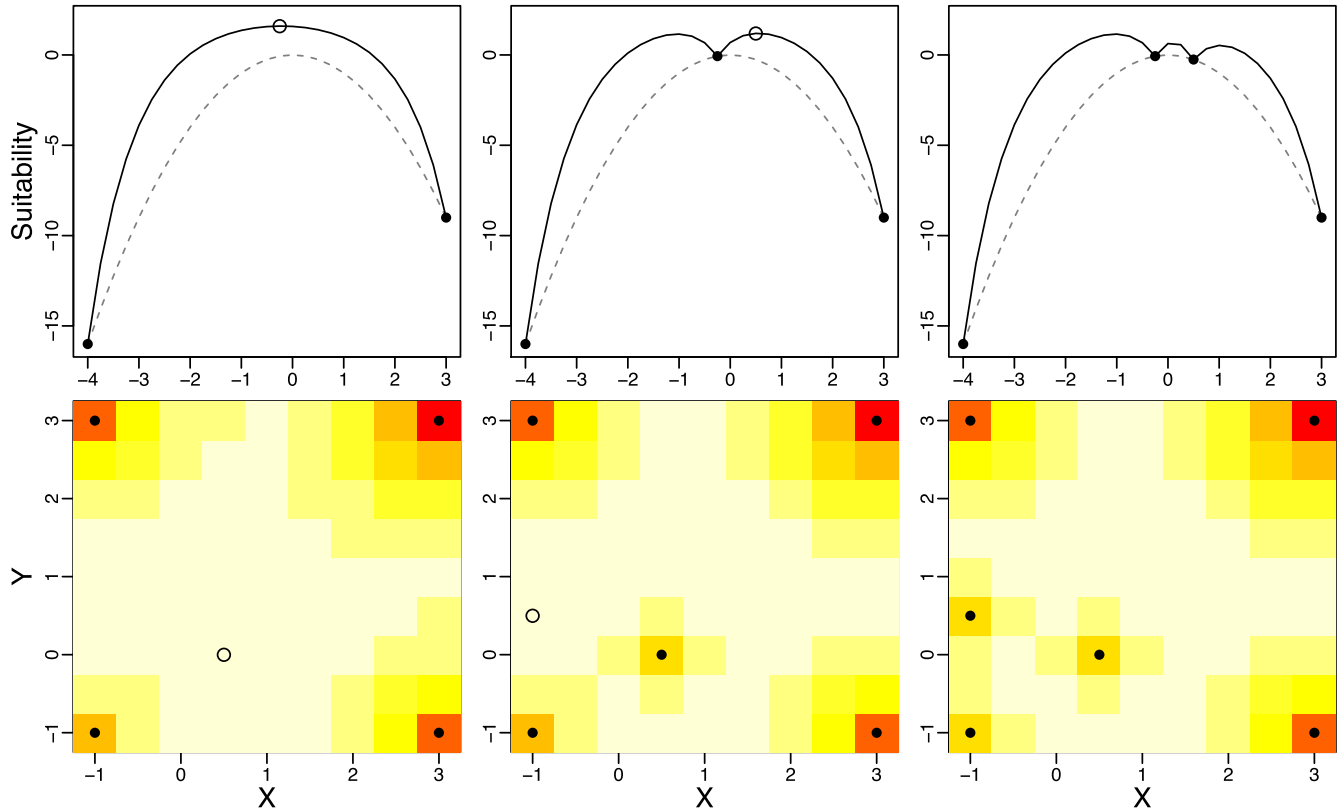
Figure A2. Kriging parameter optimization technique example. Two iterations of the Kriging search algorithm on a one-dimensional example (first row) and a two-dimensional example (second row). In the one-dimensional scenario, we are attempting to optimize the suitability function $-x^2$, shown as the dashed grey line. The algorithm starts by interpolating a pseudo-confidence manifold from two known points (filled points), and finding the greatest value (unfilled point). The algorithm then calculates the true suitability value for that point, and repeats the process. The two-dimensional scenario is similar – we attempt to optimize the suitability function $-\mathbf{x}^2$. Here, the pseudo-confidence manifold is shown as a heatmap, with red/darker representing lower suitability and white/brighter representing higher suitability. The point selection and evaluation process (filled and unfilled points) is identical to the one-dimensional scenario.

A3 The Kriging Approach to parameter search

The Kriging Approach allows us to efficiently traverse parameter space and know when we have converged on the 'best' set of parameters without the use of a large number of simulations as would be required of other techniques such as MCMC.

MCMC requires (1) a joint prior distribution on the parameter space, from which initial points can be drawn, (2) a likelihood function describing the distribution of the observables given a particular parameter set and (3) a proposal distribution that generates the next parameter values to examine, given the current values. MCMC then uses these functions to iterate over the parameter space, deciding whether or not to jump to the next point depending on how likely the next point is to explain the data relative to the current values. After a very large number of iterations (sometimes millions), the accepted points become a sample from the posterior distribution of parameter values given the observables, and useful inferences can thereby be derived, including point estimates of the best parameter values, and 95 per cent credible regions for where the best parameter values may lie.

Because cosmological simulations consume a large amount of computing resources, simulating so many iterations is not possible. Our approach trades the unattainable statistical properties of MCMC for the ability to make direct use of human expertise and intuition, the flexibility to adapt to changing measures of fitness, and keeping the certainty of knowing that every iteration makes a distinguishable contribution to our knowledge of the parameter space. While we lose access to the posterior distribution (i.e. a full sampling and ranking of parameter space), that is not really necessary. Instead, we gain an efficient means of finding the region of the parameter space that produces the most realistic galaxies, which is our goal.

We achieve all this by adapting Gaussian process Kriging techniques into a more intelligent and efficient grid search algorithm (see Fig. A2). We start by constructing a suitability function – a function that takes in a simulation and compares it to observed relationships and returns a score describing how realistic the simulation is (see above).

Using the following formula (MacKay 1998), we then interpolate the suitability function between all of our simulated points and put pseudo-confidence bounds around where the suitability function will actually fall. We see,

$$\mathbf{f}_*|X_*, X, \mathbf{f} \sim N(K(X_*, X)K(X, X)^{-1}\mathbf{f},$$
$$\times K(X_*, X_*) - K(X_*, X)K(X, X)^{-1}K(X, X_*)),$$

where X is the matrix of already-simulated parameter values, $\mathbf{f}$ is the corresponding vector of known suitability values, $X_\star$ is a matrix of new parameter values that we wish to examine, $\mathbf{f}_*$ is the corresponding as yet unknown suitabilities and $K(\ \cdot, \ \cdot)$ is a covariance matrix derived from a pre-specified covariance function $k(x_1, x_2)$.

Since we aren't seeking statistical properties, only utilitarian properties, we don't estimate the covariance scale so much as

choose one that spreads the first few suggested points away from the initial points, in our case a 99 per cent 'pseudo confidence surface' with covariance scale of 1, meant to ensure that parameter space is widely sampled. Since it is a suggestion algorithm rather than a statistical method, the covariance scale can even be adjusted freely before or after points have been selected and tested.

If we wish to take a hands-off approach, we would then examine the upper pseudo-confidence manifold, and instruct the algorithm to find the point with the highest potential suitability (at a fixed confidence level), and then numerically simulate that point. This is the approach we used for the initial search where we optimized the SF parameters (see above). However, if human intuition can be sufficient, we may also examine the pseudo-confidence manifold manually and select the next point ourselves without concern over losing statistical rigor. This is the approach we used for tuning the SMBHs parameters. One complication is that in regions of the parameter space where the Kriging process is extrapolating rather than interpolating, the confidence regions become extraordinarily wide, leading a naive algorithm to always select an extrapolated point. This has at least two solutions. One is to restrict any automation to the convex hull of already simulated points and use manual intervention to select points outside the convex hull if it becomes clear that such a point would make a good candidate. The second is to only calculate the Kriging bounds for a predefined, a priori reasonable region of the parameter space. The algorithm will quickly explore the outer boundary and then turn inward. From experience we learned that a good approach is to start the parameter exploration from a coarse grid of parameters values, including a range over which simulations will provide 'bad' results (e.g testing SN efficiency ranging from 0 to 4, values that will surely over- and underproduce stars). An option for future work would be to include higher-z constraints from the progenitors of massive present-day haloes, this would allow to constrain the high end of the present-day galaxy stellar mass function using a limited amount of computational resources.

A sample result of this process is in Table 1. By starting with a coarse grid of values for each of our three parameters, we utilize Kriging to traverse parameter space. After each iteration, Kriging sees both the current 'best' point and the algorithm will then run a simulation in a region not yet well enough constrained. With time, each parameter space realization gets closer to the 'best' values until Kriging tells us that it has sufficiently converged. Regions of parameter space that behave the worst are then sampled much less often while regions nearby the 'best' parameter set are sampled in more detail. The results presented in Table 1 are from simulations that do not oversample DM particles and therefore have lower mass resolution for DM than the ROMULUS simulations. We find that this increased resolution results in higher SF in dwarf galaxies. Thus, the ROMULUS simulations use the values from run 26, but with a higher SN efficiency of 0.75, a combination that we find results in final properties very similar to run 26 in Table 1.

APPENDIX B: DUST EXTINCTION APPROXIMATION

When comparing the colours of simulated galaxies to observations, it is important to account for dust attenuation. Because we only care about the average attenuation across all lines of sight integrated over all stars in a given galaxy, we utilize a simple 'spherical cow' approach similar to that of Shimizu, Yoshida & Okamoto (2011).

For a given dust distribution, the amount of attenuation can be calculated at any wavelength using the Calzetti Law (Calzetti et al. 2000), but first it must be properly normalized. For this, it is convenient to use far-UV light, since the extinction cross-section is roughly equal to the dust grain size. We choose 1600 Angstroms as our normalizing far-UV wavelength. We then make the assumption that the dust is uniformly distributed in a sheet around the stars, which allows us to relate the dust extinction by a simple function of the dust optical depth (Calzetti 2001).

$$A_\lambda \sim \frac{\tau_\lambda}{0.921} \tag{B1}$$

This is obviously not true in reality, but is not a bad assumption if we think of this calculation as an average over all lines of sight. Assuming spherical symmetry also makes the optical depth a simple function of average dust properties.

$$\tau\lambda = \int \sigma_d(\lambda) n(r) r \sim \frac{\sigma_d(\lambda) \Sigma_d}{m_p} \tag{B2}$$

In the above equation σ_d is the dust cross-section, Σ_d is the column density and m_p is the mass per dust grain. When dealing with far-UV light, the cross-section is just the cross-sectional area of the average dust particle. Because we are not accounting for structure within the gas, we can instead estimate the average column density using the total mass, M_d of dust within the galaxy and the half-mass radius, $R_{1/2,d}$ of the dust.

$$\Sigma_d \sim \frac{(1/2)M_d}{\pi R_{1/2,d}^2} \tag{B3}$$

The total mass in dust for a halo is given by the following relation from Draine et al. (2007) summed over the HI mass, $m_{HI,i}$ of every gas particle in a halo. We follow Shimizu et al. (2011) and normalize instead to the solar metallicity, rather than galactic O/H values as in the original paper.

$$M_d \sim \sum_{i=1}^{N_{gas}} 0.01 \frac{Z}{Z_\odot} m_{HI,i} \tag{B4}$$

These equations, put together with the physical properties of dust grains and applied to 1600 Angstroms, give us A_{1600}, which we can use to set the normalization of the Calzetti Law. We take the dust particle size to be $0.1\mu m$ and density to be $2.5g/cc$ (Todini & Ferrara 2001; Nozawa et al. 2003). We cap A_{1600} at a value of 2, given that more advanced dust models show that attenuation deviates significantly from its linear relationship with optical depth as column densities increase due to the fact that dustier systems will tend to be clumpier (Calzetti 2001). This normalization, combined with our adopted value of $R_v = 4.0$, gives us the ability to estimate the dust attenuation at any wavelength. When dealing with bands of wavelengths, we calculate attenuation using the central wavelength of the band.

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