Fueling Galaxy Growth Through Gas Accretion in Cosmological Simulations

Fueling galaxy growth through gas accretion in cosmological simulations

A dissertation presented

by

Dylan Rubaloff Nelson

to

The Department of Astronomy

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Astronomy & Astrophysics

Harvard University

Cambridge, Massachusetts

May 2015

Dissertation Advisor: Professor Lars Hernquist Dylan Rubaloff Nelson

Fueling galaxy growth through gas accretion in cosmological simulations

Abstract

Despite significant advances in the numerical modeling of galaxy formation and evolution, it is clear that a satisfactory theoretical picture of how galaxies acquire their baryons across cosmic time remains elusive. In this thesis we present a computational study which seeks to address the question of how galaxies get their gas. We make use of new, more robust simulation techniques and describe the first investigations of cosmological gas accretion using a moving-mesh approach for solving the equations of continuum hydrodynamics.

We focus first on a re-examination of past theoretical conclusions as to the relative importance of different accretion modes for galaxy growth. We study the rates and nature of gas accretion at $z = 2$, comparing our new simulations run with the Arepo code to otherwise identical realizations run with the smoothed particle hydrodynamics code Gadget. We find significant physical differences in the thermodynamic history of accreted gas, explained in terms of numerical inaccuracies in SPH. In contrast to previous results, we conclude that hot mode accretion generally dominates galaxy growth, while cold gas filaments experience increased heating and disruption.

Next, we consider the impact of feedback on our results, including models for galactic-scale outflows driven by stars as well as the energy released from supermassive black holes. We find that feedback strongly suppresses the inflow of "smooth" mode gas

at all redshifts, regardless of its temperature history. Although the geometry of accretion at the virial radius is largely unmodified, strong galactic-fountain recycling motions dominate the inner halo. We measure a shift in the characteristic timescale of accretion, and discuss implications for semi-analytical models of hot halo gas cooling.

To overcome the resolution limitations of cosmological volumes, we simulate a suite of eight individual 10^{12} solar mass halos down to $z = 2$. We quantify the thermal and dynamical structure of the gas in and around these halos. A radial sightline analysis allows us to measure the angular variability of halo gas properties, and demonstrate its increasing complexity at higher numerical resolution. We study the presence and characteristics of a strong virial shock, and make the link to recent observations of the circumgalactic medium surrounding galaxies.

We conclude with a technically oriented presentation of the full public data release of the Illustris simulation. Our goal is to facilitate a new era of robust comparisons, between state of the art theoretical models of galaxy formation and the many rich observational surveys of galaxy populations across cosmic time. We describe the data itself, as well as the comprehensive interface and set of tools we have developed for its analysis. We discuss scientific issues relevant when interpreting the simulations, technical details of the release effort, and future goals.

Contents

Acknowledgments

After a brief hiatus romping about the world I find myself, with six intervening years having passed rather suddenly, contemplating the onwards journey. In those six years, in Boston and at Harvard, my life has intersected with an innumerable number of people who have encouraged, shaped, and joined me along the way. As I draw up here the prominent names and memories, the improbability of doing an adequate job is clear, and I can hope only to recognize those few who have most influenced me throughout.

I would like to first thank my advisor, Lars Hernquist, for guiding me on my academic journey, and for teaching me something fundamental – not only about what science is and how it comes to be – but what our role as scientist is, and how, critically, to proceed. Lars always provided unending encouragement and unending resources, to pursue any topic or direction in which my interests veered. He spearheaded my interaction with the community, including several rather memorable trips abroad: Barcelona, Cefalu, Dublin, Heidelberg, Hamilton Island, Cambridge, Munich, Copenhagen. Perhaps most significantly, in leaving Harvard I feel immeasurably and confidently prepared to tackle independent research going forward, thanks to his unfailing guidance and mentorship.

My closest collaborators and mentors – Shy Genel, Mark Vogelsberger, and Annalisa Pillepich – who have worked with me through many long nights, deserve my deepest gratitude. I owe my growth in large part to their tireless leadership, assistance, and of course, criticism. I must also thank Paul Torrey, Volker Springel, Dušan Kereš, and Debora Sijacki, all of whom have been instrumental in my success and each irreplaceable in the roles they have played.

The Hernquist cohort, continually evolving over the years, has been a constant source of support, ideas, good company, and of course, good wine. I would like to thank those older students who made sure to shatter any illusions I might have held and teach me the way of things: Diego Muñoz, Paul Torrey, Greg Snyder, Chris Hayward, Gurtina Besla, and Laura Blecha. And Robert Harris. And Wen-fai Fong. I would also like to thank the younger students for being compatriots, and for always challenging me: Vicente Rodriguez-Gomez, Josh Suresh, Sarah Wellons, Philip Mocz, Aaron Bray, Luke Kelley, Eddie Chua, Ben Cook. Finally, I would like to thank the greater sphere of intersecting postdocs and past students, particularly Federico Marinacci, Rahul Kannan, Brendan Griffen, Laura Sales, Cameron McBride, Elena D'Onghia, Simeon Bird, Akos Bogdan, Patrik Jonsson, Claude-Andre Faucher-Giguère, Phil Hopkins, and the entire HITS team. I thank Eliot Quataert for inspiring me, and undoubtedly an entire generation of young students, to first set foot along this path.

I would like to thank all of the astronomy graduate students for many shared adventures, including my year: Yucong Zhu, Bekki Dawson, Max Moe, Tanmoy Laskar, Li Zeng, and Ragnhild Lunnan. Office mates, past and present – Diego Muñoz and Max Moe, for many distracting conversations, shared commiserations, and late nights. And, the tireless efforts of Peg Herlihy, Nina Zonnevylle, and Robb Scholten.

Finally, I would like to thank my friends and family, who have truly shaped my dreams, my self, my outlook on life. I thank Kat Deck for joining me on an extended and rather enjoyable arc through life, which will always be inseparable from the memories of my time in Boston. I thank my father, Stuart Nelson, my mother, Cinde Rubaloff, my brother Cameron and sister Anai – all of whom I love deeply – for supporting my every step through life. Where I am today and where I go tomorrow are because of you.

For friends, family, and all who share the journey.

Chapter 1

Introduction

In this first chapter we motivate the investigations carried out in this thesis. The present study is placed in the relevant context with an overview of the role of simulations in astrophysics – those concerned with cosmology and galaxy formation in particular. The fundamental physical models that we adopt are described, and the basic computational methods for their numerical solution introduced.

1.1 Background and motivation

In the long history of science the use of numerical computation is a relatively recent addition. Although many ancient civilizations used analog devices for the purposes of computation – the Antikythera mechanism from Greece being a prime example (Freeth et al., 2006) – such systems were fixed function, and not generally programmable. The ancestors of modern digital computers arose in earnest in the 1940s out of military considerations during the Second World War. Their uses included ballistic calculations,

cryptography, weather prediction, and weapons research and development.

The first uses of computation in astrophysics were concerned principally with the gravitational dynamics of few body systems. This problem lends itself to numerical solution, as Newton's law of gravitation implies that, to directly solve for the forces on N point masses, $N(N-1)$ pairwise distances must be calculated – the to be avoided $\mathcal{O}(N^2)$ scaling for "direct" N-body simulations. The complexity of the problem increases rapidly with N. Although students are likely to derive and study the analytical solution for the $N = 2$ body problem in their first year course, no general analytical solution for $N \geq 3$ is known to exist.

In a heroic pioneering effort, Holmberg (1941) studied the tidal structure induced by two interacting systems of 37 particles each. In order to compute the gravitational forces he replaced gravity with light (the strength of both declining as $1/r^2$ with distance), representing each particle as a light bulb, measured the light intensity at each point with a photo-sensitive cell, and integrated the system forward in time by physically moving the light bulbs in the table-top experiment. Moving into the realm of digital computers, von Hoerner (1960) and Aarseth (1963) studied stellar clusters with direct simulations including up to $N = 100$ particles. The size and methods for the gravitational N-body problem have since branched, according to whether the problem deals with a collisional or collisionless system. For cosmological as well as galactic scales, the relaxation time (Chandrasekhar, 1943) is either greater than, or significantly greater than, than the age of the universe, implying that to good approximation these systems can be regarded as collisionless.

In the 1970s, and first to study the gravitational collapse of structures in an

expanding universe, Press & Schechter (1974) performed numerical experiments with $N = 1000$ particles, as did Frenk et al. (1983) in the study of hierarchical clustering and filamentary collapse in large-scale structure. In the 1980s several methods were devised which improved the scaling from $\mathcal{O}(N^2)$ to $\mathcal{O}(N \log N)$. Namely, hierarchical multiple methods, or tree algorithms (Barnes & Hut, 1986; Hernquist, 1987), and particle-mesh techniques (Hockney & Eastwood, 1981). Under these approximations, $N = 10^6$ simulations became possible (Park, 1990), and resulted in fundamental explorations of the cold dark matter (CDM) hypothesis (e.g. Davis et al., 1985), as well as the discovery of the universal density profile of dark matter halos (Navarro et al., 1997).

The rapidly increasing speed of generally available compute clusters, and codes developed to exploit their parallel, distributed memory architectures led to $N = 10^9$ (Evrard et al., 2002) scale cosmological realizations, as well as perhaps the most well-known of all such simulations, the Millennium Run (Springel et al., 2005b) with $N = 10^{10}$ particles. As of 2015, the largest N-body, gravity only calculations now routinely include one trillion (i.e. $N = 10^{12}$) individual particles, enabling simulations of volumes which can exceed that of the visible universe (Kim et al., 2011; Skillman et al., 2014). Newer numerical techniques have since been developed, including the fast multiple method (FMM; Greengard & Rokhlin, 1987) which reduces the theoretical complexity further from $\mathcal{O}(N \log N)$ to $\mathcal{O}(N)$ (Dehnen, 2000). However, they have found limited use in astrophysical simulations, due mostly to a significantly higher algorithmic complexity. In cosmology, many of the largest runs are also now constrained more so by available memory, and not by raw floating point performance.

The natural successors to N-body simulations, which include only the self-gravitating evolution of the dark matter component, are cosmological hydrodynamical simulations,

which model the coupled evolution of dark matter and cosmic gas. Outside of astronomy, some of the earliest examples of such computational fluid dynamics (CFD) simulations were weather related and performed in the 1940s on the ENIAC machine, using a finite difference approach advocated by Lewis Richardson in the 1920s (Hunt, 1998). Shortly thereafter, VonNeumann & Richtmyer (1950) introduced a method to handle the presence of discontinuous shocks in grid solutions of the Euler and Navier-Stokes equations, ushering in the modern era for numerical studies of fluid dynamics.

The Lax-Friedrichs method was developed as a way to estimate the fluxes across interfaces, while Godunov (1959) described a scheme of the same name whereby finite volume fluxes are calculated as the solution of Riemann problems on the boundary of each face between neighboring cells. Godunov's method was extended from first to second order in space by van Leer with the Monotone Upstream-centered Schemes for Conservation Laws (MUSCL; van Leer, 1979), who also pioneered the use of limiters to prevent spurious numerical oscillations. The piece-wise parabolic method (PPM; Colella & Woodward, 1984) allows a third order accurate conservative scheme. All such methods and their successors, which operate through the discretization of space into disjoint sub-volumes, are broadly referred to as grid-based, or Eulerian. They provide a description of the flow field by focusing on specific locations in space through which the fluid flows with time. They were, however, not the first methods used in astrophysics.

The alternative to an Eulerian formulation is a Lagrangian approach, whereby the flow field is described by following individual fluid parcels as they move through space. Smoothed particle hydrodynamics (SPH), developed by Lucy (1977) and Gingold & Monaghan (1977), takes this starting point, discretizing not the volume but instead the mass, into a finite number of individual gas particles. In the 1990s SPH methods were

coupled to the tree approach for gravity and used to study cosmological hydrodynamics for the first time (Hernquist & Katz, 1989a; Katz & Gunn, 1991; Katz et al., 1992; Navarro & White, 1993). Their principal attraction is a natural adaptivity in space and time – since gas particles naturally follow the flow and concentrate in high density regions, numerical resolution is maximized inside of collapsed structures. SPH methods can simultaneously conserve total energy, linear and angular momentum, entropy, and of course mass, and there are no advection errors as in bulk flow across a fixed grid. Finally, they can be coupled in a straight forward and accurate manner with many methods for calculating self-gravity. The size of the largest gas particle simulations in cosmological volumes trails their N-body relatives, nowadays by a rather constant factor, flagship projects having reached $N = 10^9$ (Gottlöber & Yepes, 2007) and $N = 10^{10}$ (Khandai et al., 2014) gas elements.

Grid methods for cosmological and galaxy formation problems have also become common. Early efforts focused on the properties of the low density intergalactic medium (Cen et al., 1994), including the process of cosmic reionization (Norman et al., 1998). The escape from uniform grids through the use of adaptive mesh refinement (AMR) techniques (e.g. Teyssier, 2002) enabled dynamic adaptivity in space. The principal attraction of grid-based techniques is a high order, accurate representation of the fluid field, including excellent capture of discontinuous shocks, fluid instabilities, and phase boundaries. These simulations now reach similar sizes as their SPH counterparts, the Horizon MareNostrum (Ocvirk et al., 2008) and more recent Horizon-AGN (Dubois et al., 2014) projects encompassing a volume with side length of order 100 Mpc with resolution reaching $\simeq 1$ kpc. This combination of scales is a frequently adopted and pragmatic choice which aims to minimally resolve the internal structure of individual

galaxies while including a sufficient volume to obtain statistically robust results across a wide halo mass range.

Various comparisons between grid and particle methods (Kang et al., 1994; Frenk et al., 1999; O'Shea et al., 2005) and idealized explorations of their fundamental differences (Tasker et al., 2008; Sijacki et al., 2012) have been made. Various deficiencies in each have been noted – the tendency of galaxy orientations to numerically align with the Cartesian grid axes in AMR simulations for instance (Hahn et al., 2010), or the inability of SPH to properly mix colder gas blobs moving through a hot ambient medium (Agertz et al., 2007). The influence of numerical issues in the interpretation of cosmological and galaxy-scale phenomena can be quite severe. We explore such a case in the second chapter of this thesis, in the context of the feeding of galaxies through hot versus cold mode gas accretion processes.

1.2 Physical models and numerical approaches

In this section we aim to introduce the principal physical models and the numerical methods used to approximate their solution in the remaining chapters of this thesis. As a comprehensive description of all the relevant aspects would be prohibitive, we present only the foundations.

1.2.1 Gravity

First and foremost is gravity. The matter density of the universe is dominated by dark matter, a presumably large ensemble of microscopic, non-baryonic particles as of yet

unidentified. We model this component under the continuum approximation of the Poisson-Vlasov (or collisionless Boltzmann) equation, given by

$$\frac{\partial f}{\partial t} + \boldsymbol{v}\frac{\partial f}{\partial \boldsymbol{x}} - \frac{\partial \Phi}{\partial \boldsymbol{x}}\frac{\partial f}{\partial \boldsymbol{v}} = 0,$$

where $f = f(\boldsymbol{x}, \boldsymbol{v}, t)$ is the six dimensional distribution function, $\boldsymbol{x}$ denotes spatial position, $\boldsymbol{v}$ spatial velocity, and Φ the scalar gravitational potential, which is given by

$$\nabla^2\Phi(\boldsymbol{x}, t) = 4\pi G\rho(\boldsymbol{x}, t) = 4\pi G m \int f(\boldsymbol{x}, \boldsymbol{v}, t) d\boldsymbol{v}.$$

Here m represents the mass of a tracer particle, although this has absolutely no relationship to the mass of the underlying dark matter particle. It is introduced to solve the continuum equation through the N-body approach. For a given density distribution, then gravitational field then gives rise to the acceleration

$$\ddot{\boldsymbol{x}}_i = -\nabla_i\Phi(\boldsymbol{x}_i)$$

at the tracer particle at position $\boldsymbol{x}_i$. As the potential at a position $\boldsymbol{x}$ due to an ensemble of N point masses m_i at positions $\boldsymbol{x}_i$ is the superposition of each $\phi_i = -Gm_i/(\boldsymbol{x} - \boldsymbol{x}_i)$ we have

$$\Phi(\boldsymbol{x}) = \sum_{i=1}^{N} \phi_i = -G\sum_{i=1}^{N} m_i \left[(\boldsymbol{x} - \boldsymbol{x}_i)^2 + \epsilon^2\right]^{-1/2}.$$

In the second expression, the softening length ϵ has been introduced. Its primary purpose is to avoid the singular behavior of $\Phi \to \infty$ as $\boldsymbol{x} - \boldsymbol{x}_i \to 0$ which results in the large accelerations, with correspondingly small time steps, arising from close encounters. As ϵ represents the smallest spatial scale on which the collisionless dynamics are resolved,

increasing this value decreases resolution as well as computational cost. It cannot however be reduced arbitrarily to achieve higher precision, as spurious two-body effects of the massive tracer particles must be avoided.

There are several methods to calculate the gravitational potential itself. By straightforward evaluation of the sum on the right-hand side, we can *exactly* compute Φ through direct summation over all particle pairs, resulting in a $\mathcal{O}(N^2)$ scaling. All other techniques are simply less accurate but faster approximations. We can deposit the tracer particle mass into a grid representation, and use either Fourier techniques to solve for the potential in $\boldsymbol{k}$ space (Klypin & Shandarin, 1983; White et al., 1983), or iterative methods to solve Poisson's equation directly (Teyssier, 2002). Alternatively, we can use hierarchical multipole (tree) expansions, replacing long range particle-particle interactions with particle-cell (Barnes & Hut, 1986) or cell-cell (Greengard & Rokhlin, 1987) interactions which account for groups of particles at once. Hybrid combinations of the two such as TreePM are often used in practice (Xu, 1995; Springel, 2005).

In the first case, for particle mesh (PM) techniques, any reasonable, conservative interpolation method can be used to assign masses to each grid cell from the scattered tracer point set. At zeroth order, the mass of each particle can be assigned to its nearest grid cell (NGC). At first order, each particle can be assigned some characteristic size (constant or adaptive) and its mass distributed over all overlapping cells (as in CIC, cloud in cell). The mass can be distributed not in a top-hat but with a kernel of a given shape, such as triangular (TSC) or the usual SPH cubic spline, and so on. The same method can be later used to map the gravitational potential, force, and so acceleration from the grid back to the position of each tracer particle.

Alternatively, for Fourier methods we seek a solution for the potential in the frequency domain, given the Fourier transform of the density field,

$$\hat{\Phi}(\boldsymbol{k}) = -4\pi G \frac{\hat{\rho}(\boldsymbol{k})}{k^2}.$$

Here $\hat{q}(\boldsymbol{k})$ denotes the Fourier transform of the real space $q(\boldsymbol{x})$, computed in practice using the discrete, Fast Fourier Transform (FFT; Cooley & Tukey, 1965) which has a scaling of $\mathcal{O}(N \log N)$. Alternatively Poisson's equation can be solved in real space for the N tracers by casting it terms of a linear system of equations $\boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{b}$ where $\boldsymbol{x} = (\Phi_i) \in \mathcal{R}^n$ and $\boldsymbol{b} = 4\pi G \rho_i / \Delta_x^2$ assuming a uniform grid with spacing Δx. The matrix $\boldsymbol{A} \in \mathcal{R}^n \times \mathcal{R}^n$ couples mesh points through a differencing scheme to approximate ∇^2. Approximate, iterative solutions to this system including Jacobi or Gauss-Seidel iteration, or multigrid methods (first approximating the solution on a coarser mesh) can be used. In each case the cost scaling is the same as with the FFT approach, $\mathcal{O}(N \log N)$ where N is the number of grid points. For both Fourier and real space methods the grid need not be uniform and can have adaptive mesh cell size in different regions of space.

In contrast, tree methods are based on a hierarchical grouping of tracer particles. The multipole moments for each are then computed, up to some order: monopole or quadrupole are both common choices. Any reasonable, adaptive tree structure can be used: oct-trees and kd-trees are both common choices. The error of the method, and so also the inverse of the computational cost, is controlled by a parameter θ_c – the tree opening angle criterion. The geometrical interpretation of θ is the angle a given tree node subtends from the point of view of a position $\boldsymbol{x}$ where we seek to compute the gravitational potential – roughly, the node size normalized by its distance. The force is computed at each particle by recursively considering each tree branch, starting at the

root node. If $\theta < \theta_c$ then the expansion is used and the branch descent terminated. Note that if the tree is opened down to its leaf nodes, which contain only a single tracer, then the expansion is not used and the force computation is exact. Therefore for $\theta_c \to 0$ we recover the direct summation method. The expected number of nodes which will need to be opened scales as $N_{\text{nodes}} \propto \log N/\theta_c^3$ (Hernquist, 1987), so the overall cost of the gravity computation is again reduced to $\mathcal{O}(N \log N)$.

Finally, three general aspects remain for a gravity only simulation. First, in cosmological volumes, a comoving coordinate system in $\boldsymbol{x'}$ is typically used instead of a physical coordinate system in $\boldsymbol{x}$. The two are related by the time evolving scale factor as $\boldsymbol{x'} = a(t)\boldsymbol{x}$, where the evolution of $a(t)$ is governed by the background cosmology. Second, cosmological simulations model a periodic volume assumed to be embedded in an expanding, infinite space. The equations of motion and the calculation of the gravitational potential in comoving coordinates for periodic boundary conditions must be suitably modified (Springel et al., 2001b). Finally, the system must be advanced in time. After spatial discretization of a PDE system we are left with a first-order ODE system of the form $d\boldsymbol{q}/dt = \boldsymbol{f}(\boldsymbol{q}, t)$ for a solution vector q^n with respect to the time variable t at the time step t^n, which is updated in discrete steps from q^n to q^{n+1}. Common choices include second order Runge Kutta (predictor-corrector) schemes, or second order (symplectic) leapfrog. The latter is time-reversible and nicely avoids any secular growth of the energy error, but presents difficulties in terms of computational efficiency after inclusion of our particular hydrodynamic solver (Pakmor et al., 2015).

1.2.2 Hydrodynamics

Dark matter is the dominant component of the universe by mass, and dominates the gravitational forces and dynamics on large scales. However, virtually all empirical evidence arising out of observational astronomy probes instead the two most important baryonic components, gas and stars, through measurements of their radiative outputs. As stars form through the gravitational collapse of gas clouds, our principal goal is the modeling of the relevant gas dynamics. In the low viscosity limit, appropriate for both the cosmological and galactic regimes considered in this thesis, we model gas dynamics using the Euler equations. This is a nonlinear, hyperbolic conservation law system derived by considering the conservation of mass, momentum, and energy. It is given by

$$\begin{aligned} \frac{\partial \rho}{\partial t} + \nabla(\rho \boldsymbol{u}) &= 0, \\ \frac{\partial}{\partial t} + \nabla(\rho \boldsymbol{u}\boldsymbol{u}^{\mathrm{T}} + p) &= 0, \\ \frac{\partial E}{\partial t} + \nabla[(E+p)\boldsymbol{u}] &= 0, \end{aligned}$$

where ρ denotes density, $\boldsymbol{u}$ velocity, p pressure, and E total energy. With three equations and four unknowns a closure is required and comes from the ideal gas law

$$p = (\gamma - 1)\left(E - \frac{1}{2}\rho \boldsymbol{u}^2\right),$$

where we take $\gamma = 5/3$ for a mono-atomic gas. The solution of this system for any non-trivial configuration can be complex, and has been a fundamental focus of the computational fluid dynamics field for many decades. It must be discretized in the time variable t, but more critically in space, over some domain Ω, a task for which many

approaches exist. In astrophysics two are commonly used, both relevant for the present study – the finite volume method (FVM), and smoothed particle hydrodynamics (SPH).

The basis of the finite volume scheme is as follows (Toro, 1999; LeVeque, 2002). In one dimension, for some conserved scalar quantity q we can write, in integral form,

$$\frac{d}{dt}\int_{x_1}^{x_2} q(x,t)dx = f(q(x_1,t)) - f(q(x_2,t))$$

with $x_1, x_2 \in \mathcal{R}$ the endpoints of some interval, and $f(q)$ some flux function. The interpretation of the above equation is that the change in time of q in the interval equals the amount that "flows in" minus the amount that "flows out". The equivalent differential form is

$$q_t(x,t) + f(q(x,t))_x = 0,$$

with partial derivatives indicated by subscripts. We discretize the domain into a number of disjoint cells (that is, finite volumes) C_i and store for each the cell average of the quantity $q(x)$ at timestep t^n in that cell

$$Q_i^n = \frac{1}{\Delta x}\int_{C_i} q(x,t^n)dx.$$

The goal is then to update the quantity Q_i^n for all cells, based on the fluxes into and out of C_i, thereby arriving at Q_i^{n+1}. Integrating the integral form from t_n to t_{n+1} and normalizing by $\Delta x = x_{i+1/2} - x_{i-1/2}$ (assuming equally spaced cells) gives

$$Q_i^{n+1} = Q_i^n - \frac{1}{\Delta x}\left[\int_{t_n}^{t_{n+1}} f(q(x_{i+1/2},t))dt - \int_{t_n}^{t_{n+1}} f(q(x_{i-1/2},t))dt\right].$$

The right-hand side cannot be computed because we do not know a priori the evolution

of $q(x_{i\pm1/2}, t)$ for $t \in [t_n, t_{n+1}]$. If, however, we have some approximation of these *fluxes* $F^n_{i\pm1/2}$ averaged across the time interval

$$F^n_{i\pm1/2} \simeq \frac{1}{\Delta t}\int_{t_n}^{t_{n+1}} f(q(x_{i\pm1/2}, t))dt$$

then we arrive at the update method of the finite volume scheme:

$$Q^{n+1}_i = Q^n_i - \frac{\Delta t}{\Delta x}\left(F^n_{i+1/2} - F^n_{i-1/2}\right) \quad ; \quad i = 1, 2, ..., N_i.$$

This formulation is clearly conservative, since the flux lost from one cell is gained by its neighbor. The sum $\sum_{i=1}^{N_i} \Delta x Q^{n+1}_i$ equals the fluxes at the boundary of the domain, or in the case of closed or periodic boundary conditions, zero. We can also see that if we interpret Q^n_i not as a cell average, but instead as a grid value, then this method is equivalent to the usual first-order finite difference (FD) scheme with grid spacing Δx in the space dimension and Δt in the time dimension. However, the FVM has two key advantages over FD: it can be generalized to cells of arbitrary shape and size, and so to adaptive and unstructured meshes, and it explicitly enforces conservation of the cell average quantities.

The FVM concept extends naturally from a scalar conservation law in one dimension to a system of m conservation laws in $n \geq 2$ dimensions. Now, $q = (q_1, q_2, ..., q_m)$ is a vector instead of a scalar, and $f \in \mathcal{R}^n$ is the flux vector. For the Euler equations in 3D, we have $q = (\rho, \rho u_1, \rho u_2, \rho u_3, E)$ with three momentum components. The cell averages $Q^n_i \in \mathcal{R}^m$ are then also vectors, and their change within the volume C_i equals the integral of the divergence, $\int \boldsymbol{\nabla} \cdot \boldsymbol{f}(q(x,t))d\boldsymbol{x}$. Using the divergence theorem, this volume integral is transformed into a surface integral over the cell boundary δC_i, resulting in the update

$$Q_i^{n+1} = Q_i^n + \Delta t \int_{\delta C_i} \boldsymbol{f}(q(s,t)) \cdot \boldsymbol{n}(s) ds$$

where $\boldsymbol{n}(s)$ is the outward normal vector on the cell surface. In practice, this integral is split into a sum over the finite number of geometric faces of each cell.

A fundamentally different and rather prevalent approach to solving such a conservation law system is smoothed particle hydrodynamics (Monaghan, 1992, 2005; Springel, 2010b). Rather than discretize in space, we discretize in mass, into a finite number of tracer particles which are followed in a Lagrangian sense. The method at its heart is an interpolation scheme, whereby for each continuous property of the fluid field $Q(\boldsymbol{x})$ a smoothed representation $Q^\star(\boldsymbol{x})$ is defined as

$$Q^\star(\boldsymbol{x}) = \int_\Omega Q(\boldsymbol{x'}) W(\boldsymbol{x} - \boldsymbol{x'}) d\boldsymbol{x'}.$$

The integral is over the entire domain Ω, and the convolution (or smoothing) kernel is W. In practice, W is chosen to have finite extent with some characteristic size h, and this integral is computed as a weighted sum over the N nearest sample points within that region

$$Q^\star(\boldsymbol{x}) \simeq \sum_{i=1}^{N} V_i Q_i W(\boldsymbol{x} - \boldsymbol{x}_i; h).$$

Here Q_i is the value of the field for each sample point, while the weight $V_i = m_i/\rho_i$ is the adaptive volume associated to each point. This estimate $Q^\star$ is defined everywhere in space and is also differentiable, so long as $W(\boldsymbol{x})$ is likewise. Equations of motion corresponding to the gas dynamics can then be derived from these interpolated values. The modern approach proceeds from the Lagrangian corresponding to the Euler equations for an ideal gas (Springel & Hernquist, 2002):

$$L = \int \rho \left(\frac{1}{2} v^2 - u_{\rm therm} \right) dV = \sum_{i=1}^{N} \left(\frac{1}{2} m_i v_i^2 - m_i u_i \right).$$

The second equality gives the discretized form for N particles with masses, velocity, and internal energies per unit pass of m_i, v_i, and u_i, respectively. Application of the Euler-Lagrange equation,

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{\boldsymbol{x}}_i} - \frac{\partial L}{\partial \boldsymbol{x}_i} = 0,$$

where the first term is the convective derivative, leads to the equation of motion for each particle ($\ddot{\boldsymbol{x}}_i$) which can then be updated forward in time. In detail, as with the FVM, many choices and additional numerical considerations can apply, although the method is remarkably simple, depending only on interpolated estimates of density, pressure, velocity, and their derivatives. In general, the existence of multiple, fundamentally different numerical approaches, including more advanced and hybrid techniques, enables a powerful consistency check for any scientific conclusions we seek to draw from simulations.

In this thesis we employ a relatively new simulation code, Arepo, which is described extensively in Springel (2010a). It makes use of a hybrid TreePM scheme for self-gravity, with long range forces calculated through a Fourier particle-mesh method, and short range forces with a hierarchical tree algorithm. Gravity is coupled to a novel "moving mesh" technique for solving the equations of continuum hydrodynamics. This technique can be viewed as a hybrid between the FVM and SPH methods, aiming to inherit the strengths of both while simultaneously avoiding their deficiencies. The spatial discretization of the fluid is provided through an unstructured Voronoi tessellation which is dynamic in time. In all investigations in the present work, Godunov's

method is employed on the volumes defined by each cell, with a directionally unsplit MUSCL-Hancock scheme and an iterative, exact Riemann solver. Therefore the input fluid states to the Riemann problem, on each side of each face, are derived not only with cell average values as discussed above, but also with linear gradients in each cell, providing second order accuracy in space. This fluid state is boosted into the moving frame of reference of each Voronoi face, and also interpolated to its time centered ($t^{n+1/2}$) value, whereby a leapfrog integrator with hierarchical, adaptive time stepping provides second order accuracy in time. The Voronoi mesh is generated uniquely from a set of control points which move with the local fluid velocity, modulo mesh regularization corrections. We employ throughout a new Monte Carlo tracer particle scheme (described in Genel et al., 2013) to follow the Lagrangian evolution of baryons.

1.3 Making a model universe

A few crucial ingredients are still missing if we hope to conduct a numerical investigation which has some correspondence to the actual universe. Perhaps the most important is a starting point – the initial conditions. Cosmological simulations have the advantage that the properties of the early universe, in the statistical sense, are now known to remarkable precision. In particular, the WMAP and PLANCK missions (Spergel et al., 2003; Planck Collaboration et al., 2014) have observed the cosmic microwave background (CMB) radiation in exquisite detail, measuring the properties of temperature fluctuations just a few hundred thousand years after the big bang. They also constrain what is now considered the standard model of cosmology; namely, a flat ($\Omega_{\rm tot} = 1$), Λ-dominated universe seeded by gaussian fluctuations and with relative mass-energy contents of

$\Omega_\Lambda \sim 0.7$, $\Omega_m \sim 0.3$, and $\Omega_b \sim 0.04$ for dark energy, matter, and baryons, respectively. Using well-understood linear perturbation theory (Zel'dovich, 1970) we can evolve the initial field of density fluctuations to the beginning of the non-linear regime, typically at a redshift of $z \sim 100$, and thereby construct the initial conditions for either a hydrodynamical or N-body cosmological simulation.

If we do so in a volume with side-length $L = 20\mathrm{Mpc}/h$ and containing only a toy number of 16^3 gas cells using the Arepo code and pure gravitational collapse with non-radiative gas physics, the final $z = 0$ evolved structure of the three-dimensional Voronoi tessellation is shown in Figure 1.1. As the volume of each Voronoi cell is inversely proportional to density, we can already see hints of the large-scale structure of the universe in voids, filaments, and collapsed structures, through a dynamic range in spatial scale as large as the simulation allows.

Zeroth order study of the galaxies which form in the centers of these collapsed halos requires three further ingredients: (i) gas cooling, (ii) star formation, and (iii) feedback. At cosmological and galactic-scale densities, two-body processes dominate the radiative cooling of a primordial or metal-enriched gas. These include collisional excitation, collisional ionization, recombination, free-free emission at high temperatures, and Compton cooling off the CMB. These processes are included implicitly as an additional term in the thermal energy update equation, $du/dt \mathrel{+}= \Lambda_{\rm net}/\rho$ where $\Lambda_{\rm net} = \Gamma - \Lambda$ is the net, heating minus cooling rate. For a primordial gas, a rather simple chemical network including only hydrogen and helium (H^0, H^+, He^0, He^+, He^{++}) is calculated explicitly following Katz et al. (1996). Ionization equilibrium with a meta-galactic radiation field is assumed in the optically thin regime – energy emitted as cooling radiation is simply lost in the simulation. Contributions to $\Lambda_{\rm net}$ from metal-line cooling can be optionally

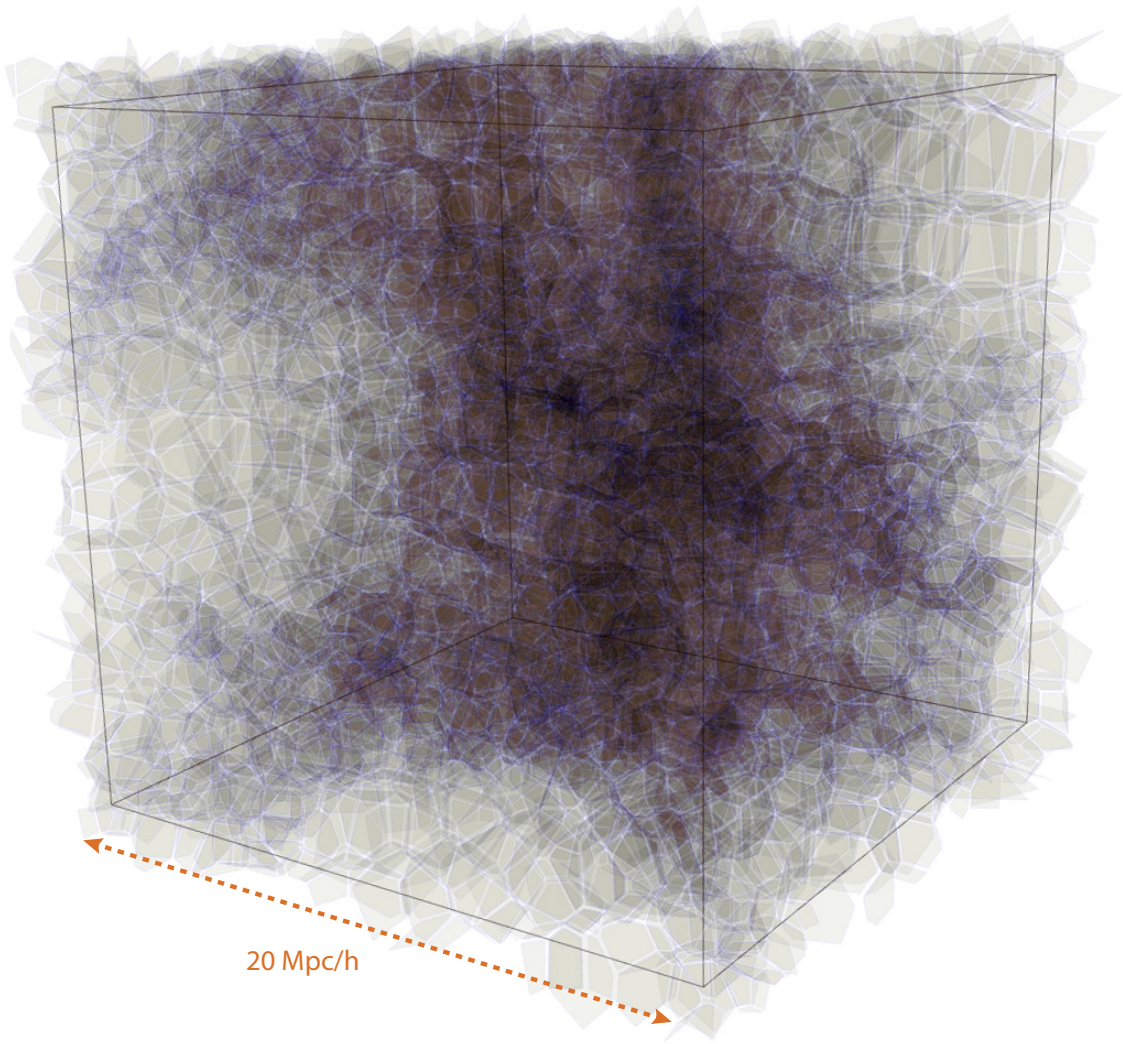

Figure 1.1 Geometry of the 3D Voronoi tessellation at the present day, $z = 0$, in a periodic volume with side length equal to 20Mpc/h. This example contains only 16^3 gas cells for visual clarity, but is initialized with the same type of primordial density fluctuations as the Illustris simulation with 1820^3 cells.

incorporated following Sutherland & Dopita (1993) and Wiersma et al. (2009a), with non-negligible effects around the temperature of virialized hot halo gas. Chapters 2 and 4 of this thesis include only cooling from a primordial gas, while chapters 3 and 5 discuss simulations including metals. Below a temperature of $\sim 10^4\,\mathrm{K}$ molecular and metal fine-structure cooling channels dominate, which are not considered.

For a sufficiently cool, dense, self-gravitating mass of gas, star formation can proceed. For cosmological simulations, the entirety of this process occurs on spatial scales smaller than the resolution limit of the gas. A sub-resolution model is therefore required, and we adopt throughout the method of Springel & Hernquist (2003). In particular, fluid parcels representing the ISM are pictured as containing a mixture of colder condensed clouds in pressure equilibrium with a hot ambient gas, following the theoretical picture of McKee & Ostriker (1977). Star formation is assumed to convert the cold phase into stars above some density threshold and on a characteristic time scale $t_\star$ which is proportional to the local dynamical time, such that $t_\star(\rho) = t_0^\star(\rho/\rho_{\rm th})^{-1/2}$. This results in a Schmidt-Kennicutt type star formation law, with $\rho_{\rm th}$ a free parameter ($t_0^\star$ is fixed for a given value of the threshold). In practice, given the instantaneous star formation rates $\dot{M}_\star$ of each gas cell, collisionless star particles are formed probabilistically over the subsequent time step such that the correct amount of stellar mass is created in the time average.

Once stars form, they inject energy as well as heavy elements into their surroundings. We include this energetic *feedback* in two fundamentally different ways: unresolved and resolved. By the latter we mean the resolved generation of galactic-scale outflows driven by star formation feedback, as well as the energy injection from supermassive black holes. As with metal contributions to cooling, resolved feedback is considered in chapters 3 and

5 of this thesis, where the fiducial physics model of the Illustris simulations is adopted (Vogelsberger et al., 2013). That model is described in more detail therein. In chapters 2 and 4 we restrict our study to the "no-feedback" case, where only unresolved energy input from supernovae is included, again following Springel & Hernquist (2003). In this case, each newly formed stellar population will have some number of massive stars which rapidly die as supernovae, the amount depending on the adopted initial mass function. A canonical amount of energy per SN is injected back into the gas, as a heating term for the hot phase and an evaporative term for the cold phase of each ISM gas cell. This establishes local self-regulation of SF, preventing runaway collapse through an effective pressure, increasing as a function of gas density, and arising from supernova feedback.

Such a minimally comprehensive simulation including gas dynamics and these relevant baryonic processes, which is initialized with cosmological initial conditions, can be used to address a number of questions. Does it form a population of galaxies, groups, and clusters in resemblance to what is seen in the observed universe? If so, how do these objects form in the early universe, and how do they evolve across cosmic time? What is their relation to the underlying dark matter structures? What physical processes dominate and in what regimes? What underlying properties correspond to observed galaxy characteristics, and how are the two related?

The focus of this thesis is the baryonic growth of galaxies through gas accretion: how does gas from the intergalactic medium fall into a galaxy and ultimately contribute to the production of stars (Rees & Ostriker, 1977; Silk, 1977)? What is the thermal and dynamical history of this accreting gas (Katz et al., 2003; Kereš et al., 2005)? How do the quasi-static reservoirs of virialized gas in dark matter halos form, and how do they cool (Birnboim & Dekel, 2003a; White & Rees, 1978)? How does inflowing gas interact

with this hot atmosphere (Dekel & Birnboim, 2006; Dekel et al., 2009)? Do feedback processes fundamentally alter the nature or rates of accretion (van de Voort et al., 2011b; Faucher-Giguère et al., 2011), and how is the regulation of star formation tied to the baryon cycle of inflowing and outflowing material, across redshift and as a function of halo mass (Oppenheimer et al., 2010; van de Voort et al., 2011a)? How can we connect to observations of the hydrogen and metal content in and around galaxy halos (e.g. Hennawi et al., 2006; Rudie et al., 2012; Prochaska et al., 2013)? In the following four chapters we tackle some of these points, and briefly introduce each investigation below.

1.3.1 Tracing cosmological gas accretion

In Chapter 2 of this thesis we investigate the nature of gas accretion onto halos and galaxies at $z = 2$ using cosmological hydrodynamic simulations. We make a quantitative comparison between identical simulations run with the moving mesh code Arepo and the smoothed particle hydrodynamics (SPH) code Gadget-3. By doing so we re-evaluate past results on the modes and rates of gas accretion, which have predominantly been quantified through simulations run with the SPH technique. To enable this comparison, we implement a Monte Carlo tracer particle scheme to determine the origin and thermodynamic history of accreting gas.

Contrasting these two numerical approaches, we find significant physical differences in massive halos above $\sim 10^{10}\,\mathrm{M}_{\odot}$. In agreement with previous work, Gadget simulations indicate that only a small amount of accreted gas heats to an appreciable fraction of the virial temperature during accretion. The same galaxies in Arepo show a significantly lower "cold fraction", due to a hot gas accretion rate which can be an order

of magnitude larger, while the cold accretion rate is also lower by a factor of two. These discrepancies increase for more massive systems, and we explain both trends in terms of numerical inaccuracies with the standard formulation of SPH.

By considering the spatial distribution of gas in galaxy halos we verify that the filamentary geometry of accreting gas near the virial radius is a common feature in massive halos above $\simeq 10^{11.5}\,\mathrm{M}_\odot$. Gas filaments in GADGET, however, tend to remain collimated and flow coherently to small radii, or artificially fragment and form a large number of purely numerical "blobs". These same filamentary gas streams in AREPO show increased heating and disruption within half the virial radius. We discuss the implications for prior conclusions in the literature about the relative importance of "hot" versus "cold" mode accretion.

1.3.2 The impact of feedback

In Chapter 3 of this thesis we extend our earlier simulations with a comprehensive physical model for baryonic feedback processes. We contrast two suites of simulations – with and without feedback. The feedback runs implement the full physics model of the Illustris simulation project, including star formation driven galactic winds and energetic feedback from supermassive black holes.

We explore the accretion rate of material contributing to the net growth of galaxies and originating directly from the intergalactic medium, finding that feedback strongly suppresses the raw, as well as the net, inflow of this "smooth mode" gas at all redshifts, regardless of the temperature history of newly acquired gas. At the virial radius the temperature and radial flux of inflowing gas is largely unaffected at $z = 2$. However,

the spherical covering fraction of inflowing gas at $0.25\,r_{\rm vir}$ decreases substantially, from more than 80% to less than 50%, while the rates of both inflow and outflow increase, indicative of recycling across this boundary.

We measure the amount of time taken by gas to cross from the virial radius to the galaxy – the "halo transit time". This timescale increases in the presence of feedback by a factor of $\simeq 2$ - 3, and is notably independent of halo mass. We discuss the possible implications of this invariance for theoretical models of hot halo gas cooling.

1.3.3 Zooming in: the structure of halo gas

In Chapter 4 of this thesis we return to a simplified physical model, without resolved feedback, and study the properties of gas in and around 10^{12} $M_\odot$ halos at $z = 2$ using a suite of cosmological hydrodynamic 'zoom' simulations. By focusing on single objects instead of the large volumes of the previous two chapters we obtain significantly higher spatial resolution, the gas cell mass reaching ∼10,000 solar masses.

We quantify the thermal and dynamical structure of these gaseous reservoirs in terms of their mean radial distributions and angular variability along different sightlines. Studying the interaction of filamentary inflow and the quasi-static hot halo atmosphere we highlight the discrepancy between the spatial resolution available in the halo gas as opposed to within the galaxy itself, and find that stream morphologies become increasingly complex at higher resolution, with large coherent flows revealing density and temperature structure at progressively smaller scales. Moreover, multiple gas components co-exist at the same radius within the halo, making radially averaged analyses misleading.

We investigate the process of gas virialization and identify different regimes for the heating of gas as it accretes from the intergalactic medium. Halos at this mass have a well-defined virial shock, associated with a sharp jump in temperature and entropy at $\gtrsim 1.25\, r_{\rm vir}$. The presence, radius, and radial width of this boundary feature, however, vary not only from halo to halo, but also as a function of angular direction. On average a strong virial shock covers roughly $\sim 85\%$ of the 4π sphere.

Our findings are relevant for the proper interpretation of observations pertaining to the circumgalactic medium, and we discuss them in terms of the evidence for large amounts of cold gas surrounding massive halos at intermediate redshifts.

1.3.4 The Illustris public data release

In Chapter 5 of this thesis we shift focus from our studies of gas accretion to a technical project, the goal of which is to enable community-wide comparisons between state of the art theoretical models for galaxy formation and current and upcoming observations of galaxy populations across cosmic time.

In particular, we present the full public release of all data from the Illustris simulation project. As previously presented, Illustris is a suite of large volume, cosmological hydrodynamical simulations run with the moving-mesh code Arepo and including a comprehensive set of physical models critical for following the formation and evolution of galaxies across cosmic time. Each simulates a volume of $(106.5\ \mathrm{Mpc})^3$ from a starting redshift of $z = 127$ to the present day, $z = 0$, and self-consistently evolves five different types of resolution elements: dark matter, gas, gas tracers, stars, and supermassive black holes.

The data release includes the snapshots at all available redshifts, halo and subhalo catalogs at each snapshot, and two distinct merger trees. Six primary realizations of the Illustris volume are released, including the flagship Illustris-1 run. These include three resolution levels with the fiducial "full" baryonic physics model, and a dark matter only analog for each.

We describe the released data products as well as the comprehensive interface and set of tools we have developed for their analysis. All data may be directly downloaded online, and we also release a comprehensive, web-based API which allows programmatic access to search and data processing tasks. Finally, we address scientific issues relevant for the interpretation of the simulations, comment on technical aspects of the release, and discuss future directions.

Chapter 2

Moving mesh cosmology: tracing cosmological gas accretion

D. Nelson, M. Vogelsberger, S. Genel, D. Sijacki, D.Kereš, V. Springel, L. Hernquist, *Monthly Notices of the Royal Astronomical Society*, Vol 429, Issue 4, p. 3353-3370, 2013

Abstract

We investigate the nature of gas accretion onto halos and galaxies at $z = 2$ using cosmological hydrodynamic simulations run with the moving mesh code AREPO. Implementing a Monte Carlo tracer particle scheme to determine the origin and thermodynamic history of accreting gas, we make quantitative comparisons to an otherwise identical simulation run with the smoothed particle hydrodynamics (SPH)

code GADGET-3. Contrasting these two numerical approaches, we find significant physical differences in the thermodynamic history of accreted gas in massive halos above $\simeq 10^{10.5}\,\mathrm{M}_{\odot}$. In agreement with previous work, GADGET simulations show a cold fraction near unity for galaxies forming in massive halos, implying that only a small percentage of accreted gas heats to an appreciable fraction of the virial temperature during accretion. The same galaxies in AREPO show a much lower cold fraction, for instance $< 20\%$ in halos with $M_{\mathrm{halo}} \simeq 10^{11}\,\mathrm{M}_{\odot}$. This results from a hot gas accretion rate which, at this same halo mass, is an order of magnitude larger than with GADGET, together with a cold accretion rate which is lower by a factor of two. These discrepancies increase for more massive systems, and we explain both trends in terms of numerical inaccuracies with the standard formulation of SPH. We note, however, that changes in the treatment of ISM physics – feedback, in particular – could modify the observed differences between codes as well as the relative importance of different accretion modes. We explore these differences by evaluating several ways of measuring a cold mode of accretion. As in previous work, the maximum past temperature of gas is compared to either a constant threshold value or some fraction of the virial temperature of each parent halo. We find that the relatively sharp transition from cold to hot mode dominated accretion at halo masses of $\simeq 10^{11}\,\mathrm{M}_{\odot}$ is a consequence of the constant temperature criterion, which can only separate virialized gas above some minimum halo mass. Examining the spatial distribution of accreting gas, we find that the filamentary geometry of accreting gas near the virial radius is a common feature in massive halos above $\simeq 10^{11.5}\,\mathrm{M}_{\odot}$. Gas filaments in GADGET, however, tend to remain collimated and flow coherently to small radii, or artificially fragment and form a large number of purely numerical "blobs". These same filamentary gas streams in AREPO show increased heating and disruption at 0.25-0.5 r_{vir}

and contribute to the hot gas accretion rate in a manner distinct from classical cooling flows.

2.1 Introduction

The manner in which forming galaxies acquire their baryonic matter through cosmic time in a ΛCDM cosmology is not yet fully understood. In the classic picture of accretion from the intergalactic medium gas shock-heats to the virial temperature of the halo and forms a hot pressure supported atmosphere in approximate equilibrium (Rees & Ostriker, 1977; Silk, 1977; White & Rees, 1978). Gas can then cool onto a centrally forming galaxy, or not, based on the relevant cooling time-scale (White & Frenk, 1991). In this regime the accretion rate onto the galaxy depends on the cooling rate and not on the cosmological accretion rate onto the halo itself. However, this holds true only for sufficiently massive halos. Below some threshold mass the dark matter halo cannot support a stable hot, gaseous atmosphere (Birnboim & Dekel, 2003a) and accreting gas does not necessarily shock heat to the virial temperature. This conclusion clearly motivates the differentiation of a "cold accretion mode" comprised of gas which does not shock heat to the virial temperature of its parent halo during accretion. Numerical simulations over the past decade have shown evidence for a cold accretion mode which dominates both at high redshift and for low mass objects at late times (Katz et al., 2003; Kereš et al., 2005, 2009; Ocvirk et al., 2008). halos have been found to transition from cold to hot mode dominated at a mass in rough agreement with analytical estimates for the stable virial shock cutoff (e.g. Birnboim & Dekel, 2003a).

Cold accretion in low mass objects below this threshold is precisely the regime

where the virial shock is inactive. More massive halos supporting a virial shock are also claimed to experience continuous, active cold mode accretion at high redshift ($z \geq 2$). This cold gas accretes along strongly filamentary geometries, termed "cold flows," which penetrate the virial radius and deliver gas deep within the halo (Kereš et al., 2005; Dekel et al., 2009; Dekel & Birnboim, 2006). In the context of analytical predictions, the non-spherical accretion of cold gas via highly collimated geometries was an unexpected result of these numerical simulations. Although smoothed particle hydrodynamics (SPH) simulations were the first to address the question of gas accretion modes in a cosmological context (Abadi et al., 2003; Kereš et al., 2005) they were soon followed by grid-based adaptive mesh refinement (AMR) efforts (Ocvirk et al., 2008, and others).

Simulations of cosmological volumes using the SPH technique typically make a cold mode gas selection based on the past temperature history of each gas element, invoking the pseudo-Lagrangian nature of the scheme. The resulting maximum past temperature $T_{\rm max}$ value is compared to some temperature threshold (Kereš et al., 2005, 2009; van de Voort et al., 2011b; Faucher-Giguère et al., 2011). This threshold was originally motivated by the location of a minimum in the strongly bimodal global distribution of $T_{\rm max}$ found by Kereš et al. (2005), using the TreeSPH (Hernquist & Katz, 1989b) code. That work compared $T_{\rm max}$ to the virial temperature of halos as well as to a constant temperature threshold, and chose $T_c \simeq 2.5 \times 10^5\,{\rm K}$ as sufficient to identify virial heating in halos, due to the high halo masses considered. Essentially all SPH simulations to date have used this approach, although with less care as to its applicability, as we discuss later. One notable exception is Brooks et al. (2009) which used the SPH code GASOLINE (Wadsley et al., 2004) and identified shocked gas particles based on an entropy jump criterion. This is a novel approach which should be explored in a full

cosmological volume. The authors concluded that their method led to a selection of shocked/unshocked gas which agreed well with the selection of hot/cold gas based on the use of a constant temperature threshold.

More recently, finite volume techniques including cosmological re-simulations using adaptive mesh refinement (AMR) have been used to study the same problem. Grid-based simulations offer a better treatment of shocks, fluid instabilities, and phase boundaries. However for the analysis of cosmological gas accretion they are limited by their Eulerian nature, as the trajectory of any gas parcel cannot be simply identified due to mass exchange between cells. This difficulty can be overcome using tracer particles, and we argue these are a necessary feature of modern cosmological simulations. One AMR work to date (Dubois et al., 2012) used tracer particles in order to study the angular momentum of accreting gas. A common alternate approach is to use the instantaneous gas properties such as mass flux through a surface (e.g. Ocvirk et al., 2008; Dekel et al., 2009) to measure rates, or the low metallicity of gas as an indicator of cosmological origin (Agertz et al., 2009). Unfortunately these techniques cannot constrain the past thermal history of any gas element, nor identify for example material tidally stripped from a galaxy or subhalo during a previous strong interaction. Given these caveats, however, AMR simulations of single halos have also found strong support for the importance of a filamentary, cold accretion mode in massive systems at high redshift.

The contribution of "clumpy" accretion due to infalling satellites and other bound substructures is also difficult to distinguish in Eulerian simulations. Some SPH studies use a stringent definition based on lack of membership to any identified bound substructure at any time prior to accretion (e.g. Brooks et al., 2009). Others only require non-membership at some particular previous time (Kereš et al., 2005; van de

Voort et al., 2011b). In either case, this gas should not be considered as contributing to cold accretion from the intergalactic medium (IGM). This is appropriate since we wish to isolate cosmological gas accretion from the IGM and not hierarchical buildup via mergers. The technical distinction is not particularly important as smooth accretion (defined as all gas not identified as clumpy) is generally found to be the dominant contribution to the total accretion rate (van de Voort et al., 2011b). However, Shen et al. (2013) for example includes a substantial $\simeq$35% contribution of infalling dwarf satellites to the cold accretion mass budget. This effectively combines gas from multiple sources under a single label and ignores the diverse physical processes driving distinct accretion mechanisms. This is especially important in the results we report in what follows using AREPO, for which the accretion rate of cold gas directly from the IGM is small.

Only recently have simulations begun to explore the effects of feedback and galactic scale outflows on the character of gas inflow (Oppenheimer et al., 2010; Faucher-Giguère et al., 2011; van de Voort et al., 2011b). Differences in morphology between different modes of accretion in massive halos and the added complexity from the physics of star formation implies that accretion rates may be quite sensitive to feedback processes. Simulations using SPH have concluded that a filamentary cold mode is largely unaffected by the presence of strong outflows arising from blastwave supernova feedback (Shen et al., 2013). van de Voort et al. (2011a) studied the impact of AGN feedback on inflow and concluded that it preferentially prevented hot mode gas (under the standard definition) from cooling from the halo onto the galaxy, while Dubois et al. (2013) found that AGN feedback morphologically disturbed cold filaments. Current treatments of feedback in cosmological simulations, however, are largely phenomenological and are not self-consistent. Furthermore, the same numerical issues explored in Vogelsberger et al.

(2012), Sijacki et al. (2012), Kereš et al. (2012) and discussed herein that compromise the accuracy of SPH studies of gas inflow will also affect the interaction of outflowing ejecta and wind material with both halo and filamentary gas. Moreover, the accretion rates onto galaxies are significantly higher in simulations with AREPO than with GADGET, implying that feedback effects need to be even more efficient than in previous simulations with SPH.

Strong feedback may also indirectly affect the characterization of inflows. In particular, galactic fountain material and other recycled gas may be a significant additional mode of accretion onto galaxies. Oppenheimer et al. (2010) included a phenomenological galactic wind model in cosmological SPH simulations and concluded that recycled gas accretion is a prominent "third" mode which in fact is the dominant accretion mechanism at $z \leq 1$, with minimal effect on this high redshift accretion with choice of outflow velocity. The ejection of hot gas from galaxies might also physically be expected to add mass and thermal support to a halo atmosphere. The result would be increased heating of cosmologically inflowing gas, an underestimated effect in all simulations lacking galactic scale outflows, including those presented here.

This paper presents results from a comparison of cosmological simulations run with the SPH code GADGET (Springel, 2005) and the moving mesh code AREPO (Springel, 2010a). We compare individual objects as well as statistics derived from cosmological volumes and show that the importance and character of cold gas accretion depends sensitively on both analysis methodology and numerical technique. In Section 4.2 we describe the numerical details of our approach, two complementary tracer particle schemes, and our analysis methodology. Section 2.3 compares the thermodynamics of gas accretion between the two codes, discusses the various definitions of the cold mode

and the sensitivity of our conclusions to those definitions, and investigates the geometry and filamentary nature of inflowing gas. Finally, Section 4.6 wraps up our discussion and summarizes the main conclusions.

2.2 Methods

We use and extend the simulations and numerical methods presented extensively in Vogelsberger et al. (2012) and refer the reader to that paper for additional details. Brief descriptions of the key hydrodynamic differences between the two codes are presented here.

GADGET (last described in Springel, 2005) is a smoothed particle hydrodynamics (SPH) implementation where the Euler equations are solved to model fluid flow by discretizing mass using a set of particles. Continuous fluid quantities are defined by an interpolation kernel over a set of nearest neighbors. This method is pseudo-Lagrangian in nature (see discussion) and is particularly well-suited to handle the large dynamic range present in cosmological simulations where it is desirable to resolve both large scale structure and the internal dynamics of forming galaxies. We use a "standard" SPH formulation in GADGET to facilitate comparison to earlier work (Monaghan, 1992) - the cubic spline kernel with 32 neighbors and the density based formulation (Springel & Hernquist, 2002) modified via an artificial viscosity $\alpha_{\rm SPH} = 1.0$ with the Balsara switch (Balsara, 1995).

AREPO (described in Springel, 2010a) is a finite volume scheme where the control volumes are defined by a Voronoi tessellation of space. Euler's equations are solved

using Godunov's method with the MUSCL-Hancock scheme to compute numerical fluxes and obtain second order accuracy. The tessellation is obtained by a set of mesh generating points which are allowed to move arbitrarily, though in our simulations follow the flow in a quasi-Lagrangian fashion. As a result AREPO retains the principal strengths of SPH including its adaptivity to a large dynamic range in spatial scales, Galilean invariance of the truncation error, and an accurate and efficient gravity solver (e.g. O'Shea et al., 2005). It also gains the strengths of finite volume codes, including the improved treatment of fluid instabilities, weak shocks (which can be missed in SPH; (e.g. Keshet et al., 2003)), phase interfaces, and shearing flows. We use the maximum cell face angle regularization scheme and allow dynamic refinement and de-refinement to maintain approximately constant mass cells (Vogelsberger et al., 2012).

2.2.1 Simulation Set

All our simulations employ a WMAP-7 cosmology ($\Omega_{\Lambda,0} = 0.73$, $\Omega_{m,0} = 0.27$, $\Omega_{b,0} = 0.045$, $\sigma_8 = 0.8$ and $h = 0.7$) with a $20h^{-1}$ Mpc side-length volume at several different numerical resolutions, initially with equal numbers of dark matter particles and gas elements. The particle counts, initial gas element masses, dark matter particle masses, and Plummer equivalent gravitational softening lengths are given in Table (2.1). Unless otherwise noted, all results in this paper compare results between the 512^3 resolution runs at a redshift of $z = 2$.

GADGET and AREPO simulations include identical physics. We account for optically thin radiative cooling assuming a primordial H/He ratio which sets an effective temperature floor for gas in these simulations just below $\sim 10^4$ K (Katz et al., 1996).

Table 2.1: Details on the three different simulation resolutions and relevant numerical parameters. $N_{\rm gas}$ and $N_{\rm DM}$ denote the total number of gas cells and dark matter particles, respectively. $m_{\rm gas}$ indicates either the exact SPH particle mass, or the Voronoi gas cell "target" mass, $m_{\rm DM}$ the dark matter particle mass, and ϵ the gravitational softening length for the collisionless component in comoving units.

$N_{\rm gas}$	$N_{\rm DM}$	Vel Tracers	MC Tracers	$m_{\rm gas}$ $[h^{-1}\,{\rm M}_\odot]$	$m_{\rm DM}$ $[h^{-1}\,{\rm M}_\odot]$	ϵ $[h^{-1}$ kpc]
128^3	128^3	1 x 128^3	10 x 128^3	4.8 x 10^7	2.4 x 10^8	4.0
256^3	256^3	1 x 256^3	10 x 256^3	6.0 x 10^6	3.0 x 10^7	2.0
512^3	512^3	1 x 512^3	10 x 512^3	7.4 x 10^5	3.7 x 10^6	1.0

We do not include metal line cooling. A redshift-dependent ionizing UV background field (Faucher-Giguère et al., 2009) is included as a spatially uniform heating source. Star formation and the associated ISM pressurization from unresolved feedback events are included following Springel & Hernquist (2003). Gas elements are stochastically converted into star particles of a constant mass when the local gas density exceeds a threshold value of $n_{\rm H} = 0.13$ cm^{-3}.

We do not include any explicit forms of strong stellar feedback that would drive galactic-scale winds. There is no treatment of black holes, radiative transfer and its effects, or magnetic fields. The key point is that we use a "standard physics" implementation which has been well studied and probes the interaction of gravity, hydrodynamics, and star formation in the cosmological context. The gravity implementation as well as the chosen sub-grid physics are identical between our SPH and moving mesh simulations, allowing us to attribute differences in the solutions to differences in the method used to solve the equations of hydrodynamics. They are also similar to Kereš et al. (2009), in terms of the code and physics employed as well as the simulation resolution, enabling us

to make a direct comparison to that work.

2.2.2 Tracer Particles

In order to trace the evolution of gas properties over time in our moving mesh calculations we use a new "Monte Carlo" tracer particle technique (see Vogelsberger et al., 2013; Genel et al., 2013). This probabilistic method associates tracers with parent gas cells and exchanges them based explicitly on the mass fluxes through each face. Specifically, all cells start with an equal number of tracers, after which each child tracer transfers from its original cell (i) to a neighboring cell (j) with a probability equal to $\Delta M_{ij}/M_i$, the ratio of the mass flux through face (ij) to the current mass of the originating cell. By construction, the tracer density is guaranteed to follow the underlying fluid density, at the cost of Poisson noise due to the probabilistic nature of the scheme. Furthermore, at each computational timestep each tracer records several fluid quantities of its parent gas cell - for this work we use only the maximum previous temperature and time of that $T_{\rm max}$ event. We initialize each gas cell with 10 Monte Carlo tracers. Due to the target mass refinement scheme the number of child tracers per parent cell remains roughly constant throughout the simulation, with a Gaussian distribution extending a factor of two above and below the initial number. In particular, in galaxies at $z = 2$ the mean number of child tracers is $\simeq$10 with a standard deviation of $\simeq$3.

We also evaluated a "velocity field" tracer particle scheme (Genel et al., 2013). In this more typical approach (e.g. Vazza et al., 2010; Seitenzahl et al., 2010; Dubois et al., 2012), tracers represent massless, passive particles which are purely advected by the local velocity field of the fluid. Our testing shows that this approach exhibits a systematic

bias in its Lagrangian ability to follow mass fluxes. We suggest that the technique does not accurately recover the flow of mass in astrophysical situations involving convergent flows, including cosmological cases where the problem manifests as a "pile-up" of tracers in the centers of dark matter halos. The Monte Carlo approach is unaffected by this bias seen with the velocity field tracer particles. Furthermore, the flux-minimizing moving mesh scheme allows us to use a reasonably small number of such tracers while achieving an acceptable level of noise (Genel et al., 2013). All results presented herein which involve tracking gas properties through time in the AREPO runs use our Monte Carlo tracer method. We however checked that our conclusions are qualitatively robust against choice of tracer scheme.

2.2.3 Identifying halos

We identify dark matter halos using the SUBFIND algorithm (Springel et al., 2001a) which begins with a friends-of-friends procedure (linking length $b = 0.2$) applied to dark matter particles. Gas and stars are associated to their nearest dark matter particle. Tracers are naturally associated with their nearest (parent) gas cells. Next, an iterative unbinding procedure which accounts for thermal energy identifies substructures within each FOF group which are gravitationally bound groups, with a minimum of 20 particles each. We refer to the largest such subgroup as the halo itself. Thus the halo selection specifically excludes gas in satellites and other orbiting substructures as well as unbound material, *at the time of selection*. However, it is important to note that material associated but not bound to substructures, including tidal features such as tails, would not be explicitly excluded by this procedure. We discuss below the separate issue of

classifying gas as either smooth or clumpy when following its thermal history.

The center of each halo is taken as the position of the particle with the minimum gravitational potential. We take as the virial radius $r_{\rm vir}$ the numerically computed $r_{200,\rm crit}$[1], and for the halo mass the mass of the largest bound subgroup. To determine accretion times, we need to estimate the time evolution of the virial radius of each halo as well as its position. However, SUBFIND is a purely 3D algorithm with no explicit time linking. We construct a parent tree (a simplified merger tree) by matching halos between successive snapshots based on the largest fraction of common dark matter particles. We require 60% particle agreement and reasonable mass and center offsets between snapshots. For consistency we restrict our analysis on a halo by halo basis to the time period for which the parent tree is well defined, which is true for $>90\%$ of halos with $M_{\rm halo} \geq 10^{10}\,{\rm M}_\odot$ over the chosen time window, as discussed below.

2.2.4 Maximum Past Temperature

We broadly consider three different means of measuring cold mode accretion. Two depend on the properties of associated halos. We introduce three important temperature values used throughout this work:

- $T_{\rm vir,cur}$ is the virial temperature for the parent halo of each tracer or gas particle at the "current" time (i.e. the time the gas selection is made, $z = 2$).
- $T_{\rm vir,acc}$ is the virial temperature evaluated at the "accretion" time on a particle by particle or tracer by tracer basis. We define the accretion time here as the most

[1]The radius within which the enclosed overdensity is 200 times the critical density $\rho_{\rm crit}$.

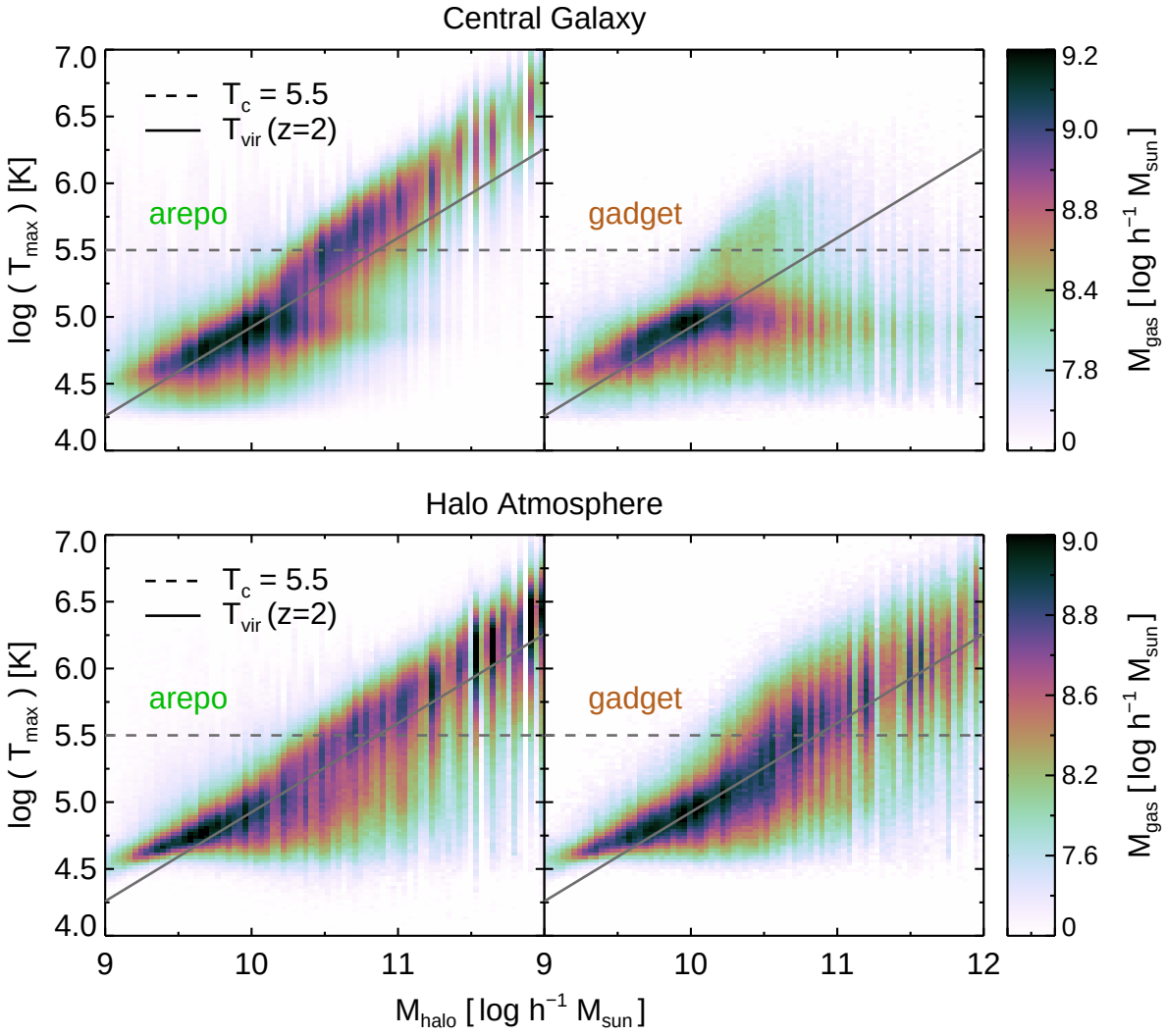

Figure 2.1 The mass-weighted histogram of past maximum temperature $T_{\rm max}$ for gas accreted onto central galaxies (top) and halo atmospheres (bottom) by $z = 2$, as a function of the parent halo mass at $z = 2$. AREPO and GADGET are shown separately in the left and right panels, respectively. In this figure we do *not* normalize these temperatures by any fraction of $T_{\rm vir,cur}$ or $T_{\rm vir,acc}$. A constant temperature of $T_c = 10^{5.5}\,{\rm K}$ is shown (dashed line) together with the virial temperature at redshift two (solid line). The two codes deviate strongly for galaxies in halos with masses above $\simeq 10^{10.5}\,{\rm M}_\odot$, where the strong contribution of hot gas in AREPO, which scales approximately with $T_{\rm vir}$, is mostly absent in GADGET. The cold $\sim 10^5\,{\rm K}$ gas which dominates the accretion in GADGET galaxies is largely absent in the AREPO simulation.

recent virial radius crossing time.

- T_c refers to a constant temperature threshold at some value, applied uniformly across all halo masses.

To estimate the virial temperature we use the definition of Barkana & Loeb (2001) which at redshift two gives

$$T_{\rm vir} = \frac{\mu m_p V_c^2}{2k_B} \simeq 4 \times 10^5 \,{\rm K} \left(\frac{M_{\rm halo}}{10^{11}\, h^{-1}{\rm M}_\odot}\right)^{2/3} \left(\frac{1+{\rm z}}{3}\right) \tag{2.1}$$

where V_c is the circular velocity at the virial radius and $\mu \simeq 0.6$ for fully ionized, primordial gas. We note that the redshift scaling as expressed here is only approximate, as it ignores small changes in cosmological parameters at the redshifts of interest.

We calculate a cold fraction for each halo or galaxy as the ratio of cold to hot accreted gas at $z = 2$ using each of the three methods. To calculate the maximum past temperature of a gas selection in GADGET runs we examine the temperature of each SPH particle at each snapshot back to the start of the simulation. We use 314 snapshots logarithmically spaced in redshift from $z = 30$ to $z = 0$, corresponding to a mean spacing at $z = 2$ of $\simeq 25$ Myr. To calculate the maximum past temperature of a gas selection using the Monte Carlo tracers in AREPO runs we first find all the tracer children of those gas cells at the target redshift. These tracers are then treated as a uniform population and the maximum past temperature of each, recorded for each active timestep of their parent gas cell, is obtained since the start of the simulation. That is, the mass weight of the cell is divided among its children tracers and each may potentially have a distinct thermodynamic history. In both cases we neglect temperature maxima while gas is on the effective equation of state (star forming).

2.2.5 Measuring Accretion

We separate gas in the "central galaxy" from the "halo atmosphere" by a cut in the density, temperature plane as

$$\log\left(T_{\rm gas}[{\rm K}]\right) - \frac{1}{4}\log\left(\rho_{\rm gas}[10^{10}M_{\odot}\,h^{2}{\rm kpc}^{-3}]\right) < 6.0 \tag{2.2}$$

following Torrey et al. (2012). This selection differentiates between hot gas in the halo and the cold, dense gas which is rotationally supported at the center of the halo. We add a radial cut at $0.15r_{\rm vir}$ to this separation criterion to prevent including any cold gas at large radii as part of the galaxy. Star particles are also included throughout our analysis and are considered as part of the galaxy if they satisfy the radial cut. Our results are quantitatively unaffected for reasonable choice of these three galaxy selection parameters.

In order to define accretion and measure accretion rates we introduce a "time window" over which accretion events are counted. A gas element is considered to have accreted onto a halo if it belonged to that halo at $z = 2$ and crossed the virial radius during the chosen time window. A gas element is considered to have accreted onto a galaxy if it belonged to that galaxy at $z = 2$ and either crossed the phase space cut in (ρ, T) or the radial cut at $0.15r_{\rm vir}$ during the chosen time window. The accretion time is taken to be the more recent (lower redshift) of these two events. The same definition is used for star particles in galaxies, where we search for the accretion event of the progenitor of each star particle only in the gas phase. That is, stars which remain stars for the duration of the chosen time window are not considered as accreted. Rates are calculated by normalizing the total accreted mass (counting particles or tracers) by the time period.

All accreted material is separated into one of three disjoint "modes" of accretion based on the following criteria, evaluated at the virial radius crossing time for each gas element:

- Smooth - not a member of any halo or subhalo, other than the main progenitor branch of its parent halo, and likewise back to $z = 6$.
- Clumpy - gravitationally bound to a halo or subhalo, other than the main progenitor branch of its parent halo.
- Stripped - not a member of any halo or subhalo (smooth), but gravitationally bound to some other halo or subhalo at any previous time back to $z = 6$.

These definitions allow us to exclusively consider accretion from the IGM by restricting our analysis to the smooth mode only. This is required in order to remove the "merger contribution", and is similar in spirit to the approach taken by previous studies of cosmological gas accretion for which it is possible to make this distinction (SPH and not grid codes). However, we note that our definition of clumpy is somewhat more restrictive than in previous works (e.g. Kereš et al., 2005) which would remove only the central galaxy component of each subhalo. By taking all gas bound to substructures as a merger contribution we impose a stronger restriction, and therefore remove a larger fraction of material in our analysis of smooth accretion. Finally, we measure the importance of tidally stripped material with the third mode, which is not direct accretion from the IGM but rather gas which has been previously gravitationally bound to some halo or galaxy. Unless otherwise noted, all results in this paper include only smoothly accreted material over an accretion time window of 1 Gyr. The only exceptions are

Figure 2.6 which considers the total accreted baryonic mass over an "integrated" time window extending back to $z = 6$, and Figure 2.11 which compares the smooth, clumpy, and stripped accretion modes.

2.3 Results

In this section we investigate differences between GADGET and AREPO in the thermodynamic history of gas accreted onto halos and the galaxies forming at their centers. Our primary proxy for this potentially complicated history is $T_{\rm max}$, the maximum past temperature of each gas element. In Figure 2.1 we show the mass-weighted histogram of $T_{\rm max}$ for smoothly accreted gas, as a function of parent halo mass. We also indicate a nominal constant temperature threshold value along with the virial temperature at redshift two.

Firstly, we see that GADGET and AREPO agree well on the relative distribution of past maximum temperatures contributing to *halo atmospheres* across the entire mass range (lower panels). For accretion onto galaxies (upper panels) at all halo masses the AREPO temperature distribution scales approximately with $T_{\rm vir}$, while in GADGET this behavior holds only for low mass halos below $\simeq 10^{10}\,{\rm M}_\odot$. In fact, we find that gas heating proceeds in relation to the virial temperature even for low mass systems. At $\simeq 10^{9}\,{\rm M}_\odot$ the $T_{\rm max}$ distribution may be beginning to scale somewhat shallower than $T_{\rm vir}$, flattening towards a uniform, cold temperature set by the IGM of a few times $10^4\,{\rm K}$. However, our simulation resolution is not sufficient at these low mass systems to convincingly demonstrate this. There is also an indication of a roughly constant temperature component at $T_{\rm max} \simeq 10^5\,{\rm K}$ in the GADGET simulation, which we discuss

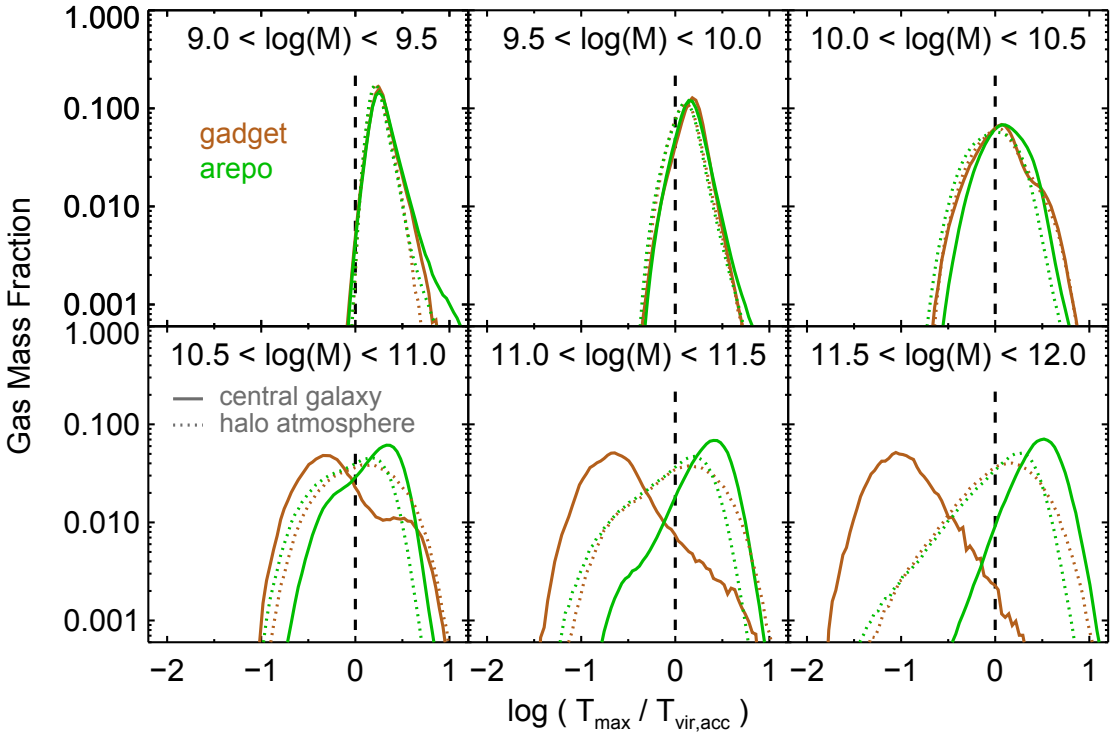

Figure 2.2 Relative distributions of past maximum gas temperature normalized by the virial temperature of the parent halo at the time of accretion, without any cold or hot mode differentiation. Each of the four curves in each panel is separately normalized, so that only their positions and shapes, but not their amplitudes, should be compared. At low halo masses (top three panels) and for gas accreted onto halos (dotted lines) the two codes agree well. For more massive halos above $\simeq 10^{10.5}\,\mathrm{M}_\odot$ significant differences arise in the thermal history of gas accreted onto galaxies, where gas in GADGET has predominantly $T_{\rm max} < T_{\rm vir,acc}$ while gas in AREPO has predominantly $T_{\rm max} > T_{\rm vir,acc}$.

further below. The scaling with $T_{\rm vir}$, together with the lack of any distinguishable segregating feature about a constant temperature threshold on the order of a few times $10^5\,{\rm K}$, motivates the comparison of $T_{\rm max}$ to $T_{\rm vir}$ on a halo by halo basis.

In Figure 2.2 we show the distribution of $T_{\rm max}$ for smoothly accreted gas, onto both galaxies and halos, divided into six bins of halo mass. In each case we now normalize temperatures by the virial temperature of the dark matter halo of each gas element, at the time of accretion. For accretion onto central galaxies for halos with $M \leq 10^{10.5}\,{\rm M}_\odot$ the two codes show excellent agreement. For such low mass systems, however, this agreement does not necessarily indicate a similar level of shock heating. The virial temperatures of these halos are sufficiently low that IGM pre-heating begins to set an effective temperature floor for the least massive systems. This heating outside the halo occurs through a combination of radiative energy input from the UV background, as well as shocks during large scale collapse, both of which we discuss further in the context of cold fractions. It is only for $M_{\rm halo} \geq 10^{10.5}\,{\rm M}_\odot$ where gas accreted onto centrally forming galaxies begins to show different thermal histories between GADGET and AREPO. At these masses, the majority of gas in the SPH calculation is never heated to temperatures comparable to the halo virial temperature at accretion. In contrast, a dominant fraction of the same accreted gas in AREPO reaches approximately the virial temperature, or above. This striking difference in the temperature history of gas accreted onto galaxies in halos with $M_{\rm halo} \geq 10^{10.5}\,{\rm M}_\odot$ is the first main result of this paper.

In this same mass regime both codes show evidence for potentially bimodal temperature distributions. This is most evident for GADGET galaxies between $10.5 \leq \log({\rm M}) \leq 11.0$ where $T_{\rm max}/T_{\rm vir,acc}$ could be well described by the sum of two symmetric distributions with different peak temperatures. As normalized, the colder peak moves

to lower temperatures with increasing halo masses approximately as $T \propto M^{2/3}$ and is consistent with a constant physical temperature of $T \simeq 10^5\,\mathrm{K}$, as seen in Figure 2.1. We note that common choices for a constant temperature threshold adopted by previous studies are above this value. The warmer peak shows no scaling with halo mass and is consistent with a heating process scaling with virial temperature. These two contributions represent, respectively, accretion of roughly constant temperature material from the IGM which has minimal interaction with the halo potential, and material which virializes to a mass dependent temperature. Results from both GADGET and AREPO are consistent with this picture, though they strongly disagree on the relative contributions of these two accretion channels.

2.3.1 Accretion Rates

To understand the nature of the differences, we measure the accretion rate onto both central galaxies and halos themselves as a function of halo mass. Figures 2.3 and 2.4 separate cold from hot mode accretion based on comparison of T_{max} to either a constant temperature threshold T_c or the halo virial temperature T_{vir}, respectively. The lower panels of both figures again indicate that the total accretion rates and thermal history of gas which accretes onto the halo itself is similar between the two codes. Our accretion rates for the low mass end are somewhat lower than in Kereš et al. (2009) for instance, which may be due to our more stringent definition of smooth accretion as well as our better resolution of the hierarchical nature of structure growth. When using a $T_{\mathrm{max}}/T_{\mathrm{vir,acc}}$ comparison this scaling is well described by a power-law in the mass range where systems are both well resolved and sufficiently numerous. We now discuss how the

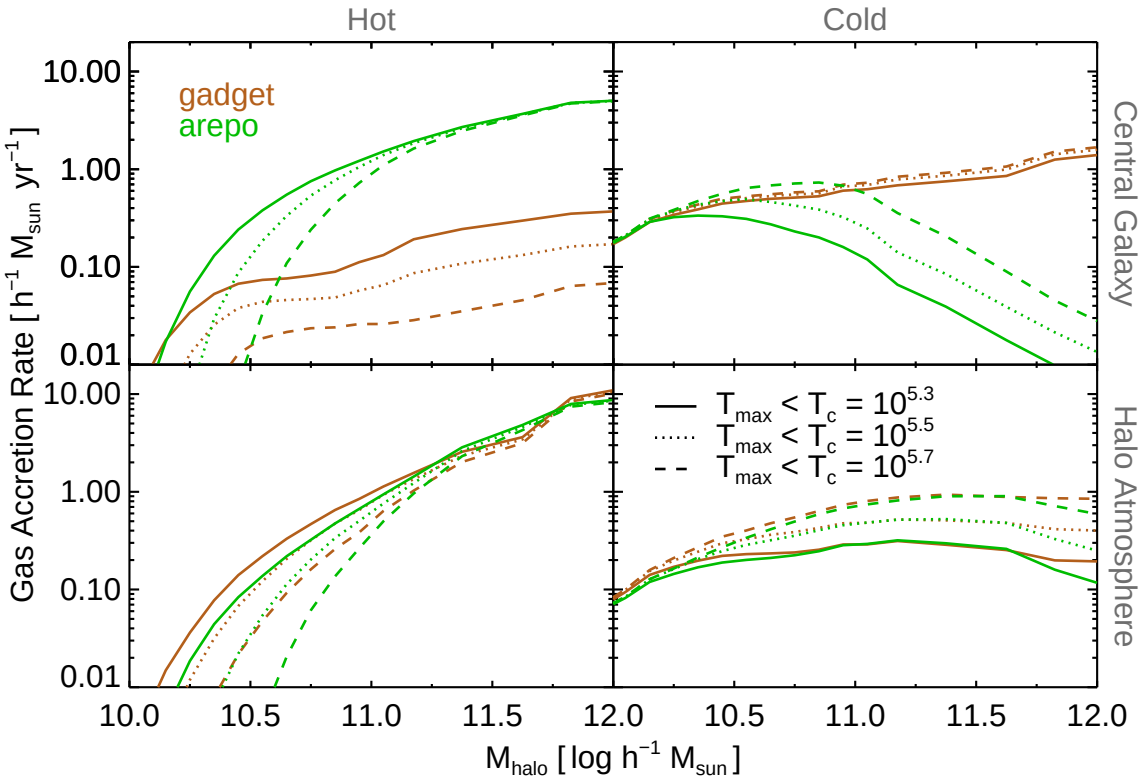

Figure 2.3 Gas accretion rate onto galaxies (top) and halos (bottom). In all four panels we define the cold mode as requiring $T_{\rm max}$ be less than a constant value T_c, as listed, demonstrating the sensitivity of the derived rates to the method of measuring the cold versus hot mode.

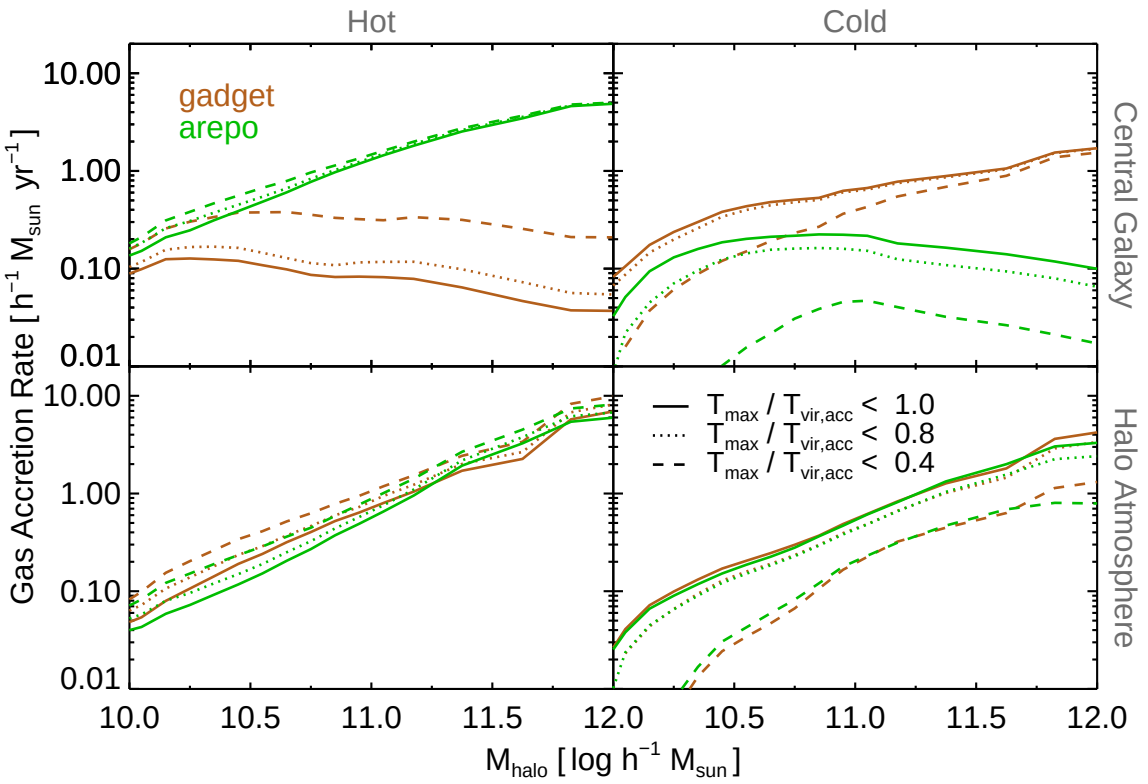

Figure 2.4 As in the previous figure, exploring the sensitivity of the gas accretion rate onto galaxies (top) and halos (bottom) to definition of the cold mode. Here we take cold mode gas as having $T_{\rm max}/T_{\rm vir,acc} < 1.0, 0.8, 0.4$, as listed.

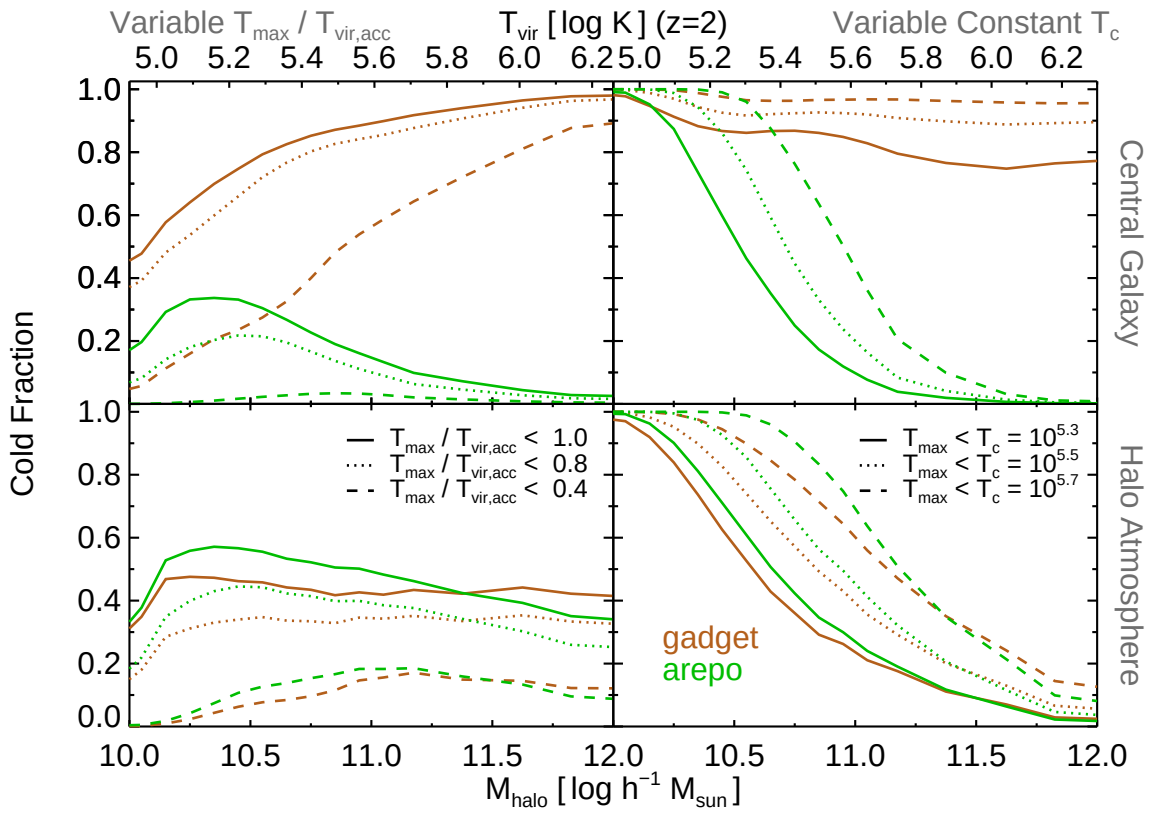

Figure 2.5 Sensitivity of the derived cold fraction to the manner in which it is measured. The top two panels consider gas accreted onto central galaxies whereas the bottom two panels consider gas accreted onto halo atmospheres. The left two panels consider a gas element to have been accreted cold if the fraction $T_{\rm max}/T_{\rm vir,acc}$ is less than $\{1.0, 0.8, 0.4\}$. The right two panels define cold mode gas as having $\log(T_{\rm max}[{\rm K}])$ less than a constant value of $\{5.7, 5.5, 5.3\}$. We find that the most massive systems which have $f_{\rm cold} \simeq 1$ in GADGET have a cold fraction approaching zero in AREPO, while the relatively sharp transition from cold to hot mode dominated, seen clearly in the constant temperature moving mesh case, is not evident when using the virial temperature comparison.

accretion rates compare between the two codes based on Figure 2.4 and $T_{\rm vir,acc}$, returning later to this particular choice.

The accretion rate onto galaxies in GADGET is dominated by gas which has always been relatively cold, across the entire mass range of $10^{10} \leq M_{\rm halo} \leq 10^{12}$. In AREPO the same galaxies are predominantly fed by gas which has experienced significant heating. This contrast is robust under any reasonable comparison of maximum past temperature to either T_c or $T_{\rm vir}$. The accretion rate in the hot mode is different by roughly an order of magnitude at $M_{\rm halo} \simeq 10^{11}\,{\rm M}_\odot$ between the two codes, having essentially zero contribution to halos above this mass in the SPH simulation. The primary reason for this difference is the efficient cooling of hot halo gas in AREPO which is artificially suppressed in GADGET by spurious viscous heating. Bauer & Springel (2012) studied the dissipation of energy from subsonic turbulence in SPH and found that spurious heating arises both from the premature dissipation of turbulent energy on large spatial scales and the viscous damping of SPH-noise on small scales. In the context of our current cosmological simulations, this prevents gas at the cooling radius from forming a proper Kolmogorov-like spectrum. The viscous heating offsets part of the gas cooling, keeping the gas locked in the diffuse, hot phase. In AREPO this same turbulent energy is more realistically dissipated by cascading to smaller spatial scales and higher densities, allowing significantly more efficient cooling of hot gas in the atmospheres of massive halos (Kereš et al., 2012). This cooling channel is a dominant accretion mode of massive halos in the moving mesh calculation.

The accretion rate of cold gas onto galaxies also differs between the two calculations. At $M_{\rm halo} \simeq 10^{11}\,{\rm M}_\odot$ GADGET galaxies have a factor of two larger cold accretion rates as corresponding AREPO galaxies, and this discrepancy grows with increasing halo mass

to an order of magnitude effect for $M_{\rm halo} \simeq 10^{12}$. We identify two primary numerical issues that cause the observed artificial increase of the cold gas accretion rate in the SPH calculation. Both operate by increasing the efficiency with which gas is heated in various regimes, resulting in a redistribution of gas to higher temperatures in the AREPO halos.

First, Creasey et al. (2011) demonstrate the issue of numerical overcooling due to shock broadening which arises due to the artificial viscosity treatment in SPH. At comparable resolution, the shock-capturing Godunov scheme in AREPO better resolves the resulting post-shock gas properties (including temperature), critical since smoothly accreting gas undergoes multiple small shocks as it virializes with hot halo gas. Secondly, spurious pressure forces in regions of steep density gradients (Agertz et al., 2007), particularly contact discontinuities, lead to the numerical fragmentation of gas structures into artificial "blobs" due to surface tension (see Torrey et al., 2012). A discussion of the formation mechanism of these blobs and their prevalence in different formulations of SPH can be found in Hobbs et al. (2012). Their presence is a significant contribution to the accretion rate of cold gas in GADGET. We have checked explicitly that gas in these blobs have sufficiently low $T_{\rm max}$ to be included in the cold accretion rate regardless of definition. For $M_{\rm halo} \geq 10^{11}\,{\rm M}_{\odot}$ these radially penetrating blobs contribute to the cold, smooth accretion rate in GADGET, directly transporting low temperature, low angular momentum material into the centers of halos (e.g. Kereš & Hernquist, 2009). No comparable gas features are present in the moving mesh calculation, and this cold gas contribution to galactic accretion is entirely absent.

2.3.2 Cold Fractions and the Transition Mass

The ratio of total accreted mass with $T_{\rm max}$ below some temperature threshold to the total mass above – the cold fraction, as a function of halo mass – is a widely reported quantity in studies of cosmological gas accretion, dating back to Kereš et al. (2005). Figure 2.5 explores how robust the derived cold fractions are to different types of parameter choices in the measurement of a cold mode. Using a constant temperature comparison (upper right panel), we note that the cold fraction for galaxies derived from our GADGET simulations are in agreement with Kereš et al. (2009). As expected from our previous discussion of accretion rates, this cold fraction is near unity for all halo masses, while in AREPO the cold fraction drops to zero for galaxies hosted in sufficiently massive halos above $\simeq 10^{11}\,{\rm M}_{\odot}$. We have seen already that lower cold fractions in AREPO arise from both a lower total gas mass accreted through a cold channel, as well as a larger total gas mass accreted through a hot channel.

We define the "transition mass" as the halo mass at which the cold fraction equals 1/2. Using the constant temperature cut $T_c \simeq 5.5$ (Kereš et al., 2005) leads naturally to a transition from $f_{\rm cold} = 1$ to $f_{\rm cold} = 0$ at a halo mass of $\simeq 10^{11.0}\,{\rm M}_{\odot}$in our highest resolution simulation. It is tempting to compare this to model predictions for the critical halo supporting a virial shock. Birnboim & Dekel (2003a) showed that this transition mass, defined as the point where the cold and hot modes have equal contribution, scales due to the contribution of metals to the cooling curve from $\simeq 10^{10.9}\,{\rm M}_{\odot}$ at $Z = 0$ to $\simeq 10^{11.6}\,{\rm M}_{\odot}$ at $Z = 0.3$ (redshift two). However, we observe that the appearance and location of this transition is by construction due to the choice of T_c. Given the scenario where all halos heat some large fraction of infalling gas to their virial temperature, this

use of a constant temperature threshold would also reproduce a transition in this mass range due solely to the changing virial temperatures. Consider the case of a halo with a virial temperature just below the chosen constant temperature threshold, corresponding to a mass of $\simeq 10^{11}\,\mathrm{M}_{\odot}$ taking the usual choice for T_c. Even if this halo shock heats all accreting gas to its virial temperature it would nevertheless be attributed a cold fraction of 100% under this method of measuring the cold accretion, which is clearly not a physically meaningful statement. We find that the subsequently derived transition mass increases monotonically with the chosen T_c if a constant temperature criterion is used to separate the two modes. From $T_c = 5.3$ to $T_c = 5.7$ (a factor of $\simeq 2.5$) the transition mass increases by roughly half a decade, which is consistent with the $T^{3/2}$ scaling expected solely due to the increasing virial temperature.

If we instead use a virial temperature comparison (upper left panel), the derived cold fractions show markedly different behavior. In this case GADGET galaxies exhibit a monotonic increase in their cold fraction from $\simeq 0.4$ to 1.0 from $M_{\mathrm{halo}} = 10^{10}$ to $10^{12}\,\mathrm{M}_{\odot}$. Galaxies in low mass halos below $M_{\mathrm{halo}} \simeq 10^{10}\,\mathrm{M}_{\odot}$ are no longer cold mode dominated, while those in high mass systems are. In contrast, AREPO galaxies still show the broad trend of decreasing cold fraction with increasing halo mass, though it tends towards continual evolution as opposed to sharp transition. Using the virial temperature as the dividing criterion on a halo by halo basis makes any transition significantly broader in halo mass. The second key result of this paper is then, given this more physically motivated means of measuring cold mode accretion in lower mass halos, there is no longer a clear feature in the cold fraction as a function of halo mass.

Although we argue for a comparison against T_{vir} and not T_c in this work, it is important to note that the two methods work equally well within a specific halo mass

range. In particular, for halos with $T_{\rm vir} \gg T_c$ such that virialization related heating is well separated by this cut. At $z = 2$ this must require as a minimum $M_{\rm halo} \geq 10^{11}\,{\rm M}_\odot$ which is indeed the mass range considered in early studies (e.g. Kereš et al., 2005). Unfortunately this minimum halo mass falls well within the cold-hot transition range as reported at this redshift as well as at redshift zero, where the restriction becomes $M_{\rm halo} \geq 10^{12}\,{\rm M}_\odot$ – the midpoint or even tail of the transition (van de Voort et al., 2011b; Faucher-Giguère et al., 2011). Caution must also be taken using any $T_{\rm vir}$ comparison as the more restrictive selections of the cold gas selection $(0.4, 0.8)$ for low mass halos fall below the effective temperature floor of our simulations. We run a separate experiment at the 256^3 resolution which excludes the heating contribution from the UV background. The resulting temperature distribution of IGM gas extends to significantly lower temperatures, and the turnover in cold fraction around halo masses of $\simeq 10^{10.5}\,{\rm M}_\odot$ essentially disappears. We then attribute this feature in the cold fraction to the high IGM gas temperature prior to accretion. This also indicates the mass scale below which a temperature criterion of any kind will be unable to separate physically distinct accretion modes.

Finally, we note that switching from the "current" (not shown) to the "accretion" virial temperature criterion leads to a $\sim 10\%$ change, lowering the cold fraction across all halo masses. This change is unlikely to have great impact on the general interpretation of the cold mode importance. However, the virial temperature of a particular halo can change significantly between the time material is accreted and $z = 2$. To physically motivate our analysis requires that we either consider the virial temperature when it was relevant to the question of shock heating - at the time of accretion - or restrict the analysis window to a short enough time-scale over which the virial temperature does not

appreciably change.

2.3.3 Integrated Baryonic Budget

The strongly discrepant history of gas which eventually builds up galaxies by $z = 2$ has important ramifications for the morphology and structure of these forming galaxies. In Figure 2.6 we show the total gas mass smoothly accreted into both galaxies (top) and halo atmospheres (bottom), decomposed into hot and cold contributions based on a comparison of $T_{\rm max}$ to $T_{\rm vir,acc}$. Here, in contrast to earlier figures, we extend the "time window" over which accretion is counted to include a longer evolutionary history of these systems. We extend back in time as far as each halo is successfully tracked in our parent tree scheme, up to $z = 6$. Consequently Figure 2.6 shows the integrated contribution of smoothly accreted baryons to the $z = 2$ mass budgets of these systems. When compared to the accretion rates calculated only over the past 1 Gyr we find that the previous conclusions regarding the comparisons between the two codes and the balance between the hot and cold accretion modes remain qualitatively the same. Since the two simulations have the same initial conditions, we naturally expect their properties to tend towards better agreement at sufficiently high redshift. In particular, comparing the total accreted mass onto galaxies at halo masses of $\simeq 10^{12}\,{\rm M}_\odot$ we find that AREPO has a factor of $\simeq$10 higher mass accreted with $T_{\rm max} > T_{\rm vir}$, and a factor of $\simeq$3 lower mass accreted with $T_{\rm max} < T_{\rm vir}$ when compared to GADGET. This represents a significant discrepancy in the gas accretion history integrated down to $z = 2$, which is directly related to the morphological and kinematic differences of the respective galaxy populations.

A long standing difficulty of SPH has been its ability to produce central objects

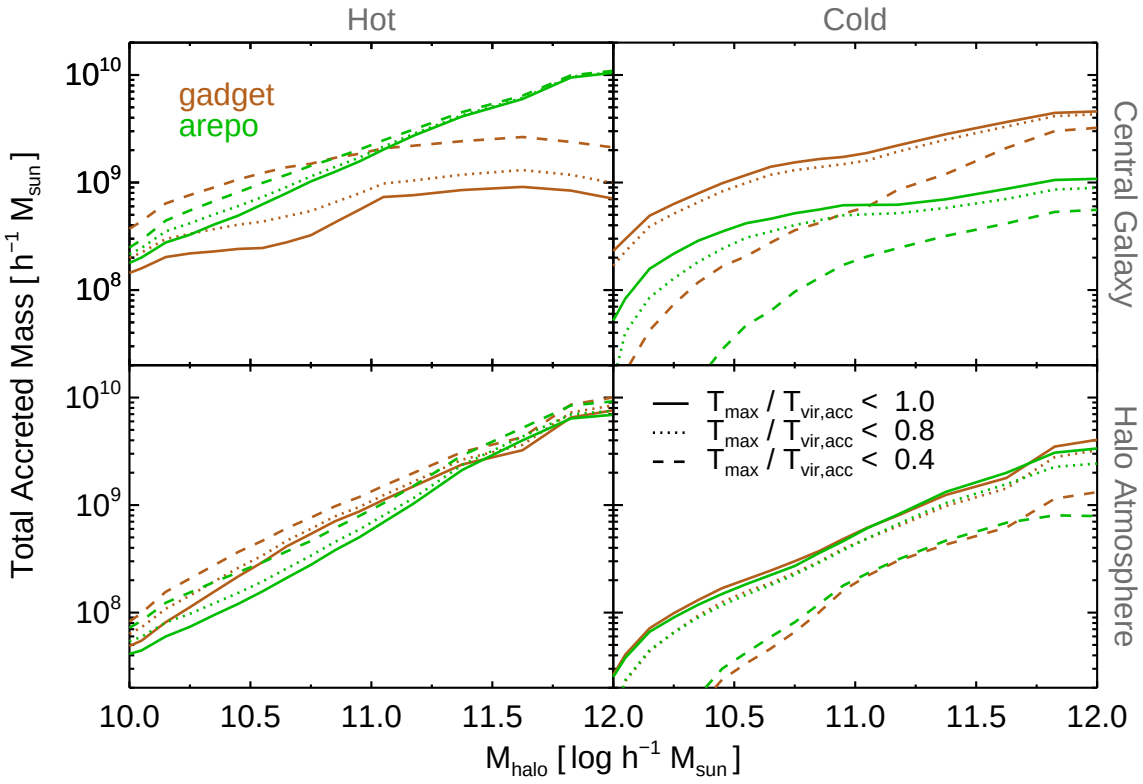

Figure 2.6 The total amount of gas smoothly accreted onto either galaxies or halo atmospheres by $z = 2$ decomposed into hot and cold contributions. Following our earlier discussion we make this separation based on some ratio of $T_{\rm max}$ to $T_{\rm vir,acc}$, as shown. Unlike in previous figures, we include gas accreted over an "integrated" time window encompassing the evolution of these systems back to $z = 6$ or as far as they have well-defined parents.

with angular momentum sufficient to form rotationally supported disks with realistic scale lengths (Vogelsberger et al., 2012). The relative differences between the angular momentum content of cold gas accretion in GADGET when compared to hot gas accretion in AREPO explains why galaxy properties differ between these two approaches. AREPO galaxies are more disk-like and more rotationally supported, with systematically larger disk scale lengths (Torrey et al., 2012) in part because they are significantly built from the diffuse, hot halo gas, which has a large reservoir of angular momentum (see also Sijacki et al., 2012). In contrast, GADGET galaxies are more compact and less rotationally supported in part because they are being built mainly from the cold gas, which is clumpy due to numerical artifacts and which tends to artificially lose its angular momentum. That realistic disk morphologies are related to the angular momentum reservoir from which they form has also been found regardless of numerical technique, in the context of galactic fountain material (Brook et al., 2012) and cooling from hot coronae (Sales et al., 2012). Here the differences in the hot halo itself may also play a role. As we discuss below, hot halo gas in AREPO has a smaller spatial extent and reaches somewhat higher central densities than in GADGET (Kereš et al., 2012), which leads to shorter cooling times and thus larger hot accretion rates onto centrally forming galaxies.

2.3.4 Geometry of Accretion

We find that the filamentary nature (geometry) of accreting gas at and around the virial radius of massive halos is in good qualitative agreement between GADGET and AREPO. At high redshift ($z = 2$) filaments are a common feature in a majority of

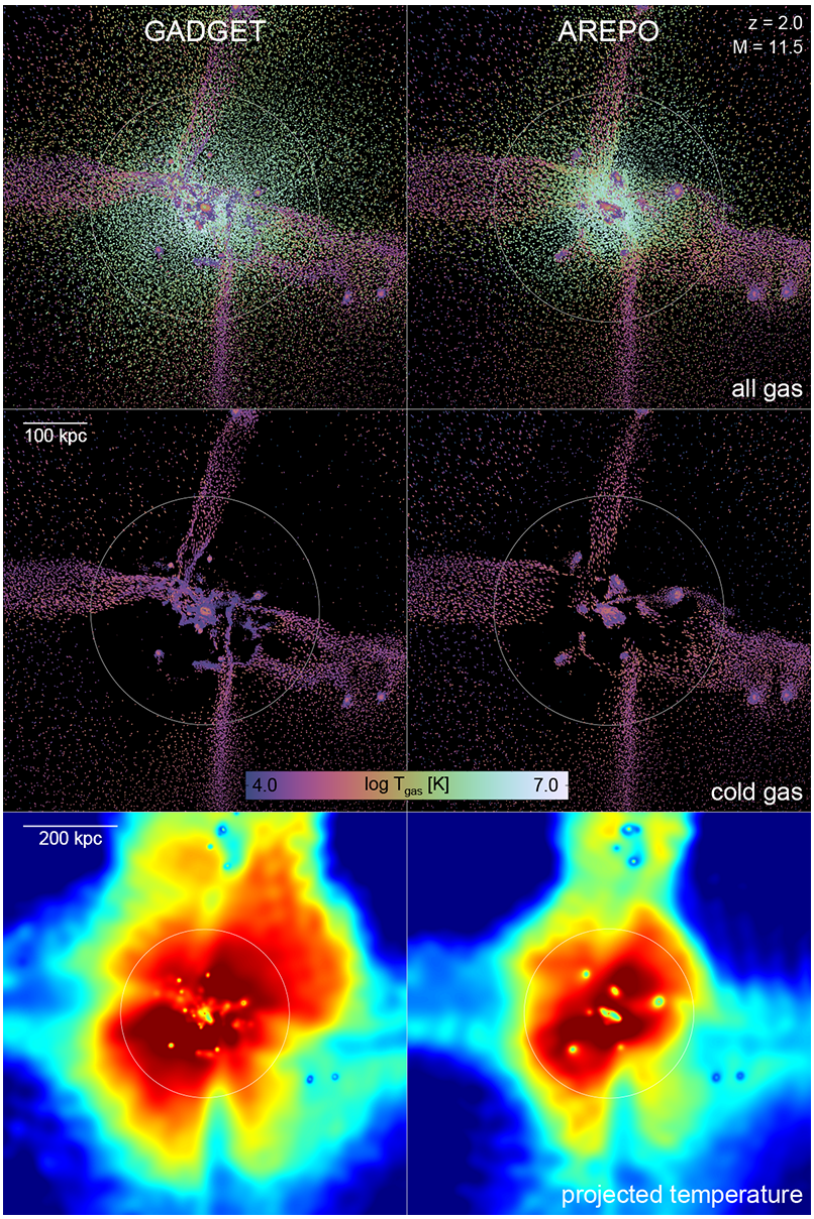

Figure 2.7 Density and temperature distribution of individual gas elements for a $\simeq 10^{11.5}\,\mathrm{M}_\odot$ halo at $z = 2$ in GADGET (left panels) and AREPO (right panels). The top panels show gas of all temperatures, while the middle panels show only gas with instantaneous temperature $T < 10^5$ K. Ticks are color coded by instantaneous temperature and with directions representative of the local velocity field. The bottom panels show the mass-weighted temperature projection for lines of sight through a larger cube of side length $5r_{\mathrm{vir}}$. In both cases locally overdense and overcool filamentary gas structures penetrate the hot halo at r_{vir}, but evolve quite differently towards smaller radii.

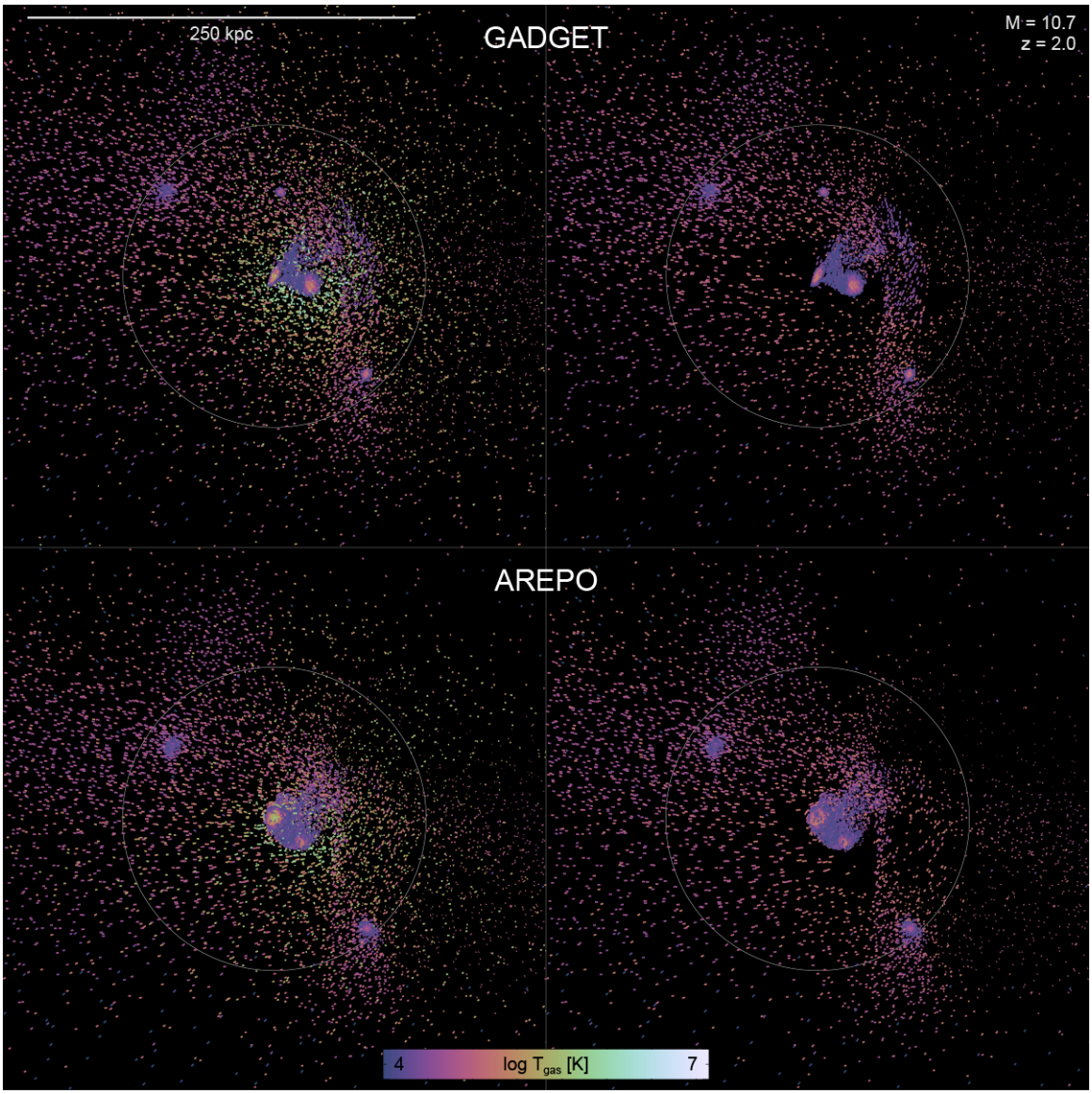

Figure 2.8 Density of individual gas elements for a $\simeq 10^{10.7}\,\mathrm{M}_\odot$ halo at $z = 2$ ($T_{\rm vir} \simeq 10^{5.4}$ K) in GADGET (top panels) and AREPO (bottom panels) color coded by instantaneous temperature and with directions representative of the local velocity field. The left panels show gas of all temperatures, while the right panels show only gas with instantaneous temperature $T < 10^5$ K. This halo, as with a majority of lower mass systems, does not show the same prominence of filamentary gas inflows.

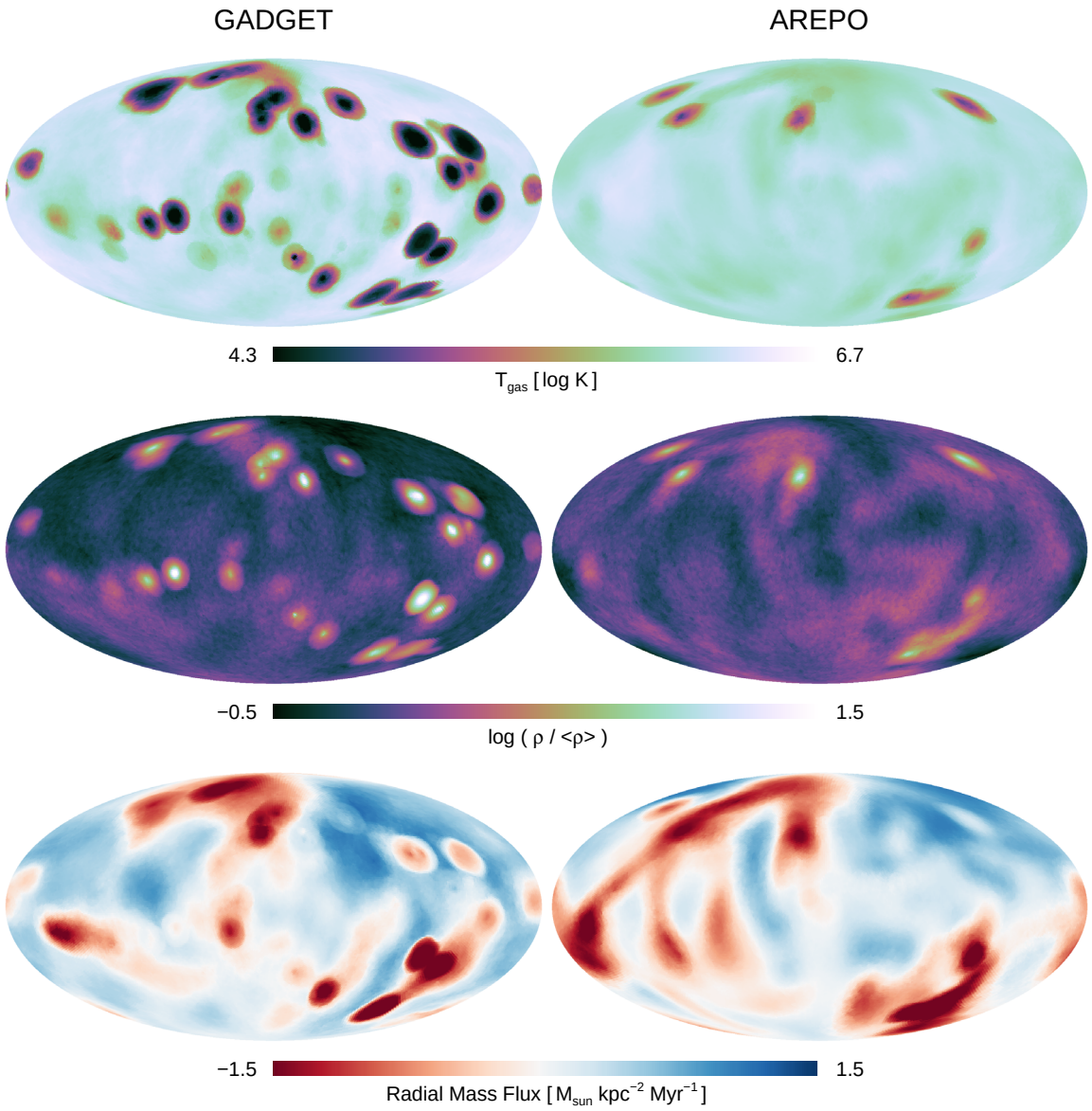

Figure 2.9 An all-sky Mollweide projection of temperature, overdensity, and radial mass flux on the surface of the half virial sphere centered on one halo with $\log(M) \simeq 11.8$. All gravitationally bound substructures have been removed. The clustered regions of negative radial mass flux represent cross sectional views of gas filaments. They are associated with mass overdensities and cold temperatures with respect to the shell averages. The larger arcing features in the map are infalling sheets of gas with large angular extent on the sky.

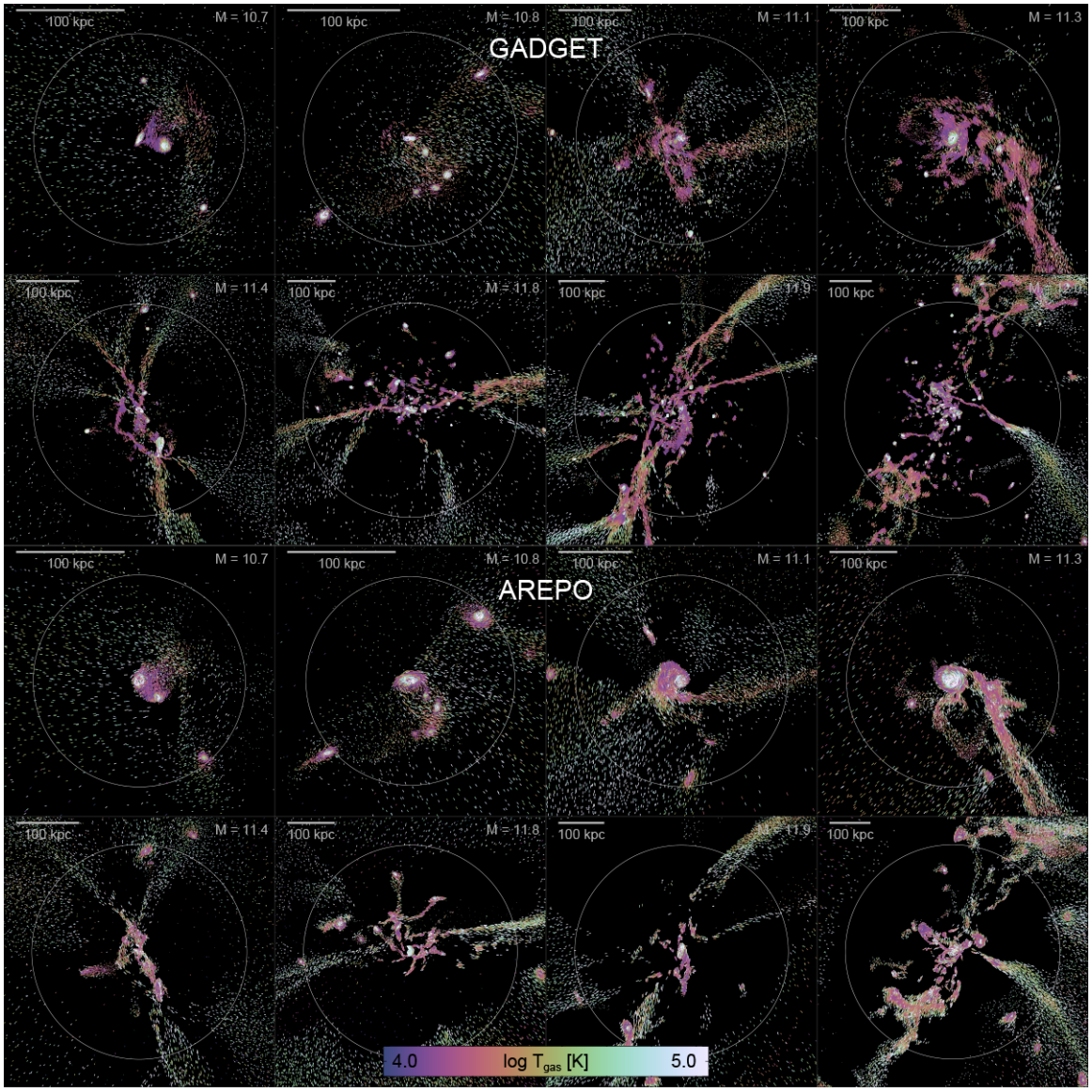

Figure 2.10 Comparison of strictly cold gas with instantaneous temperature below 10^5 K for eight matched systems at redshift two. The top eight panels show systems from the GADGET simulation while the bottom eight panels show the identical systems in AREPO. We commonly find examples of predominantly spherical accretion in low halo masses and differences in filamentary gas structures for larger halo masses, as discussed in the text.

massive ($M \geq 10^{11.5}\,\mathrm{M}_{\odot}$) halos. Prior to this section, we have included only smoothly accreted gas in our analysis and discussion. When visualizing the geometry of accretion, we show all gas elements, and do not restrict to only the smooth accretion mode as in previous plots. Figure 2.7 shows a gas projection of a particular halo matched between the two simulations, where in the top two rows temperature is mapped to particle color and the local velocity field by tick direction. The top panels shows all temperature gas, while the middle panels show only gas with instantaneous temperature below 10^5 K. In GADGET (left column) we observe that filamentary gas structures often coherently flow to $r \leq 0.25 r_{\mathrm{vir}}$ and become progressively smaller in cross section. They can connect directly to central gas structures at $r \leq 0.1 r_{\mathrm{vir}}$, although the contribution of these thin streams to the total accretion rate of gas onto the galaxy is unclear. Their temperature profiles remain well below that of the virialized halo gas and show little evidence for shock heating. In AREPO (right column) these same gas filaments are more diffuse and experience significant heating at comparable radii. For instance, the two filaments from the top and left which seem to "disappear" in the right panel in fact heat to the temperature of the halo gas while remaining coherent. We discuss this important contribution to the accretion rates of hot gas below.

Another significant difference between the two results is the distribution of the hot halo gas itself. The bottom row of Figure 2.7 shows the mass-weighted temperature projection for the same halo in both AREPO and GADGET. In order to make a fair visual comparison we integrate each sightline by spatially distributing the gas mass of both SPH particles and Voronoi cells using the same SPH kernel approach. We observe that hot halo gas in GADGET extends to larger radii than in AREPO. To quantify this effect, we measure the stacked radial temperature profile of all halos in the box with

$10^{11.5}\,\mathrm{M}_\odot < M_{\rm halo} < 10^{12.0}\,\mathrm{M}_\odot$. At the virial radius, the gas temperature in AREPO halos is a factor of $\simeq 2/3$ lower than in GADGET – the mean gas temperature at $r_{\rm vir}$ in GADGET is obtained at $\simeq 0.7 r_{\rm vir}$ in AREPO. As we discuss in the context of accretion rates, the smaller hot gas reservoir in AREPO halos results from more efficient cooling out of the halo, while in the case of GADGET this cooling is suppressed due to spurious numerical heating.

An independent point which has led to confusion is the nature of accretion in lower mass halos. Figure 2.8 shows an example of such a system with $\mathrm{M}_{\rm halo} \simeq 10^{10.7}\,\mathrm{M}_\odot$ at redshift two. In this case there are no prominent filaments, nor do we expect any (Katz et al., 2003; Kereš et al., 2009). The virial temperature of this system is $\simeq 10^{5.4}$ K and would lie at the low mass end of the transition mass region from cold to hot mode dominated. Given the usual definition of $\log T_c = 5.5$, gas which had accreted and shock heated to the virial temperature would be considered cold mode accretion. In contrast, under a definition based on some fraction of $T_{\rm vir}$ a large fraction of that same gas would be considered hot mode. Indeed, for this mass range in Figure 2.5 we expect that approximately half of the halo gas has $T_{\rm max} > T_{\rm vir}$ and half has $T_{\rm max} < T_{\rm vir}$. It is then critical to differentiate between gas which shocks to $T_{\rm vir} \leq T_c$ and gas which accretes "cold" into low mass halos with short cooling times. This is particularly the case since the signature of both mechanisms is similar in maximum past temperature, making any separation based on temperature alone difficult if not impossible.

In order to better compare the properties of filaments we construct maps of various fluid quantities which reveal the cross-sectional profiles of these structures. Figure 2.9 shows gas temperature, density and radial mass flux calculated on the surface of a sphere centered on one particular halo, the $\log(M) = 11.8$ example from Figure 2.10, and with

a radius equal to half the virial radius. We use a simple mass-weighted tophat kernel to interpolate these quantities onto equal area pixels using the HEALPIX scheme (Górski et al., 2005), and display the result using a Mollweide all-sky projection.

Pronounced differences appear in images of projected gas density. We find that the radial mass flux inward at these radii is dominated in the SPH halos by a large number of thin filaments and gas blobs with progressively smaller cross section. The moving mesh case shows no evidence of gas structures of comparably small angular extent. The density profiles of the filaments are also not as centrally concentrated and have larger angular extent. Although filaments reach roughly the same minimum temperature, the transition to the hot $10^{5.5} - 10^6$ K gas in the halo is more gradual in AREPO. That is, filaments have less well-defined boundaries, indicative of increased mixing, in the moving mesh calculation, though they appear to be roughly at the same level of pressure equilibrium. Although the angular covering factor of infalling material is slightly larger in AREPO, it is interesting that for this particular halo, where we have accretion onto the central galaxy dominated by $T_{\rm max} > T_{\rm vir,acc}$ gas, the mass flow is still highly aspherical. That these filaments of hot gas, originating from large scale features in the IGM, contribute to the hot gas accretion rate distinctly from classical cooling flows out of hot halo gas is the third key point of this work. The presence of this coherent hot accretion implies that a simple cooling rate calculation based on halo properties cannot accurately describe high redshift accretion in massive halos. This may have important consequence for semi-analytical models (SAMs) which assume radiative cooling from a quasi-static atmosphere.

Comparing large $\simeq 1\,$Mpc scale gas and dark matter density fields, the cold $\simeq 10^4\,$K filaments which are relatively narrow within the virial radius can clearly be identified as

extensions of filaments of larger diameter outside the halo. Given that GADGET and AREPO share the same treatment of dark matter and the same gravity solver, and show good agreement in IGM properties outside halos (Bird et al., 2013), it is unsurprising that at the virial radius the differences we observe above minimize and the two codes show similar gas properties.

We show eight further examples of strictly cold gas below 10^5 K in matched systems spanning the mass range $10.7 \leq \log(M_{\rm halo}[M_\odot]) \leq 12.0$ in Figure 2.10 illustrating that the similarities and differences we describe are commonly found in the simulated cosmological volume[2]. halos in both simulations support approximately the same frequency of filamentary structures over any given range of halo mass. The exception appears to be filaments that are tidal in nature, which are much more pronounced – longer and more concentrated – in AREPO. The cold gas distribution in the most massive GADGET halos is dominated by the large blob populations, which are absent in the corresponding AREPO systems.

2.3.5 The Clumpy Contribution

We seek to understand the accretion of material onto halos and galaxies originating directly from the intergalactic medium. If we include all material, the resulting cold fractions and cold mode accretion rates are only upper limits, due to some nonzero contribution from minor mergers. Past numerical simulations run with SPH and focusing on cold mode accretion have specifically addressed this issue of merging substructures,

[2]A catalog of similar visual comparisons for all objects with $M > 10^{10.75}\,{\rm M}_\odot$ at $z = \{0, 1, 2, 3\}$ is available online at `www.cfa.harvard.edu/itc/research/movingmeshcosmology/`.

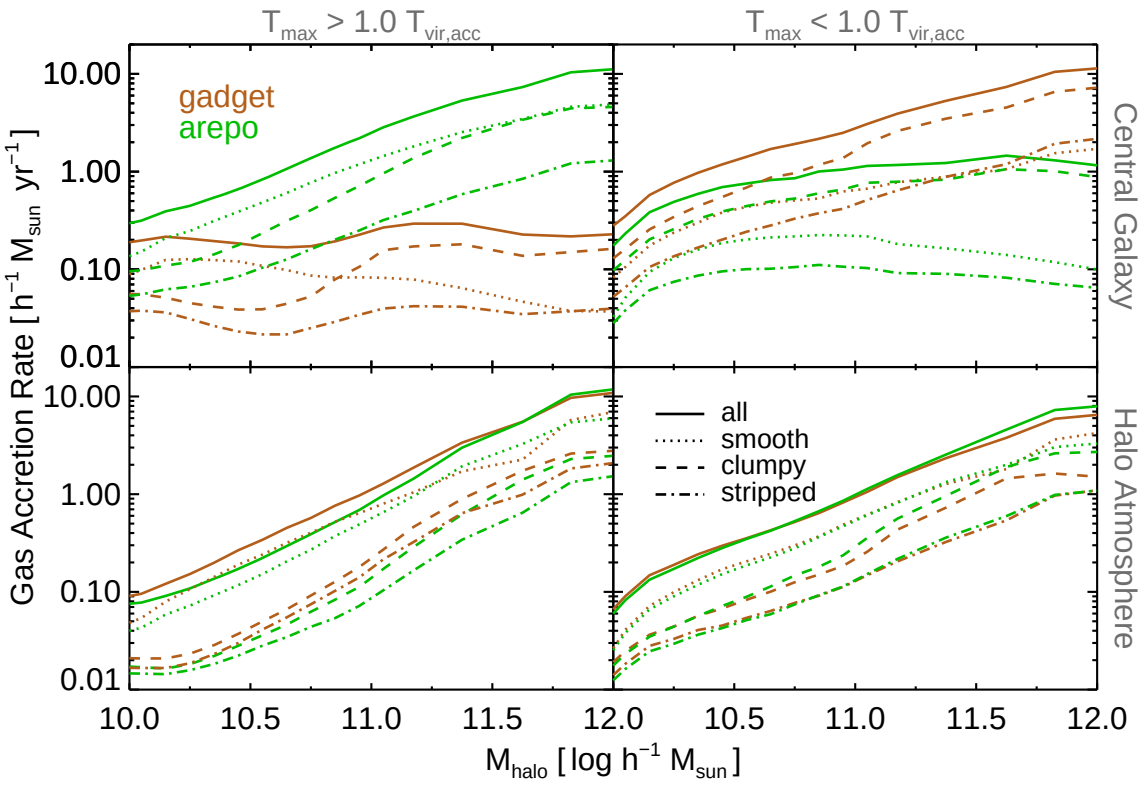

Figure 2.11 The total accretion rate separated into contributions from our three different modes: smooth (dotted), clumpy (dashed), and stripped (dot-dashed).

both resolved and unresolved (Kereš et al., 2005, 2009; Brooks et al., 2009; van de Voort et al., 2011b). The exact approach in separating out the merger contribution differs. In Brooks et al. (2009) and in this study we remove material that belonged to any galaxy halo other than the main galaxy under consideration, while in Kereš et al. (2009) for instance only the central galaxy and not its associated circumgalactic (or halo) material is removed. Other studies have neglected this distinction and included the accretion of dwarf satellites as part of a cold mode (Shen et al., 2013). Although simulations using the AMR technique note the presence of substructure embedded within filaments, they also neglect to make this separation (Dekel et al., 2009; Agertz et al., 2009) though due to technical constraints. One option as in Ocvirk et al. (2008) is to use an instantaneous criterion such as gas density to identify a clumpy component. However, without some additional tool for tracing gas properties back in time (tracer particles) the contribution of substructures and associated material cannot be unambiguously accounted for, and a measurement of the cold gas accretion from the IGM represents an upper limit.

Applying the three definitions for "smooth", "clumpy", and "stripped" from Section 2.2.5 we decompose the accretion rates separately for hot and cold accretion in Figure 2.11, where we here use $1.0 \times T_{\rm vir,acc}$ as the separating criterion. As previously discussed, our definition of clumpy includes all material bound to satellites – both galaxies and their associated circumgalactic medium. We focus first on cold material accreted onto central galaxies, for which the clumpy mode is the $\simeq 60\%$ dominant contributor at all halo masses in both GADGET and AREPO. Smooth accretion accounts for only $\simeq 10\%$ of the cold mass accumulation at the massive end, where most of the cold accretion onto central galaxies arises from merging substructure. For the central galaxies of the lowest mass systems we consider with $\simeq 10^{10}\,{\rm M}_\odot$ the smooth contribution is slightly larger

($\simeq 20\%$). This fraction is relatively converged with resolution; however, any estimate of the merger contribution for the lowest mass systems may be underestimated due to numerical resolution. For all halo masses the stripped component is comparable to the smooth component. The accretion of hot material onto central galaxies shows similar behavior, except that the contribution from clumpy mode gas scales up with increasing halo mass, while the contribution of smooth gas scales down. The only notable difference between the two codes is for material accreted hot onto galaxies hosted in massive halos with $M_{\rm halo} \geq 10^{11}\,{\rm M}_{\odot}$ for which AREPO shows a lower clumpy fraction (40% vs 70%) and a higher smooth fraction (40% vs 20%), consistent with the picture that smooth gas is more efficiently heated in AREPO.

The importance in separating different gas origins in the simulations can be seen by reference to observations of our local environment, in which the Magellanic Stream is a large reservoir of cold, neutral gas thought to arise from the tidal interaction of dwarf satellites (Besla et al., 2010, 2012). Its significant spatial extent through the halo of the Milky Way does not arise from a gas filament of direct cosmological origin, although it might easily be characterized as such. Similar features in external galaxies will be even more difficult to characterize. Although separating the contribution of minor mergers and accreted tidal debris from gas infalling directly from the IGM presents a technical challenge for simulations, it is necessary to truly identify the origin of the gas that ends up forming galaxies, without mixing up contributions from disparate physical origins.

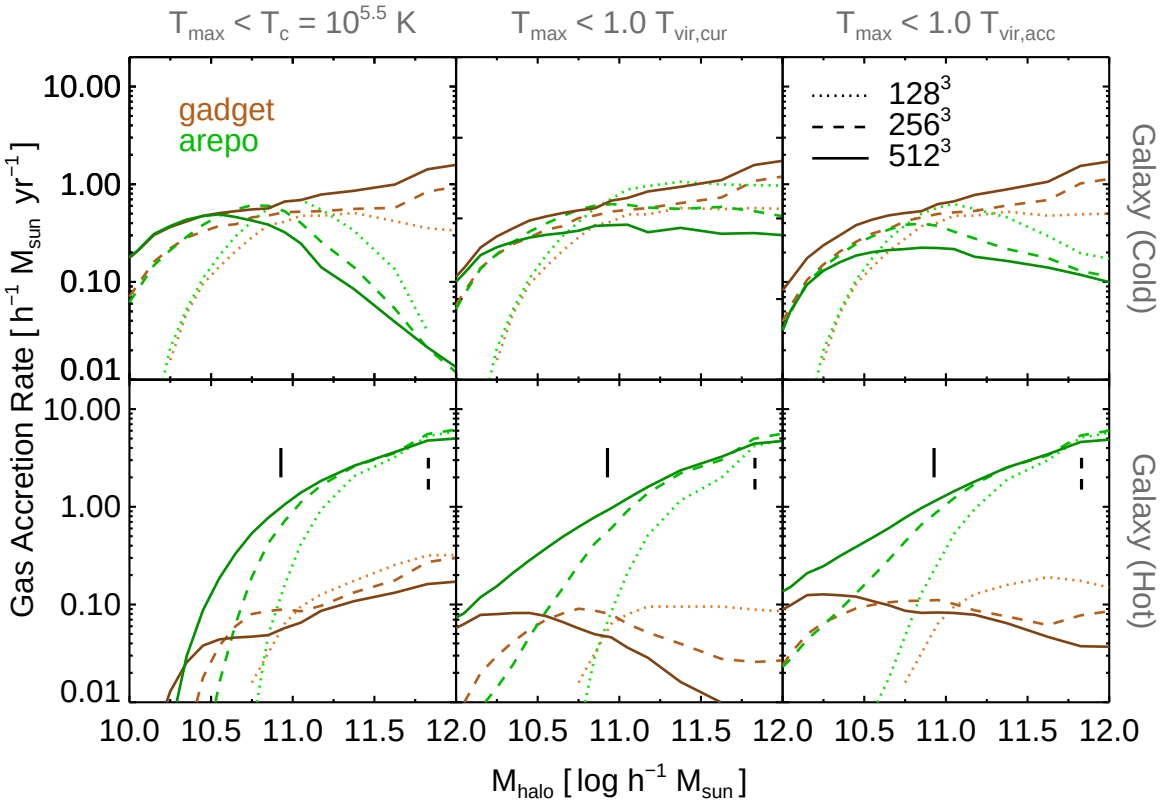

Figure 2.12 Convergence of the accretion rate onto galaxies with resolution. The vertical black lines show the halo masses at which $\Omega_b M_{\rm halo}/m_{\rm target} \simeq 5 \times 10^4$ for the 512^3 and 256^3 runs, a rough requirement on the number of gas mass elements for which gas accretion physics begins to be adequately resolved and converged in AREPO.

2.3.6 Resolution Convergence

In this last section we address some of the numerical issues related to our moving mesh calculation. In Figure 2.12 we show the convergence of the gas accretion rate onto central galaxies for our three resolution levels. The total (the sum of the top and bottom panels for any given column) is well converged across the given halo mass range, and all resolution trends are qualitatively similar regardless of the separating line drawn between hot and cold. In general, Vogelsberger et al. (2012) showed the increase in high redshift global star formation rates due to better resolving dense gas in low mass halos. We expect that smooth accretion rates will then decrease as resolution increases, with a corresponding increase in either clumpy or stripped accretion, as more gas mass collapses into substructures at early times. Galaxies in AREPO exhibit this behavior for material accreted cold, i.e. directly from the IGM. In contrast, the accretion rate of hot material from the halo increases strongly with resolution for halo masses which are poorly resolved in our simulations. This enhanced cooling from the halo occurs due to a combination of the changing temperature distribution of hot halo gas, together with inadequately resolved mixing in two particular regimes, as we discuss below.

First, Kereš et al. (2012) shows that the radial temperature profiles of AREPO halos become steeper and reach higher maximum temperatures with increasing resolution. We verify this and furthermore find that it is particularly problematic for poorly resolved systems where the number of gas cells in the entire halo is of order a few thousand or fewer. For more well-resolved halos ($M_{\rm halo} \geq 10^{11}\,{\rm M}_{\odot}$) the temperature profiles are converged between our medium and high resolution runs. This occurs in part because the halo material more fully populates the high temperature tail of the thermal distribution

of the virialised gas.

The strong mixing between the rotating galactic disk and the halo gas identified by Sijacki et al. (2012) is also artificially enhanced at low resolution where this interface is poorly resolved. This shortens the cooling times of hot halo gas near the disk and drags down the temperature profile, which ultimately results in accreting material heating to a lower peak temperature. The potential for overmixing at low resolution is a deficiency inherent to grid codes, including AREPO. In this case the moving mesh nature of the scheme acts to minimize the effect, though it cannot be suppressed entirely as in SPH codes. The same issue arising at the disk-halo interface also occurs in a second regime, with infalling substructures which penetrate into the halo and lower the cooling times of the hot gas due to overmixing. This is in contrast to SPH, where a relatively poor treatment of fluid instabilities prevents efficient stripping of gas from infalling satellites (Sijacki et al., 2012).

Galactic disks in the low mass halos are also poorly resolved in our lowest resolution simulations (Kereš et al., 2012), which may extend them beyond our $0.15r_{\rm vir}$ cut, particularly in AREPO where they are larger and more extended to start with. The result is an underestimation of accretion rates, while at higher resolution these disks are more compact and the accretion is largely recovered. This effect may preferentially influence gas with larger $T_{\rm max}$ which is found at systematically larger radii and lower density in galactic disks.

2.4 Discussion and Conclusions

We are motivated to define a cold accretion mode based on the markedly distinct channel of cold filamentary streams penetrating the hot atmospheres of massive halos. Given the existence of two distinct modes, how can we distinguish between the two? We observe that this form of gas accretion enables gas to infall onto galaxies without shock heating to an appreciable fraction of the virial temperature. This motivates a selection criterion based on virial shocking, or lack thereof. This is not a new idea. Indeed, Kereš et al. (2005) explored this option but concluded that the distinction between hot and cold was less clear than with a constant temperature cut. van de Voort et al. (2011b) likewise compared hot fractions obtained using either a threshold $\log T_c = 5.5$ or various fractions of $T_{\rm max}/T_{\rm vir}$. That work concluded, as we do, that the hot fraction depends strongly on the choice of criterion. Furthermore, van de Voort et al. (2011b) acknowledges that gas accreted in their "cold mode" may have in fact gone through a virial shock, indicating that it is critical to investigate the gas selection made with a global constant temperature threshold over the range of halo masses being considered.

Furthermore, even comparing the temperature history of gas elements to the virial temperatures of their parent halos is an inexact probe of whether or not gas shock heats. The picture that gas shock heats up to $T_{\rm vir}$ at $r_{\rm vir}$ is an over simplification of a virialization process that may involve multiple shocks at increasingly higher temperatures as infalling material encounters the temperature gradient of halo gas. Indeed, shock heating may not even be the most fundamental question as to how galaxies obtain their gas. This might be better answered by considering the cooling time of gas (White &

Rees, 1978; White & Frenk, 1991), and so the time-scale for collapse onto a forming galaxy. The radial stalling associated with incorporation into the hot halo gas may be more useful than the gas temperature itself. In the same manner, the angular momentum history of infalling gas may provide a more physically motivated segregation of different accretion modes. We have found that the interpretation of the angular momentum, for instance, is not nearly as clear as with maximum past temperature and we suggest this is the primary reason why $T_{\rm max}$ remains the predominant technique used to study cold mode accretion.

Based on additional temperature cuts, recent simulations have also identified a third "warm" mode of gas accretion (Joung et al., 2012; Murante et al., 2012) in the $10^5\,{\rm K} < {\rm T} < 10^6$ K regime and selected by constant temperature bounds over either instantaneous temperature (in Joung et al. (2012) using AMR) or maximum past temperature (in Murante et al. (2012) using SPH). In the case of Murante et al. (2012) this prominent mode is attributed to effective SN thermal feedback. Unfortunately, since these studies are focused on low redshift and include galactic outflows we cannot make a direct comparison to our results. We have not therefore investigated this particular temperature selection criterion, although a quantitative comparison with AMR results and an exploration of the impact of feedback are crucial and natural directions for future work. The commonly perceived agreement between SPH and AMR simulations, of cold mode accretion as a dominant contributor to galaxy formation, is based on a mixed analysis of instantaneous gas properties and qualitative comparisons. An as of yet unperformed comparison of the thermal history of accreted gas (using tracer particles) is required to demonstrate quantitative agreement between these two methods.

2.4.1 Numerical Uncertainties and Included Physics

Particularly when studying gas accretion, a perceived strength of SPH, embodied in codes such as GADGET, is what should accurately be termed its pseudo-Lagrangian nature (Vogelsberger et al., 2012). The consequence of this behavior is the apparent ease by which the thermodynamic properties of individual mass elements can be traced through time. However, consider the example of mass represented by a single SPH particle in a shearing flow, as illustrated in Fig. 20 of Vogelsberger et al. (2012). This mass is initially drawn from a volume centered on the SPH particle with radius of order the local smoothing length. Shearing of this volume deforms its shape and the associated mass should follow. However, the nature of SPH is such that this mass is forced to remain bound to its SPH particle. The consequence is that the evolution of the fluid is not fully consistent with the equations of motion and cannot then be truly Lagrangian. In this regard AREPO is also not strictly Lagrangian since the gas cells are not allowed to become arbitrarily distorted. However, the fundamental difference is that mass exchange between cells occurs in a manner consistent with the equation of mass conservation and we can therefore term the scheme quasi-Lagrangian, because unlike in SPH the solution is faithful to the underlying equations of motion.

It is also important to note the differences arising due solely to different formulations of SPH. Kereš et al. (2005) used TreeSPH (Hernquist & Katz, 1989b) and found a strongly bimodal histogram of the maximum past temperature of accreted gas, and comparable contributions of the hot and cold modes. Kereš et al. (2009) used GADGET-2 and concluded that at $z \geq 2$ the accretion is dominated by the cold accretion mode at all halo masses, with little evidence for bimodality in the same $T_{\rm max}$ histogram. The original

motivation for a constant threshold temperature used to separate hot and cold modes is in fact the location of the temperature minimum resulting from this bimodality. Yet, this same constant value has been used in virtually all simulation based investigations of cold mode accretion since.

Yoshida et al. (2002) compared these two SPH formulations and concluded that the TreeSPH technique leads to a significant over-cooling stemming from its geometrical symmetrization of hydrodynamical forces. In the context of cooling from halo gas this problem manifests as increased hot mode accretion. Accordingly, Kereš et al. (2009) attributes the much higher hot mode accretion rates in their earlier simulations to numerical deficiencies in the older SPH formulation. The larger relative contribution of high $T_{\rm max}$ gas and the dominant contribution of hot gas accretion in the most massive halos found in Kereš et al. (2005) is in fact in better agreement with this work than that found in Kereš et al. (2009). However, this agreement is a coincidental result of the aforementioned numerical inaccuracies in TreeSPH.

It is not altogether surprising, then, that our moving mesh results with AREPO are in tension with previous studies, particularly SPH simulations made using the standard formulation found in GADGET. Significant uncertainties remain in the numerical simulation of cosmological volumes where the inclusion of a diverse number of physical processes over a large dynamic range in spatial and temporal scales is both computationally intractable and yet phenomenologically required. What is surprising is the extent to which differences arise solely due to the numerical scheme employed to solve the hydrodynamics, and not to differences in sub-grid prescriptions, star formation recipes, feedback implementations, and the like. The simulations presented in this work implement a basic set of physical processes in addition to gravity and hydrodynamics.

We leverage this relative simplicity to make a clean comparison between the SPH and moving mesh techniques. However, it is probable that strong feedback from stars or active galactic nuclei could significantly alter the nature of gas inflow from the IGM, and this remains an open question for future work. The moving mesh implementations of ideal magnetohydrodynamics (Pakmor et al., 2011) and physical viscosity (Muñoz et al., 2012) in AREPO will likewise enable us to investigate these potentially important effects for cosmological gas accretion.

2.4.2 Conclusions

We conclude with two main points. Firstly, that the constant temperature threshold commonly used to study cosmological gas accretion is only reasonable above some minimum halo mass, and application in an overly broad context biases conclusions regarding the relative importance of hot or cold gas accretion. Secondly, that numerical deficiencies inherent to smoothed particle hydrodynamics (SPH) simulations non-trivially modify the relative contributions of hot and cold mode accretion, under any definition. An identical analysis of gas accretion utilizing our Monte Carlo tracer scheme with the new moving mesh code AREPO demonstrates significant physical differences in the thermal history of accreted material. We summarize our primary results as:

- In agreement with previous work, GADGET simulations imply that at $z = 2$ only a small fraction of gas in centrally forming galaxies of massive halos above $10^{11}\,\mathrm{M}_\odot$ heats to an appreciable fraction of the virial temperature during accretion. In AREPO we find a decrease in the accretion rate of cold gas, by a factor of ~ 2 at $M_{\rm halo} \simeq 10^{11}\,\mathrm{M}_\odot$, as well as a significantly enhanced accretion rate of hot gas,

by an order of magnitude at the same halo mass. These discrepancies increase for more massive systems. We attribute the decrease in the cold accretion rate primarily to the large population of numerical "blobs" which efficiently deliver cold gas to central galaxies in our SPH simulations, but are completely absent in our moving mesh calculation. The increase in the hot accretion rate is dominated by significantly more efficient cooling from halo gas in AREPO, where spurious heating from the dissipation of turbulent energy on large scales prevents the same behavior in GADGET.

- We argue that comparison of the maximum past temperature $T_{\rm max}$ of a gas element to a fixed temperature threshold T_c makes physical sense only for halos with $T_{\rm vir} \gg T_c$. For lower mass systems the past temperature history should instead be compared to some fraction of the virial temperature of the dark matter halo. However, at sufficiently low halo masses, when the virial temperature becomes comparable to the gas temperature in the IGM, neither a constant temperature threshold nor one scaled with the virial temperature are sufficient to probe gas shock heating and virialization.

- We observe that the "transition mass" from cold to hot dominated accretion which has been reported between $10^{11-11.5}\,{\rm M}_\odot$ is a consequence of the constant temperature criterion. When comparing the thermal history of gas instead to a fraction of $T_{\rm vir}$ we find no sharp transition, only a gradual decline from $f_{\rm cold} \simeq 20\%$ to $\simeq 0\%$ over the mass range from $M_{\rm halo} = 10^{10}\,{\rm M}_\odot$ to $10^{12}\,{\rm M}_\odot$.

- The filamentary geometry of accreting gas near the virial radius is a common feature of massive halos above $M \sim 10^{11.5}$ at high redshift ($z = 2$). Although

characterized by the same large scale morphology, filamentary gas structures in GADGET tend to either remain cold and flow coherently to small radii within a halo, or artificially fragment and form a large number of "blobs" which are purely numerical in origin. In contrast, filaments in AREPO simulations are more diffuse and experience significant heating at comparable radii. The geometry and angular covering factor of material accreted with $T_{\rm max} \geq T_{\rm vir,acc}$ indicates that coherent, filamentary flows associated with large scale IGM filaments contribute significantly to hot accretion rates in these massive systems.

Chapter 3

The impact of feedback on cosmological gas accretion

D. Nelson, S. Genel, M. Vogelsberger, V. Springel, D. Sijacki, P. Torrey, L. Hernquist, *Monthly Notices of the Royal Astronomical Society*, Vol 448, Issue 1, p. 59-74, 2015

Abstract

We investigate how the way galaxies acquire their gas across cosmic time in cosmological hydrodynamic simulations is modified by a comprehensive physical model for baryonic feedback processes. To do so, we compare two simulations – with and without feedback – both evolved with the moving mesh code AREPO. The feedback runs implement the full physics model of the Illustris simulation project, including star formation driven

galactic winds and energetic feedback from supermassive blackholes. We explore: (a) the accretion rate of material contributing to the net growth of galaxies and originating directly from the intergalactic medium, finding that feedback strongly suppresses the raw, as well as the net, inflow of this "smooth mode" gas at all redshifts, regardless of the temperature history of newly acquired gas. (b) At the virial radius the temperature and radial flux of inflowing gas is largely unaffected at $z = 2$. However, the spherical covering fraction of inflowing gas at $0.25\,r_{\rm vir}$ decreases substantially, from more than 80% to less than 50%, while the rates of both inflow and outflow increase, indicative of recycling across this boundary. (c) The fractional contribution of smooth accretion to the total accretion rate is lower in the simulation with feedback, by roughly a factor of two across all redshifts. Moreover, the smooth component of gas with a cold temperature history, is entirely suppressed in the feedback run at $z < 1$. (d) The amount of time taken by gas to cross from the virial radius to the galaxy – the "halo transit time" – increases in the presence of feedback by a factor of $\simeq 2$ - 3, and is notably independent of halo mass. We discuss the possible implications of this invariance for theoretical models of hot halo gas cooling.

3.1 Introduction

Over the past decade, numerical simulations modelling the evolution of gas in a ΛCDM cosmology have made it clear that the process by which galaxies acquire their baryons across cosmic time evades a satisfactory understanding. In the "classical" theory of collapse and virialisation, gas from the intergalactic medium shock-heats to the virial temperature of a dark matter halo, subsequently forming a hot, pressure supported

atmosphere in approximate equilibrium (Rees & Ostriker, 1977; Silk, 1977; White & Rees, 1978). The timescale of energy loss from radiative cooling determines the rate at which gas can cool into the halo centre (White & Frenk, 1991), and indeed whether or not a stable virial shock can exist at all (Birnboim & Dekel, 2003a).

Numerical simulations over the past decade (starting with Katz et al., 2003; Abadi et al., 2003; Kereš et al., 2005) have shown that gas accretion in the cosmological context, without the luxury of spherical symmetry, is significantly more complex. They found that (i) coherent streams of gas can provide strong fuelling for star formation over relatively small solid angles of the virial sphere, and (ii) such streams could potentially remain cold and avoid heating up to the virial temperature, even for haloes massive enough to support a quasi-static hot atmosphere. Such flows are a natural consequence of the "cosmic web" of large scale structure, particularly at high redshift ($z > 2$), and provide an intriguing avenue for gas accretion distinct from classic hot halo cooling. Aspherical gas inflow has been connected to many key questions in galaxy formation, including the growth of the stellar populations of galaxies (e.g. Oppenheimer et al., 2010), their morphological transformations (Sales et al., 2012; Cen, 2014), kinematics (Genel et al., 2012), and star formation properties (Gabor & Bournaud, 2014; Sánchez Almeida et al., 2014). It has also been looked at in terms of the acquisition of angular momentum (Danovich et al., 2012; Stewart et al., 2013; Danovich et al., 2014), including the connection to preferred directions imposed by the filaments of large scale structure (Dubois et al., 2014), and the feeding of supermassive blackhole accretion (Dubois et al., 2012; Bellovary et al., 2013; Feng et al., 2014).

Smoothed particle hydrodynamics (SPH) simulations were the first to address the question of gas accretion modes in a cosmological context. Leveraging the quasi-

Lagrangian nature of the numerical scheme, they could use the past temperature history of each gas element to differentiate between hot and cold mode accretion. It was found that the maximum past temperature of smoothly accreted gas is bimodal, the dominant contribution arising from gas which has never experienced significant heating during infall to the galaxy (Kereš et al., 2005).

However, caution is warranted. In our previous study (Nelson et al., 2013) we compared the outcome of "classic" SPH simulations and those run with the newer moving mesh code AREPO. At $z=2$ we found a decrease in the accretion rate of cold gas, by a factor of ~ 2 at $M_{\rm halo} \simeq 10^{11}$ M$_\odot$. We also found, at this same mass, an order of magnitude larger accretion rate of gas with significant past heating. These discrepancies grew even more significant for more massive haloes. We attributed the drop in the cold accretion rate to a large population of numerical "blobs" (Torrey et al., 2012) which efficiently deliver cold gas to central galaxies in the SPH simulations, but are completely absent in the moving mesh calculations. The increase in the hot accretion rate was dominated by more efficient cooling from halo gas in AREPO, where spurious heating from the dissipation of turbulent energy on large scales prevents the correct behaviour in SPH (Bauer & Springel, 2012). Filamentary flows in the AREPO haloes were found to be warmer and more diffuse, and did not generally persist as strongly to small radii.

Grid-based adaptive mesh refinement (AMR) simulations have also been interpreted as being in agreement with respect to the importance of a filamentary, cold accretion mode in massive systems at high redshift (Ocvirk et al., 2008; Dekel et al., 2009; Agertz et al., 2009), albeit with two potentially important limitations. In particular, these studies have often not used any form of Lagrangian tracer to follow the thermal and dynamical history of accreting gas, which would permit a more direct comparison with

particle hydrodynamics codes. With the notable exception of Ocvirk et al. (2008), they have also generally focused on targeted, "zoom" simulations of individual haloes, where the significant halo-to-halo variation between codes (e.g. Vogelsberger et al., 2012) can make broadly applicable conclusions difficult. Regardless of hydrodynamical method, robust conclusions are difficult to draw from a single simulated halo, or a small sample of such haloes.

These findings made it clear that numerical deficiencies in the standard formulation of SPH used in past studies significantly biased previous quantitative conclusions as to the relative importance of cold streams or cold mode accretion (also problematic are issues with formal numerical convergence, as discussed in Zhu et al., 2014). However, in this previous comparison work we included only a simple model for baryonic physics, appropriate to make an even-handed comparison to past work, but lacking the physical fidelity of modern cosmological simulations. The large question left outstanding was then: what impact, if any, does feedback associated with galaxy formation have on the process of gas accretion. Current state-of-the-art cosmological simulations have reached the point where they can evolve a significant volume of the universe down to $z=0$, while simultaneously resolving the structure of individual galaxies, and reproducing a broad range of observational constraints (Khandai et al., 2014; Vogelsberger et al., 2014b; Schaye et al., 2015). One of the many investigations related to galaxy formation and evolution that they enable is a study of baryonic accretion.

At the outset, it would seem entirely plausible that the accretion rates and the inflow of gas from the intergalactic medium would be quite sensitive to feedback processes. The impact could be either direct or indirect, or both. For instance, inflowing streams could be disrupted by spatially coincident outflows, such that the net mass flux entirely

reversed direction. Or, energy injection from feedback could heat up the the surrounding hot halo gas, leading to a modification of the thermal history of inflow due to mixing. Alternatively, it would also seem plausible that feedback could have relatively little effect. For instance, star formation driven winds with non-isotropic outflow may simply evolve to occupy different regions of the virial volume than inflowing streams.

Recently, simulations have begun to investigate the additional complexity when feedback and galactic scale outflows are included (beginning with Oppenheimer et al., 2010; Faucher-Giguère et al., 2011; van de Voort et al., 2011b). Conclusions as to their impact have been somewhat mixed, which undoubtedly arise from a combination of different feedback implementations, numerical methods, contexts, and interpretations. In brief review, Brooks et al. (2009) simulated five haloes, included a delayed cooling, supernova blastwave feedback model, but did not explicitly consider the impact of the feedback in GASOLINE, and in general found results consistent with Kereš et al. (2005). Oppenheimer et al. (2010) included a kinetic galactic wind model in cosmological SPH simulations and concluded that recycled gas accretion is in fact the dominant accretion mechanism at $z \leq 1$, with minimal effect on high redshift accretion. Faucher-Giguère et al. (2011) included a constant velocity galactic wind model, finding that net accretion rates measured as instantaneous mass fluxes could be substantially affected. van de Voort et al. (2011b) found that while the gas accretion rates onto haloes was relatively robust against the presence of feedback, the rates onto galaxies themselves depended sensitively on stellar winds as well as metal-line cooling. van de Voort et al. (2011a) studied the impact of AGN feedback on inflow and found that it preferentially prevented hot mode gas, with high maximum past temperature, from cooling from the halo onto the galaxy. Stewart et al. (2011) simulates two relatively massive haloes, including the

supernova blastwave feedback model, but notes this has little impact at the simulated mass scale, and does not consider how it modifies gas accretion nor AGN effects. Dubois et al. (2013) investigated AGN feedback at high redshift in one halo, and found that large-scale hot superwinds could morphologically disturb cold filaments and quench cold diffuse accretion. Murante et al. (2012) used a thermal supernova feedback scheme, and found that additional heating of cold inflow due to this feedback gave rise to a significant accretion rate of intermediate temperature gas. Most recently, Woods et al. (2014) includes a combined delayed cooling supernova and early stellar feedback model, finding that overall gas accretion rates did not change with strong feedback, while the balance between cold and hot components did. Finally, Übler et al. (2014) implements a hybrid thermal/kinetic stellar feedback scheme and finds strong outflows generate substantially higher raw accretion rates, and that recycled material dominates galactic gas accretion at $z<1$.

As a caveat, we note that the same numerical issues explored in the "moving mesh cosmology" series (Vogelsberger et al., 2012; Sijacki et al., 2012; Kereš et al., 2012; Nelson et al., 2013) that compromised the accuracy of SPH studies of gas inflow will also affect the interaction of outflowing ejecta and wind material with both halo and filamentary gas. Further, because hot gas in our AREPO simulations cools more efficiently than in classical SPH, the energy input from feedback required to prevent over-cooling is even larger than in previous simulations, implying that the impact on gas accretion could be significantly altered. We are therefore motivated to extend previous investigations with the current study, which combines a comprehensive, validated feedback model with an accurate and robust numerical technique in a systematic comparison.

This paper contrasts two simulations, realisations of the same initial conditions

evolved with the moving mesh code AREPO. We compare populations of haloes and galaxies across cosmic time, contrasting the state and history of accreting gas between two runs, with and without feedback. In Section 4.2 we describe the simulation technique and analysis methodology. Section 3.3 addresses the rate and history of primordial gas accretion, while Section 3.4 compares the instantaneous state of gas in haloes. Section 3.5 considers the timescale of accretion through the halo. Finally, Sections 4.6 and 5.7 discuss our results and their implications, and summarise our conclusions.

3.2 Methods

3.2.1 The Simulations

In this work we compare two simulations, "with" and "without" feedback, which have several common features. Both employ the AREPO code (Springel, 2010a) to solve the problem of ideal continuum hydrodynamics coupled with self-gravity. An unstructured, Voronoi tessellation of the simulation domain provides a spatial discretization for Godunov's method with a directionally un-split MUSCL-Hancock scheme (van Leer, 1977) and an exact Riemann solver, yielding second order accuracy in space. Since the mesh generating sites can be allowed to move, herein with a velocity tied to the local fluid velocity modulo mesh-regularization corrections, this numerical approach falls under the Arbitrary Lagrangian-Eulerian (ALE) class. Gravitational forces are handled with the split Tree-PM approach, whereby long-range forces are calculated with a Fourier particle-mesh method, medium-range forces with a hierarchical tree algorithm (Barnes & Hut, 1986), and short-range forces with direct summation. A local,

predictor-corrector type, hierarchical time stepping method yields second order accuracy in time. Numerical parameters secondary to our current investigation – for example, related to mesh regularization or gravitational force accuracy – are detailed in Springel (2010a) and Vogelsberger et al. (2012), and are unchanged between the two simulation sets.

Both simulations evolve the same initial condition, a random realisation of a WMAP-7 consistent cosmology ($\Omega_{\Lambda,0} = 0.73$, $\Omega_{m,0} = 0.27$, $\Omega_{b,0} = 0.045$, $\sigma_8 = 0.8$ and $h = 0.7$) in a periodic cube of side-length $20h^{-1}$ Mpc $\simeq 28.6$ Mpc, from a starting redshift of $z = 99$ down to $z = 0$. Each includes 512^3 dark matter particles, an equal number of initial gas cells, and a minimum of 5×512^3 tracers (discussed below). The mean baryon mass is 1.1×10^6 $M_\odot$, and the dark matter particle mass is 5.3×10^6 $M_\odot$. The Plummer equivalent comoving gravitational softening lengths for dark matter and stars are 1.4 kpc, and gas cells have adaptive softening lengths equal to 2.5 times their volume-equivalent spherical radius. A redshift-dependent, spatially uniform, ionizing UV background field (Faucher-Giguère et al., 2009) is included as a heating source. Star formation and the associated ISM pressurisation from unresolved supernovae are included with an effective equation of state modelling the ISM as a two-phase medium, following Springel & Hernquist (2003). Gas elements are stochastically converted into star particles when the local gas density exceeds a threshold value of $n_{\rm H} = 0.13$ cm^{-3}. All of the simulations considered in this work disregard the possible effects of radiative transfer, magnetic fields, and cosmic rays.

The **no feedback runs** ("noFB") with "simple physics" additionally account for optically thin radiative cooling assuming a primordial H/He ratio (Katz et al., 1996). They do not include metal line cooling, any resolved stellar feedback that would drive

galactic-scale wind, nor any treatment of black holes or their associated feedback. This is the same simulation presented in Nelson et al. (2013), where it was used in comparison to GADGET (SPH) results.

The **feedback runs** ("FB") implement, unchanged, the fiducial physical model and associated parameter values of the Illustris simulation (Vogelsberger et al., 2014b,a; Genel et al., 2014) applied to the same initial conditions as our previous work, allowing object by object comparison. Complete details of the physics included in the model, as well as its tuning and validation, are described in Vogelsberger et al. (2013) and Torrey et al. (2014). We describe here in some detail those aspects of the model which most strongly influence gas accretion.

First, we include the radiative cooling contribution from metal lines, where heavy elements are produced from supernovae Ia/II and AGB stars in stellar population evolution modelling (Thielemann et al., 1986; Portinari et al., 1998; Karakas, 2010). In the absence of additional heating sources, this can increase the cooling rate $\Lambda(n, T, Z, \Gamma)$, for solar metallicity by an order of magnitude between $10^{4.5}\,\mathrm{K} < T_{\mathrm{gas}} < 10^{6.5}\,\mathrm{K}$, mainly due to the contribution of O, Ne and Fe (Sutherland & Dopita, 1993; Wiersma et al., 2009a), enhancing cooling from the hot halo and the buildup of stellar mass.

To balance efficient cooling in the simulations, further exacerbated by metals, we include energetic feedback from star formation driven winds as well as supermassive black holes – dominant in haloes less and more massive than $\mathrm{M}_\star$, respectively. Stellar winds are generated directly from star-forming gas, with velocity $v_w = 3.7\sigma_{\mathrm{dm}}$ where σ_{dm} is the local 1D dark matter velocity dispersion, which scales with the circular velocity maxima and so host (sub)halo total mass. In practice, gas cells are probabilistically

converted into a wind-phase gas cell/particle, which interacts gravitationally but not hydrodynamically, until it reaches either a density threshold or a maximum travel time. Specifically, 0.05 times the star formation threshold in density, or 0.025 times the current Hubble time. This typically occurs just outside the disk – that is, deep within the halo. At this point, the mass, momentum, metals, and internal energy are deposited into the gas cell in which the wind particle is located.

The energy-driven wind has a mass loading factor $\eta_w \propto v_w^{-2}$ (thereby decreasing with halo mass) and is assigned a "bipolar" outflow direction, given by the cross product of its original velocity and the gradient of the local potential. Since the wind speed is in general slightly less than the escape speed, this implementation can generate a strong "galactic fountain" effect of recycled material returning to the galaxy (Oppenheimer et al., 2010; Davé et al., 2012). As a result, it can also strongly modify the temperature and velocity structure of gas in the inner halo.

The second form of feedback originates in supermassive black holes (SMBHs, following Sijacki et al., 2007), which are seeded with a mass of 1.4×10^5 $M_\odot$ in massive haloes above 7×10^{10} $M_\odot$ and are effectively sink particles which grow through gas accretion and merging. When accretion onto the black hole is below 0.05 of the Eddington rate, a radio-mode model injects highly bursty thermal energy equal to 0.07 of the accreted rest mass energy in large ($\simeq 50$ kpc) bubbles. When the accretion rate is above this fraction, a quasar-mode model injects thermal energy into nearby gas cells, with a lower coupling efficiency of 0.01 and a smoother time profile. For accretion rates approaching Eddington, a third, radiative form of feedback modifies the cooling rate for gas in the vicinity of the BH, assuming an optically-thin $1/r^2$ attenuation. In the halo mass regime where BH feedback becomes important, the radio-mode channel generates

high velocity, high temperature outflows which influence gas at larger radii than the stellar feedback driven winds – out to the virial radius and into the intergalactic medium.

There are three other minor changes with respect to the noFB configuration, required to match the fiducial Illustris model. First, we modify the equation of state parameter (Springel & Hernquist, 2003) from $q = 1$ to $q = 0.3$, which interpolates between the effective EOS and an isothermal EOS of 10^4 K with weights of 0.3 and 0.7, respectively, in order to avoid over pressurising the ISM of star-forming gas. Secondly, we decouple the comoving gravitational softening lengths of the gas and stars from that of the dark matter at $z = 1$, allowing the former to decrease by a further factor of two down to $z = 0$. Finally, we include a correction for self-shielding of dense gas from the UV background (Rahmati et al., 2013).

3.2.2 Monte Carlo Tracers

Both simulations include our new "Monte Carlo tracer particle" technique (Genel et al., 2013) in order to follow the evolving properties of gas elements over time. This is a probabilistic method, where tracers act as unique tags in association with parent cells or particles. They have no phase space coordinates, but are instead exchanged between parents based explicitly on the corresponding mass fluxes. By locating a subset of their unique IDs at each snapshot we can, by reference to the gas cells in which the tracers reside (their parents), reconstruct their spatial trajectory or thermodynamic history. Furthermore, at each computational timestep every active tracer updates a record of its maximum temperature, density, Mach number and entropy, as well as the time of these events, enabling us to investigate these values with timestep-level resolution.

We extend the Monte Carlo tracer approach to include mass transfer between all baryonic components present in the simulations in a fully self-consistent manner. That is, tracers can reside in gas cells, star particles, wind-phase cells, and black holes, and exchange between these components in exactly the same ways that baryonic mass is exchanged during the simulation. In particular:

1. Gas cell to gas cell transfer via finite volume fluxes, refinement and derefinement.
2. Gas cell to star particle via star formation, and the reverse during stellar mass return.
3. Gas cell to/from wind-phase during the generation/recoupling of star formation driven galactic winds.
4. Gas cell to black hole as a result of BH accretion, and between two black holes during a merger.

This allows us to follow the flow of mass through all baryonic phases which are present. Finally, to explore the role of stellar feedback driven winds and recycling, each tracer also records the last time of exchange to a star/wind particle from gas, and the reverse, as well as a counter of the number of times it has been incorporated into a wind.

3.2.3 Post-processing

We identify dark matter haloes and their gravitationally bound substructures using the SUBFIND algorithm (Springel et al., 2001a; Dolag et al., 2009) which is applied on top of a friends-of-friends cluster identification. We refer to the most massive substructure

in each FoF group as the halo itself, and consider accretion onto such haloes and the central galaxies hosted therein. We follow the evolving positions and properties of haloes over time by constructing a basic merger tree as in Nelson et al. (2013), where only the "main progenitor branch" (MPB) is needed. For each halo, we restrict our analysis of accretion to the time period over which this main branch is robustly determined.

For each of a finite number of analysis redshifts, spanning $z=0$ to $z=5$, we perform a set of independent, identical analysis tasks. For all tracers (of all parent types) in all haloes at that redshift, we walk backwards and record the most recent time and direction of several particular radial crossings. We take $r_{\rm vir} = r_{\rm 200,crit}$ the radius enclosing a mean overdensity 200 times the critical density. We label the inward crossing times of two important radii, $0.15\,r_{\rm vir}$ (representative of the outer boundary of the galaxy), and $1.0\,r_{\rm vir}$ (representative of the outer boundary of the halo) the "most recent incorporation time" into the galaxy and the "most recent accretion time" into the halo, respectively. For these same two radii, we also record the earliest – that is, highest redshift – such crossing. We label the $0.15\,r_{\rm vir}$ crossing as the "first incorporation time", and the $1.0\,r_{\rm vir}$ crossing as the "first accretion time". We discuss the calculation of accretion rates based on these quantities in the following section.

Each tracer also records the virial temperature of its parent halo at the time of first accretion, labelled $T_{\rm vir,acc}$. We compare this value against $T_{\rm max}$, the maximum temperature of a tracer between the start of the simulation and the time of its first incorporation. This corresponds to the time at which each tracer accretes into the direct main progenitor of the central galaxy of the halo. As a result, the procedure is sensitive to virial shock heating in the MPB, as well as virial shock and feedback related heating in satellites prior to first incorporation, but not heating due to wind recycling after

incorporation into the central galaxy.

In addition, we also separate all accreted tracers into one of three disjoint "modes" of accretion: smooth, clumpy, or stripped, according to the following definitions applied at the time of first accretion. Smooth: not a member of any resolved substructure, other than a MPB halo, and likewise at all previous times. Clumpy: gravitationally bound to any resolved substructure which is not a MPB halo. Stripped: otherwise smooth, but gravitationally bound to some resolved substructure other than the MBP at any previous time. In addition to these three modes, each tracer is also flagged as being "recycled" if it has ever been part of a stellar feedback driven wind at some previous time.

We note that identification of the "clumpy" component – that is, the merger contribution – includes only *resolved* substructure. The combination of simulation resolution and group finder implies a minimum halo mass which can be identified. For the conclusions presented herein, this is $M_{\rm halo} \geq 10^{7.8}$ $M_\odot$, meaning that for the halo mass range we consider all mergers with mass ratio above $\simeq 10^{-3}$ are correctly identified as a clumpy mode. The contribution of unresolved mergers below this level is expected to be negligible (see Kereš et al., 2005; Genel et al., 2009), and we have further verified that the associated conclusions are qualitatively unchanged in an identical simulation with a factor of eight lower mass resolution.

3.3 Rate and Mode of Gas Accretion

In this section we compare the rate and mode of cosmological gas infall across cosmic time between the two simulations, with and without our fiducial feedback model. The

presence of star formation driven galactic winds and AGN feedback, particularly in the radio-mode, generates significant radial velocity in the baryonic component of haloes. Outflows with high outward bulk motion in turn trigger "galactic fountain" behaviour, with significant gas mass returning after some time delay with high inward radial velocity. This efficient recycling through the galaxy itself implies that any instantaneous or quasi-instantaneous measurement of the total accretion rate of material into the galaxy will be a (possibly large) overestimate of the "net" rate of material permanently joining the galaxy by, for example, forming stars. Furthermore, this recycled material is a source of accretion onto galaxies in addition to direct cosmological accretion, which does not undergo a similar phase of processing and metal enrichment.

We are first interested in the question of whether the presence of strong galactic-scale feedback modifies the rate or character of cosmological – or "primordial" – infall. Therefore, to make a sensible comparison between simulations with and without the additional motions induced by outflows, we adopt the following approach. At any given redshift, accretion rates are measured as a tracer flux over a specific time interval, restricted to those baryons which are entering or leaving the galaxy, or its parent along the main progenitor branch, for the first time. We fix this time interval to be 250 Myr, but our results are not sensitive to this choice, so long as it is not too short to run up against the finite snapshot spacing, and not too long as to effectively smooth over changing physical conditions. In practice, we calculate three rates related to baryonic fluxes with respect to the galaxy:

1. The outflow rate as the number of tracers with outward crossings through $0.15\,\mathrm{r_{vir}}$ during this time window.

2. The raw (as opposed to net) inflow rate as the number of tracers with inward crossings across this same boundary during this time window.

3. The net accretion rate as the outflow rate subtracted from the inflow rate. That is, the difference of the number of tracers with inward $0.15\,\mathrm{r_{vir}}$ crossing times and outward $0.15\,\mathrm{r_{vir}}$ crossing times during this time window.

In each case the tracer count is then multiplied by the tracer effective mass (the initial gas cell mass divided by the initial number of tracers per cell) and normalized by the time window to derive a rate. This method gives a measurement of $\dot{M}_{\rm gas}^{\rm prim}(M_{\rm halo}, z, {\rm dir}, {\rm mode})$ for $M_{\rm halo} \in [10^9, 10^{12}]$ $\mathrm{M_\odot}$, $z \in [0, 5]$, dir $\in \{{\rm in}, {\rm out}, {\rm net}\}$ and mode $\in \{{\rm all}, {\rm smooth}, {\rm clumpy}, {\rm stripped}, {\rm recycled}\}$ where "prim" denotes the primordial qualification. Alternatively, in Section 3.6.1 we discuss another approach where the requirement on first incorporation is relaxed; instead we can calculate fluxes using the most recent incorporation time, thereby measuring a sum of material with two origins – primordial as well as having previously cycled through the MPB.

3.3.1 Galactic accretion as a function of redshift

In Fig. 3.1 we compare the net smooth accretion rate between the two simulations, restricted to haloes in the mass range $11.3 < \log(M_{\rm halo,tot}/\mathrm{M_\odot}) < 11.4$ which we keep constant as a function of redshift. At $z = 5$ the virial temperature of these haloes is $\simeq 10^6\,\mathrm{K}$, decreasing to $\simeq 10^{5.3}\,\mathrm{K}$ at $z = 0$. In this box, the mass selection contains 3 haloes at $z = 5$, 42 at $z = 2$, and 53 at $z = 0$. Comparing the two simulations, we find that feedback suppresses smooth cosmological accretion by a factor of $\simeq$3 at high redshift ($z = 5$). This increases to $\simeq$10 by $z = 1$ after which the suppression continues to increase

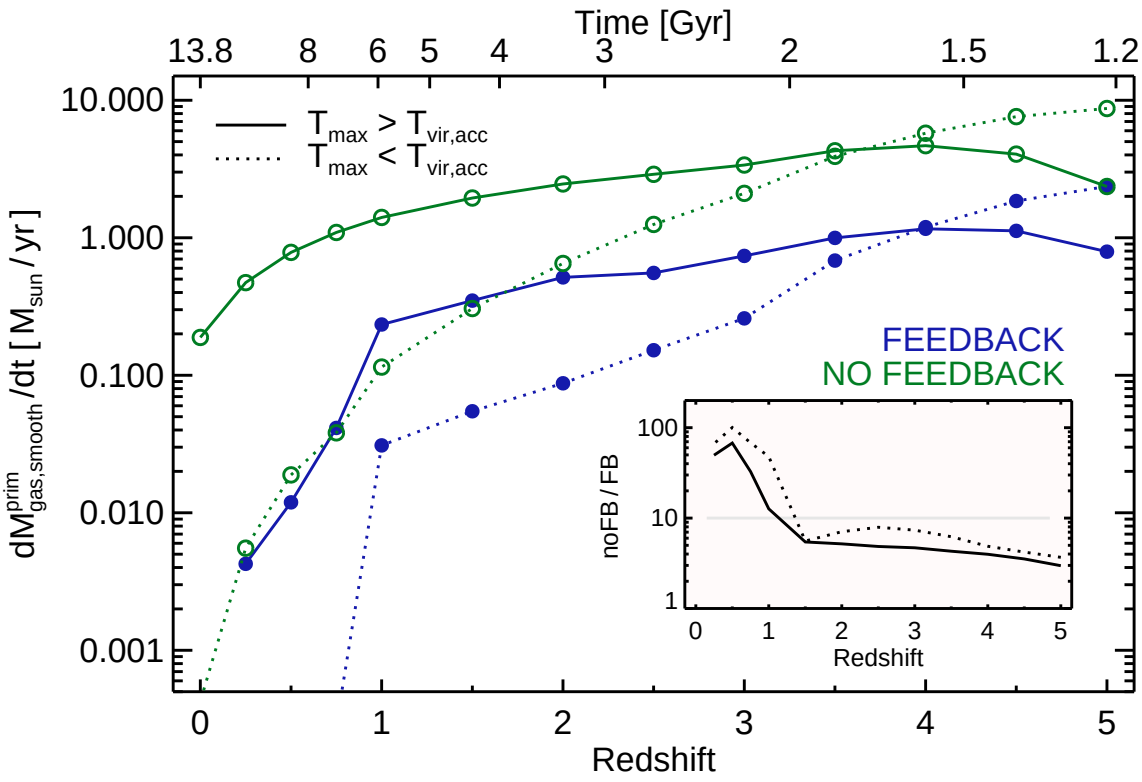

Figure 3.1 The net smooth gas accretion rate of cosmological origin onto central galaxies as a function of redshift. At each redshift accretion over a time window of $\Delta t = 250\,\mathrm{Myr}$ contributes. We include here only haloes in the mass range $11.3 < \log\left(M_{\mathrm{halo,tot}}/\mathrm{M}_{\odot}\right) < 11.4$, and separate the contribution based on the comparison of the maximum past temperature T_{max} of each Monte Carlo tracer to the virial temperature of future parent halo, at the time of accretion. Both simulations, with (blue) and without (green) our fiducial feedback implementation, indicate that relatively cold gas dominates the primordial infall onto galaxies only at redshifts $\gtrsim 4$, while at later times the reverse is true. At this mass scale, the impact of feedback is reduce the total rate of smooth cosmological accretion independent of the T_{max} comparison, by a redshift-dependent amount (see inset).

towards redshift zero, while the individual net rates also steadily decline. Indeed, the net smooth accretion rates drop by roughly an order of magnitude from $z=1$ to $z=0$ even in the noFB run, and this drop is also evident if we consider accretion over all modes, not just smooth. The star formation rate (SFR) of haloes in this mass range declines only moderately over this same time (by a factor of $\sim$2, Genel et al., 2014), implying that the component of the late time SFR in these systems supplied by smooth gas accretion is supported predominantly by material which has cycled through the direct progenitor at an earlier time.

We further split the smooth total based on the comparison between $T_{\rm max}$ and $T_{\rm vir,acc}$, available on a per tracer basis (as in Nelson et al., 2013). In this mass regime, the balance of these two components is similar between both runs, where at $z>4$ gas with $T_{\rm max} < T_{\rm vir,acc}$ dominates the smooth cosmological accretion budget, by up to a factor of three at $z=5$. Towards lower redshift gas with $T_{\rm max} > T_{\rm vir,acc}$ instead dominates, by a factor of $\sim$4 at $z=2$ and by several orders of magnitude by $z=0$. As we discuss later in more detail, the similar balance of these two temperature components between the FB and noFB runs implies that the winds have marginal impact on the temperature history of smoothly accreting material, and that the presence of winds does not preferentially prevent material of a particular temperature history. To give a sense of reference, the difference between the mean $T_{\rm max}$ of cold versus hot temperature history gas is $\simeq 0.9$ dex (at $z=5$) and $\simeq 0.6$ dex (at $z=0$).

To understand the declining net accretion rates towards redshift zero, Fig. 3.2 shows the separate measurements of net accretion, outflow, and raw inflow, disregarding the comparison between $T_{\rm max}$ and $T_{\rm vir,acc}$. The "outflow" in the no feedback run arises primarily from dynamical gas motions, particularly due to galaxy-galaxy interactions.

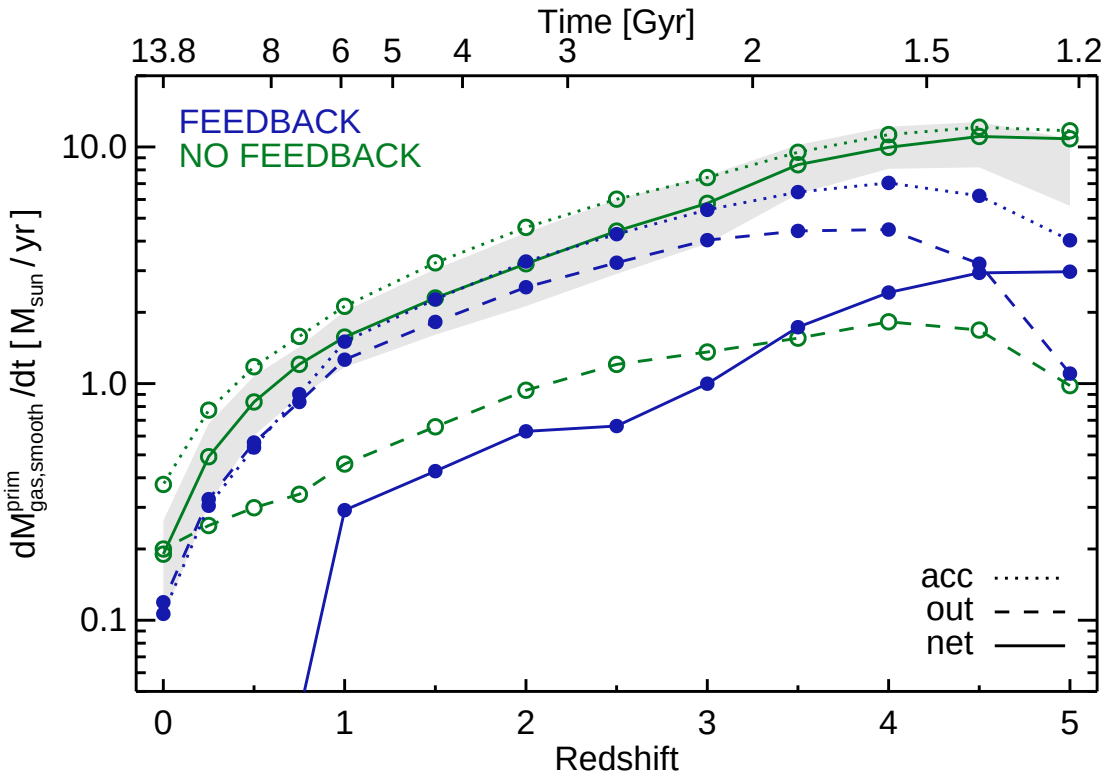

Figure 3.2 The smooth gas accretion rate of cosmological origin onto central galaxies as a function of redshift. At each redshift accretion over a time window of $\Delta t = 250\,\mathrm{Myr}$ contributes. We include here only haloes in the mass range $11.3 < \log\left(M_{\mathrm{halo,tot}}/\mathrm{M}_{\odot}\right) < 11.4$, and consider separately the contribution from net inflow, outflow, and raw accretion. The grey band indicates the upper and lower quartiles about the median for the noFB net line.

Any induced velocity which moves a tracer outside the $0.15\,r_{\rm vir}$ radial boundary will result in a non-zero outflow rate. This situation could arise during a merger or fly-by which generates any of (i) a tidal gas tail, (ii) a re-distribution of gas away from the halo centre, (iii) turbulent gas motions, (iv) or the merging companion itself passing in and subsequently out of this radius. In some sense this level of outflow represents an unrelated contribution (or noise floor) to the measurement of true, feedback induced outflows. In the case of smooth accretion, we see that this contribution is at least a factor of a few below the accretion level in the no feedback case. In the feedback run, however, large outflow rates balance about half of the inflow, this factor increasing towards redshift zero. The result is the suppression of the net smooth accretion rate at all redshifts already noted from the previous figure. Interestingly, the raw inflow rates also differ between the two runs, being suppressed in the FB case by a factor of roughly two, again at all redshifts. We find these same trends in the other accretion modes, as well as in the total accretion, disregarding mode. They also hold for all halo masses down to $\sim 10^{10.5}$ $\mathrm{M}_{\odot}$, below which the noFB simulation begins to show a similar increase in "outflow" as we reach the limit of sufficiently-resolved galaxies.

We first consider the ratio of the net rate of the $T_{\rm max} < T_{\rm vir,acc}$ component to the total primordial net rate, independent of past maximum temperature, $f_{<}/f_{\rm tot}$. This ratio declines monotonically towards redshift zero for all halo masses, such that its maximum value always occurs at the highest redshift for which a given halo mass is present in the simulation volume. We explicitly show this ratio for each mode in Fig. 3.3. Contrasting smooth accretion between the two runs at any given redshift we find that $f_{<}/f_{\rm tot}$ is nearly identical. The simulation with feedback has smaller values at high redshift ($z > 5$), and larger values at low redshift ($z < 2$), the crossover occurring between $2 < z < 3$ at all

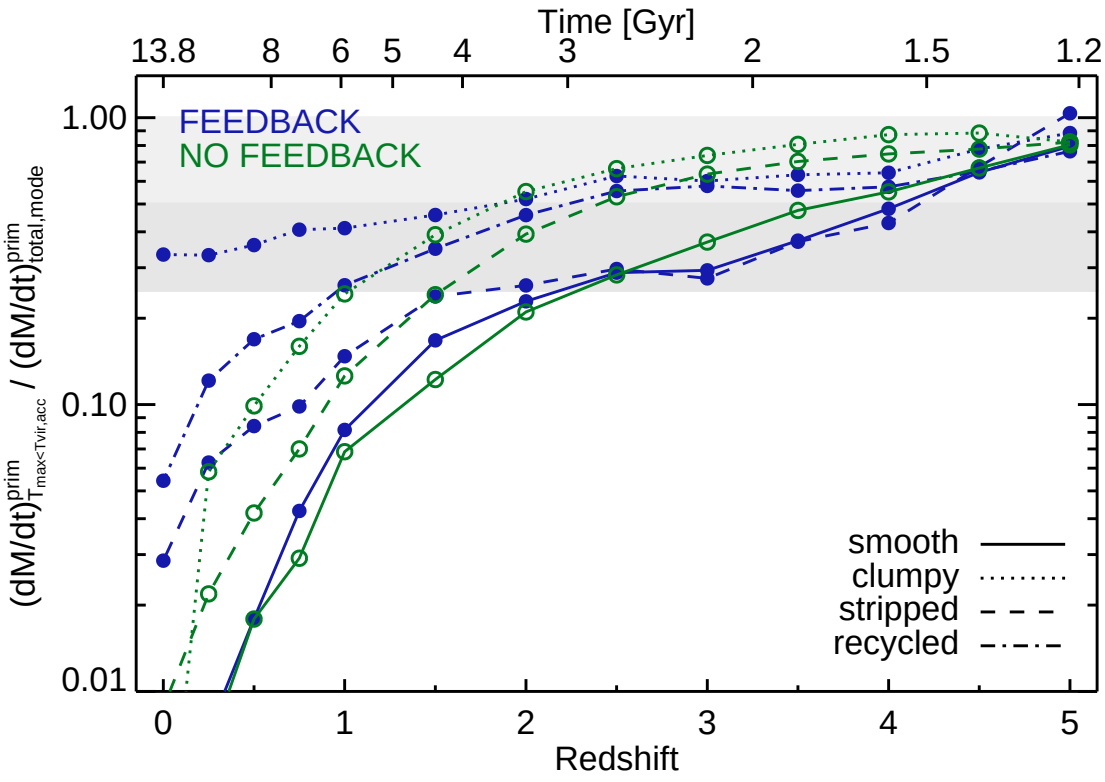

Figure 3.3 Ratio of accretion with $T_{\rm max} < T_{\rm vir,acc}$ to the total primordial rate, regardless of $T_{\rm max}$. We include accreted gas of cosmological origin acquired onto central galaxies over a time window of $\Delta t = 250\,{\rm Myr}$ as a function of redshift. The total inflow is subdivided into smooth, clumpy, stripped, and recycled components, where only gas making a net contribution to the accretion rates is included. Grey bands indicate 25%, 50%, and 100% levels.

halo masses. Therefore, while feedback in these systems strongly suppresses the rate of accretion of smooth primordial gas, it does not directly affect the temperature history of this material. Although gas which flows in cold may undergo additional heating as a result of feedback, this heating is not significant enough to generally increase $T_{\rm max}$ above $T_{\rm vir,acc}$. Furthermore, gas which flows in hot may suffer a lower $T_{\rm max}$ value due to metal-line enhancement of the cooling rates, but not at a level to suppress it below $T_{\rm vir,acc}$. The first point implies that outflowing winds do not disrupt inflowing streams, at least not enough to induce mixing with the surrounding hot halo gas. Nor does the wind material itself incorporate into the streams in a way which modifies their temperature, although here our choice of low thermal energy for stellar winds may be partly responsible and potentially masking such an effect. In the following section we consider whether the presence of outflows increases the time required for gas to inflow from the virial radius to the galaxy. Here we conclude that the temperature history of gas acquired by central galaxies in a smooth mode from the intergalactic medium is largely unmodified by our fiducial feedback model.

Significantly different behaviour is seen for baryonic mass which accretes as part of a resolved substructure as it crosses the virial radius. In both the FB and noFB runs, $f_{<}/f_{\rm tot} \gtrsim 0.5$ for this clumpy material for all redshifts $z > 2$. This ratio drops to nearly zero towards $z = 0$ in the noFB run, indicating that material incorporated into galaxies which entered the halo bound to a satellite nevertheless experiences heating comparable to smoothly accreted gas. In contrast, this ratio only drops to $\simeq 0.4$ by $z = 1$ and $\simeq 0.35$ by $z = 0$ in the FB run. We propose that metal enrichment in satellites prior to accretion into the MPB enables more efficient radiative cooling at later times, preventing a sizeable fraction of this gas from reaching a maximum past temperature as high as its

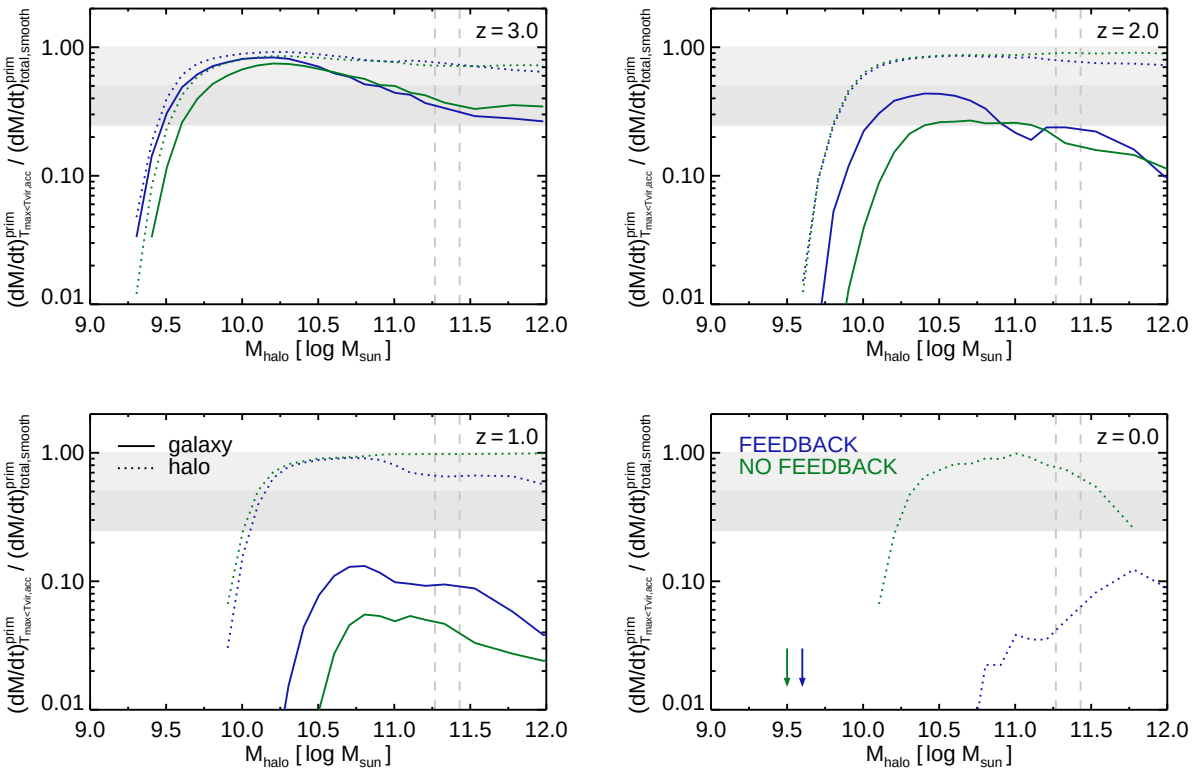

Figure 3.4 Ratio of accretion with $T_{\rm max} < T_{\rm vir,acc}$ to the total primordial rate as a function of halo mass. We include smoothly accreted gas of cosmological origin acquired onto central (solid) galaxies and their parent haloes (dotted) over a time window of $\Delta t = 250\,{\rm Myr}$ ending at four specific redshifts ($z \in \{0, 1, 2, 3\}$). Only gas making a net contribution to the accretion rates is included. The fractional contribution of the colder component onto galaxies evolves downward with redshift largely independent of halo mass, and similarly between the two runs. At low redshift, $z < 1$, the FB run suppresses this fraction for gas accretion onto haloes as a result of significant heating prior to crossing the virial radius. Horizontal grey bands indicate the 25%-50%, and 50%-100% levels, and dotted vertical lines indicate the mass range considered throughout.

counterpart in the noFB run. This would require a gradual heating process, such as the mixing and stripping of satellite gas during its orbit through a larger hot halo, processes which remain largely unexplored in the present work.

A similar though less prominent effect is also evident for stripped material. Recycled material – which was part of a stellar feedback driven wind at some previous time, but could otherwise be part of any accretion mode – has a redshift evolution comparable to that of stripped gas, supporting the idea that one or more epochs of metal enrichment in the ISM of an external galaxy, prior to accretion, can suppress $T_{\rm max}$.

3.3.2 Gas acquisition by the halo

It is notable that for material acquired by the halo itself this same ratio $f_{<}/f_{\rm tot}$ does differ between the simulations. In particular, we consider the net flux of tracers crossing $1.0\,r_{\rm vir}$ for the first time over the past 250 Myr, without necessarily also crossing $0.15\,r_{\rm vir}$. For $z < 1$, in the same halo mass range of $11.3 < \log(M_{\rm halo,tot}/{\rm M}_{\odot}) < 11.4$, we find $f_{<}/f_{\rm tot} \simeq 0.9$ in the run without feedback versus $f_{<}/f_{\rm tot} \simeq 0.6$ in the run including feedback. We expect this fraction to be large in both cases, since the temperature distribution of material near the virial radius is dominated by gas at temperatures lower than the virial temperature. Specifically, the median radial temperature profile of the hot halo gas at $z = 0$, $T(r)/T_{\rm vir}$, decreases with radius to a minimum of $\simeq 0.5$ at $r/r_{\rm vir} = 1$. We then expect that for most gas the ratio $T_{\rm max}$ / $T_{\rm vir,acc}$ will reach a maximum of ~ 0.5 in both runs. On the other hand, the exact fraction of gas satisfying $T_{\rm max}$ ¡ $T_{\rm vir,acc}$ could be lower than unity if the gas temperature distribution at the virial radius extended into a high temperature tail.

This is not true in the FB case, implying that accreting gas has already evolved following a different temperature history prior to interaction with the halo. As a check, we make a simple measurement of the temperature of gas in the intergalactic medium, defined as all gas cells outside of all friends-of-friends groups. Between the FB and noFB runs, the IGM temperature distribution is similar at high redshift ($z > 3$) and deviates strongly by $z = 0$. In particular, while the mean IGM temperature in the noFB case remains essentially constant between $0 < z < 2$ at $T_{\rm IGM} \simeq 10^{4.2}\,{\rm K}$, in the FB run it increases from that value at $z = 2$ to $T_{\rm IGM} \simeq 10^{5.3}\,{\rm K}$ by $z = 0$. The primary cause is our black hole feedback model operating in the radio-mode, which begins to alter the global IGM temperature statistics in the simulated volume as haloes of sufficiently high mass begin forming at $z \simeq 3$. In a larger simulation volume this would take place at higher redshift, and with a less homogeneous effect on the global box. We caution, however, that the gas content of massive systems is too low in our simulations (Genel et al., 2014), indicating that although the radio-mode model is efficient at moderating stellar mass growth in these haloes, its side effects – including the strong influence on IGM temperature – likely imply that the details of this model require modification.

In Fig. 3.4 we show how the fraction of gas with $T_{\rm max} < T_{\rm vir,acc}$ smoothly accreted onto haloes develops this difference between the FB and noFB cases as a function of halo mass, for four redshifts. The fraction of $T_{\rm max} < T_{\rm vir,acc}$ material accreted onto haloes agrees at $z = 2$ when comparing the FB and noFB runs. However, for sufficiently massive haloes by $z = 1$ this fraction is suppressed by a factor of two in the FB run, and drops to 10% or less by $z = 0$. It is notable that this same difference is *not* seen for accretion onto galaxies at low redshift. In this case, for the runs both with and without feedback, the fraction of gas with $T_{\rm max} < T_{\rm vir,acc}$ drops to zero at late time. This indicates that, despite

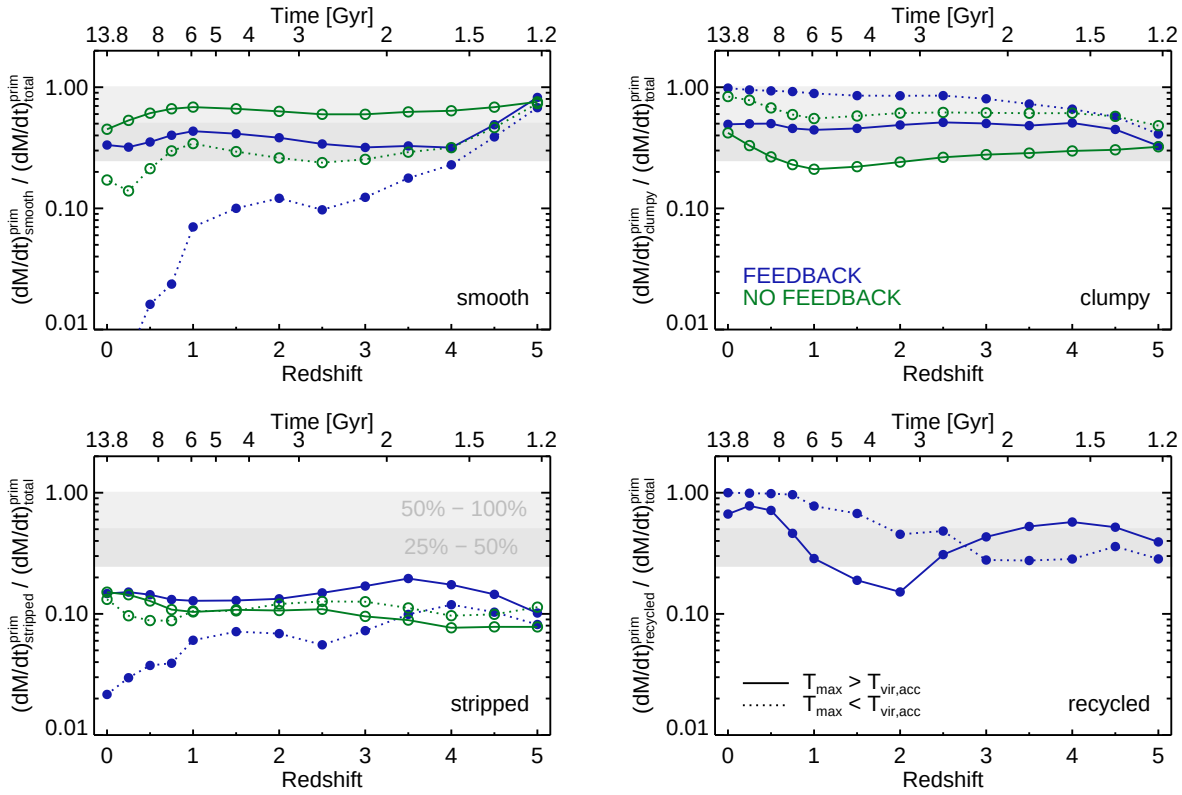

Figure 3.5 The fractional contribution of each accretion mode to the total net accretion rate of gas from cosmological origin onto central galaxies over a time window of $\Delta t = 250\,\mathrm{Myr}$, as a function of redshift. We separate the contribution based on the comparison of the maximum past temperature T_{max} of each Monte Carlo tracer to the virial temperature of future parent halo, at the time of accretion. The grey band indicates the 25% to 50% range. The four modes are smooth, clumpy, stripped, and recycled. The feedback run decreases the fractional importance of smooth accretion, regardless of temperature history and for all redshifts, particularly for cold material at late times.

the thermal history differences due to the IGM temperature, further heating proceeds within the halo before accretion onto the galaxy. The relative importance between a single strong virial shock, a series of smaller shocks within the halo, and adiabatic compression in this gas heating process remains an interesting question for future work.

3.3.3 Relative importance of different accretion modes

Having so far focused on smooth accretion, we now consider the importance of the different accretion modes. In Fig. 3.5 we show the fractional contribution of each mode to the total net accretion rate as a function of redshift. Each of the four modes is split into a separate panel. Each mode is further divided based on the comparison between $T_{\rm max}$ and $T_{\rm vir,acc}$. Most clearly, the contribution of smooth and clumpy accretion modes moves in opposite directions between the two runs. The inclusion of feedback physics suppresses the total relative contribution of smooth accretion by a factor of ~ 2. This is true at essentially all redshifts, and for all gas regardless of temperature history. However, the contribution of the smooth component to $T_{\rm max} < T_{\rm vir,acc}$ gas towards $z = 0$ is especially modified, being reduced to a negligible amount by the present day. We note that this conclusion holds without contradiction given our earlier finding from Figure 3.1 that the contribution of $T_{\rm max} < T_{\rm vir,acc}$ gas to the smooth component – that is, the "cold fraction" of smooth accretion – is largely unchanged.

The fractional importance of stripped material is similar between the two runs. By definition the hot and cold component lines both sum independently to one (excluding recycled, which is an additive attribute). Consequently, the contribution of resolved substructures increases to balance the decrease of the smooth mode. In terms of

primordial accretion contributing to the net growth of galaxies, for most of cosmic time, $z<3$, more than 80% (less than 40%) of material with a cold (hot) temperature history is acquired in the clumpy mode. This exact fraction will be sensitive to the definition of substructure – for instance, we include the entirety of a satellite halo, whereas Kereš et al. (2009) includes only the actual satellite galaxy/ISM material. Finally, tracers which have at some point in their past resided in the wind phase make up a significant fraction of the total net rate, particularly at late times – redshift zero growth is dominated by this recycled component. However, the fact that we are here only considering gas which has entered the MPB for the first time implies that this is largely recycling in satellite systems prior to incorporation into the central galaxy.

3.4 The State of Gas in the Halo

Here we consider the instantaneous properties of gas flows in the halo regime, regardless of its past or future history. To begin, Fig. 3.6 shows a prototypical, single halo in the mass range under consideration, matched between the two runs, at $z = 2$. An orthographic projection with extent in all dimensions equal to 3.5 times the virial radius (indicated with the largest white circle) shows gas density, mass-weighted gas temperature, and mass-weighted gas radial velocity, where negative denotes inflow. The run without feedback is shown on the top row, while the run with feedback is shown on the bottom row.

The temperature projection reveals that the energy injection from feedback arising in the central galaxy pushes the hot halo gas to slightly larger radii. The existence of cold and metal enriched wind material in our model suppresses the peak temperature in

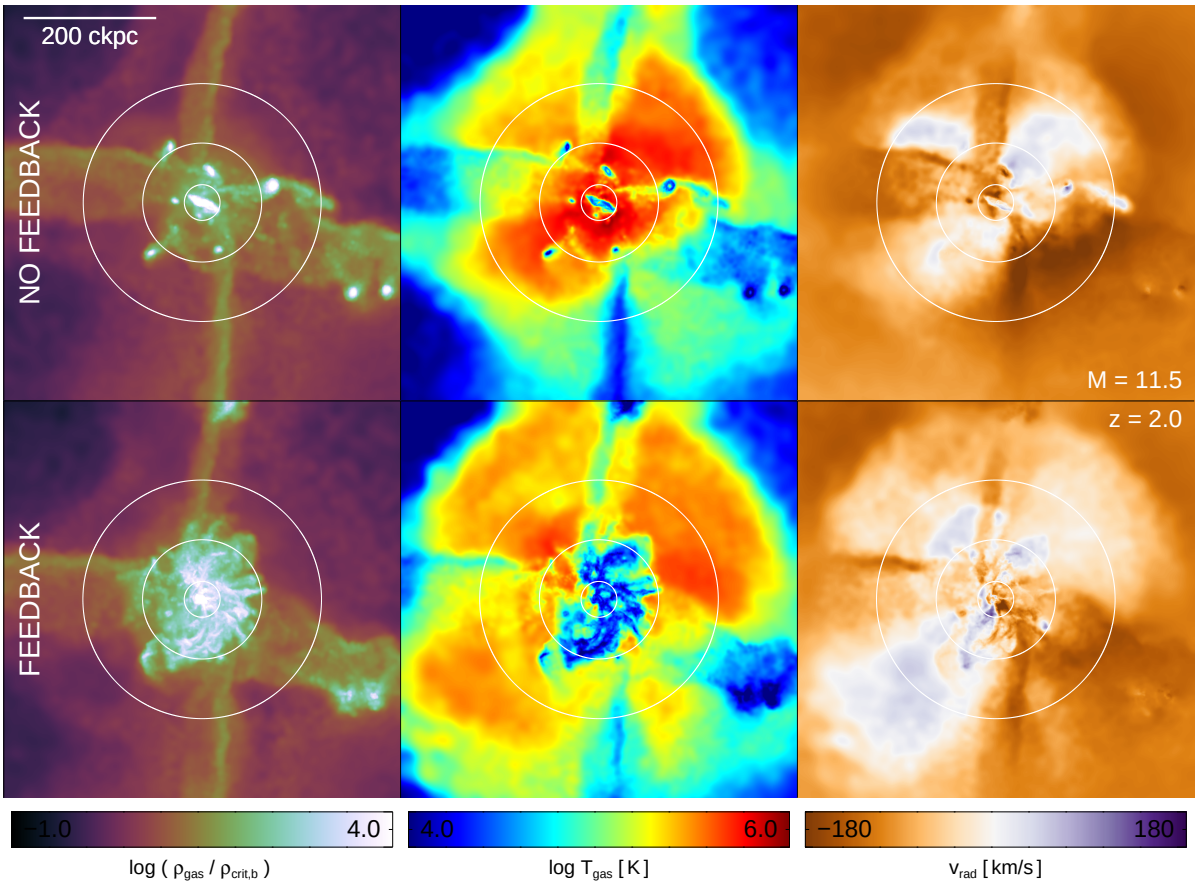

Figure 3.6 Comparison of a single halo with a mass of $10^{11.5}$ $M_\odot$ at $z = 2$ between the FB and noFB runs. Shown in projection are gas density (left), mass-weighted temperature (middle), and mass-weighted radial velocity (right). The three white circles are the same in all panels and denote $[1.0, 0.5, 0.15]$ $r_{\rm vir}$. Negative $v_{\rm rad}$ (brown) is infall, whereas positive (purple) $v_{\rm rad}$ is outflow. Introducing the feedback model pushes the hot halo gas to larger radii, while increasing the fraction of the virial sphere covered by outflow. Gas streams inflowing across the virial radius are largely unaffected, in terms of their temperature, density, or radial velocity. The virial radius of this halo is 180 kpc comoving, and the virial temperature is $\simeq 10^6$ K.

the inner halo, both directly and due to enhanced cooling. The stellar feedback driven winds clearly populate the inner halo ($r/r_{\rm vir} < 0.5$) with a large mass of cold gas with high covering fraction, substantially altering the temperature and velocity structure of gas at these small radii. Further from the galaxy, we see that more of the halo volume – and halo gas mass – is occupied by material with outward radial velocity. However, at the virial radius, gas inflow appears largely unaffected by the introduction of our fiducial feedback model. While in mass-projection the streams crossing $r_{\rm vir}$ appear somewhat more collimated, slower, and warmer, this is predominantly a side effect of this particular visualization. If we instead inspect spherical slices of gas properties at the virial radius, we find that there is e.g. no notable difference in the temperature distribution of inflowing material.

To be more quantitative, we construct spherical slices for all haloes at all analysis redshifts at a number of radii. We use a mass-weighted tophat kernel to interpolate gas quantities onto equal area pixels using the Healpix scheme (Górski et al., 2005). In Fig. 3.7 we calculate a spherical covering fraction of inflow or outflow as the fraction of pixels on this sphere with radial velocity above some threshold with the appropriate sign. As a threshold we take 10% of $v_{\rm circ}$(r) of an NFW halo of equal mass. Our results are insensitive to this choice provided it is small enough – we also considered a constant threshold of $\sim 20\,{\rm km\,s^{-1}}$, and a threshold of zero. The top panel shows the behaviour across the quarter virial sphere, while the bottom panel shows the same across the full virial sphere. In all cases we exclude all gas bound to resolved substructures prior to this calculation.

At $0.25\,r_{\rm vir}$, the impact of feedback is to substantially reduce the spherical covering fraction of inflow from 80% - 90% down to 50% - 60%, while outflow correspondingly

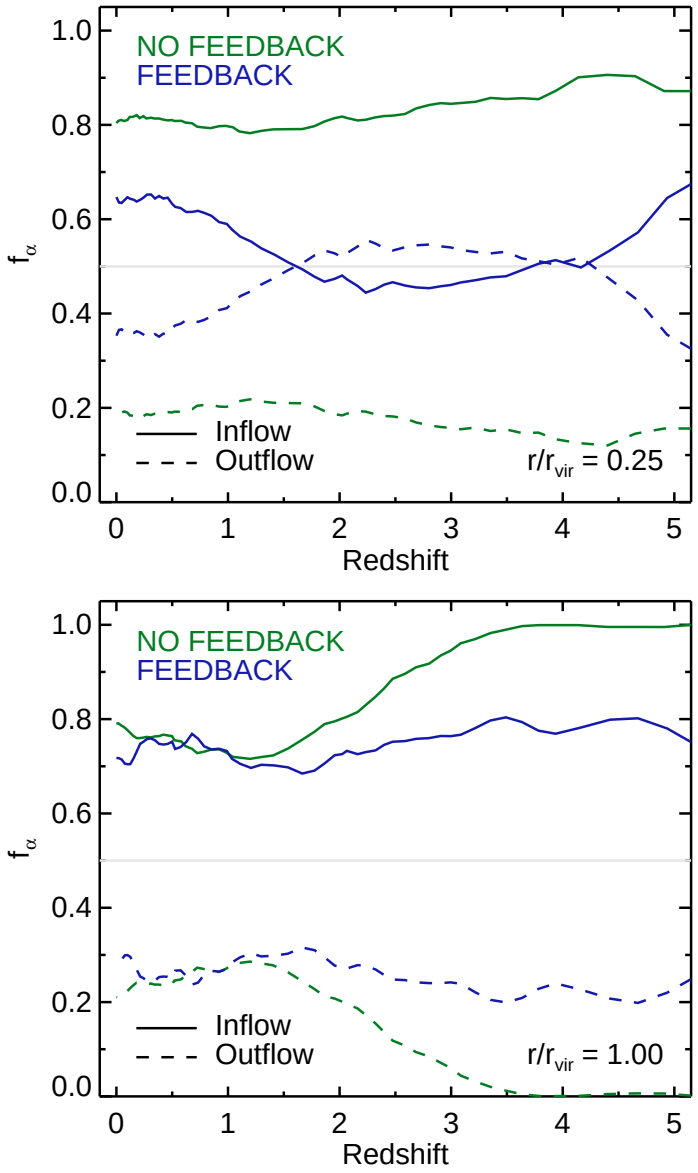

Figure 3.7 The spherical covering fraction of the quarter virial sphere (top panel) and full virial sphere (bottom panel) of inflow and outflow. Each is calculated as the fraction of the surface covered by gas of either positive or negative radial velocity which exceeds a relative threshold of 10% of the circular velocity of an NFW halo at that radius. Gravitationally bound substructures have first been removed. The halo mass range of $11.3 < \log(M_{\rm halo,tot}/{\rm M}_\odot) < 11.4$ is included. Near the galaxy, the presence of feedback increases the covering fraction of outflow above half, while at the virial radius this same behaviour is seen to a lesser degree and only at high redshift.

covers more than half of the surface at $z \sim 2-3$, up from $\simeq 20\%$. At the virial radius, both runs agree that $\geq 70\%$ of the virial sphere is covered by inflow, regardless of redshift. In the FB case, somewhat more material has outward radial velocity at high redshifts, but for this halo mass regime the difference is small and disappears towards $z = 0$. In agreement with the single halo shown previously in Fig. 3.6, we see that at $z = 2$ feedback introduces a negligible change to the spherical covering fractions of both inflow and outflow at the virial radius.

In Fig. 3.8 we calculate the radial mass flux rate of gas across the same two radii, for both inflow and outflow, with substructures removed. Since inflow and outflow occupy disjoint portions of each sphere, these rates are in principle independent, and do not necessarily correlate with the respective spherical covering fractions. At the virial radius (bottom panel) we see that the rate of gas inflow is unchanged, while feedback boosts outflow rates at high redshift substantially, although to magnitudes which are still small relative to inflow. This outflow across the virial sphere transfers mass, and therefore metals, from within virialised haloes out into the intergalactic medium. Notably, the presence of feedback does not lower the instantaneous inflow rates, which would indicate a direct impact on the accretion of intergalactic material onto the halo. Nor does it increase the inflow rates, which would indicate large-scale recycling motion across the host halo virial radius. At $0.25\,r_{\rm vir}$ (top panel) feedback has a larger impact, increasing the rates of both inflow and outflow, the magnitude of the difference growing with time. Given that the fraction of this surface covered by inflow is actually smaller in the FB run, this is a clear signature of significant gas recycling motion across this boundary.

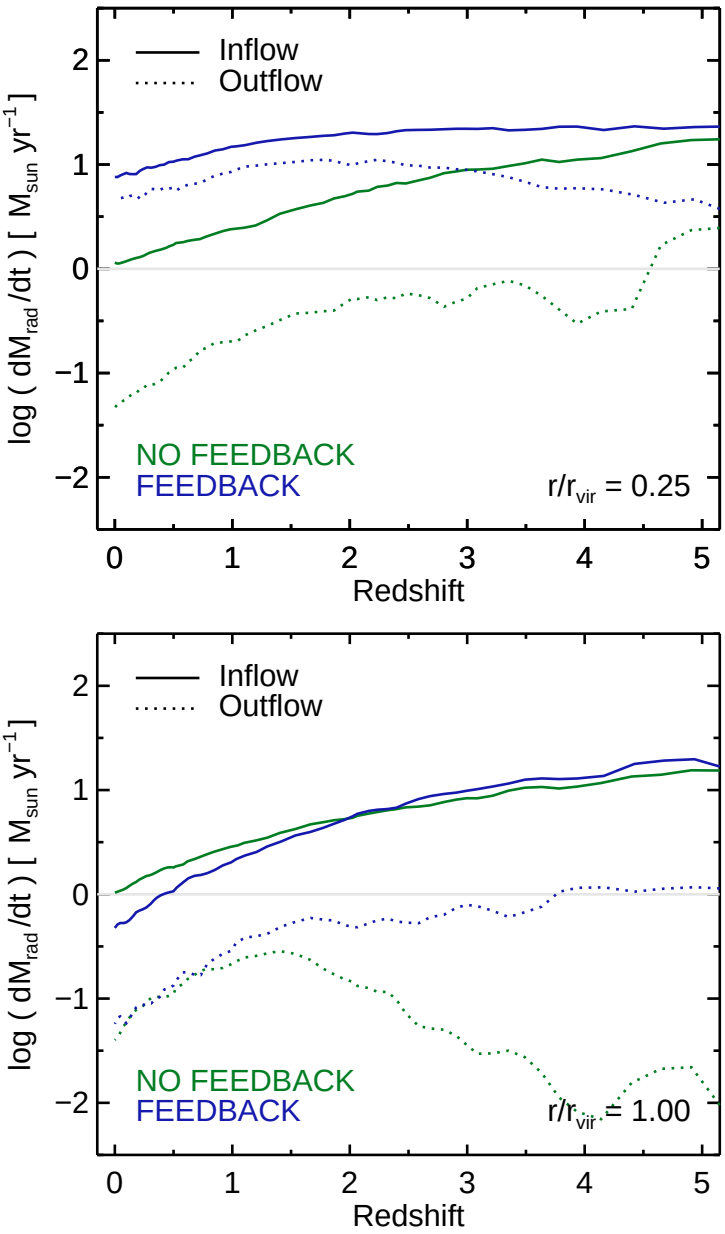

Figure 3.8 The radial mass flux rates across the quarter virial sphere (top panel) and full virial sphere (bottom panel), calculated separately for inflow and outflow. As in the previous figure, gravitationally bound substructures have first been removed, and the halo mass range of $11.3 < \log(M_{\rm halo,tot}/{\rm M}_\odot) < 11.4$ is considered. While the rate of inflow across the virial sphere is unchanged, the rates of both inflow and outflow closer to the central galaxy are both significantly increased, implying recycling occurring across the $0.25\,r_{\rm vir}$ boundary.

3.5 Timescale of Gas Accretion

In the previous section we saw that the morphology of inflow at the virial radius is largely unchanged. Here, we consider whether the presence of outflows increases the time required for gas to inflow (or "transit") from the virial radius to the galaxy. We measure this quantity as the time difference between the first virial radius crossing $t_{\rm halo}$ and the first incorporation into the galaxy $t_{\rm gal}$. In Fig. 3.9 we show the distribution of this time difference for all tracers within haloes at $z = 2$ which have previously recorded these two crossing times, split into halo mass bins from 10^9 $M_\odot$ to 10^{12} $M_\odot$. We include only smooth accretion.

First, in both runs and for all halo masses we find a broad, uni-modal distribution. That is, there is no obvious evidence for multiple channels of accretion having different halo crossing timescales. Comparing the FB (solid) and noFB (dotted) runs, we find that feedback introduces a significant delay in the halo transit time of smooth mode material. Each distribution shifts to longer times by a factor of ~ 2 - 3, and this factor is largely independent of halo mass, at least above 10^{10} $M_\odot$, where systems are well resolved. Specifically, gas smoothly accreted by $z = 2$ takes on average $\simeq 250$ Myr to cross from the virial radius to the galaxy in the noFB run, and $\simeq 700$ Myr in the FB run.

Of even more interest, we see that in both simulations, this halo transit time is again largely independent of halo mass. This implies that this time difference may be related more to the dynamical time and *not* to the cooling time, since the latter scales strongly with halo mass. Over the halo mass range of 10^{10} $M_\odot$ to 10^{12} $M_\odot$, then, we see no evidence for a transition point above and below which the process of gas accretion

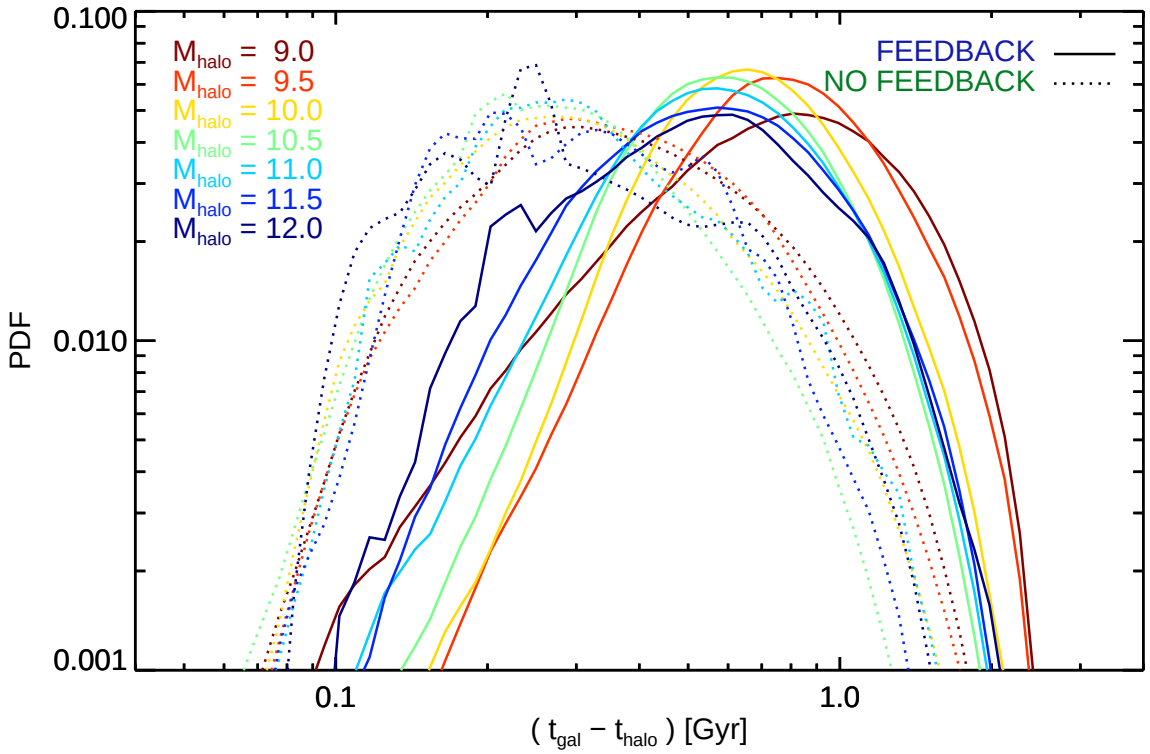

Figure 3.9 Distribution of time difference between the first $1.0\,r_{\rm vir}$ and $0.15\,r_{\rm vir}$ crossings (the "halo transit time" for each tracer during its first accretion into the halo) for smooth accretion only. Includes all tracers within haloes which have recorded these two crossings by $z = 2$. Separated by halo mass (colored lines) and for the no feedback (dotted) and feedback (solid) runs. The FB run shifts each distribution, regardless of halo mass, to longer times by a constant factor of ~ 2-3. In addition, for both simulations, the distribution of transit times is nearly independent of host halo mass.

through the halo occurs in a fundamentally different way, at least insofar as is captured by our measured "halo transit time". We return to this point and its implication for the idea of a critical halo transition mass in the discussion.

To further explore any possible relationship between this transit time and different accretion mechanisms, Fig. 3.10 shows the relation with the $T_{\rm max}$ of each tracer. The question is whether or not gas with low $T_{\rm max}$ has a shorter halo transit time, while gas with higher $T_{\rm max}$ spends longer in the halo, as might be naively expected. We see a weak correlation in this direction for the noFB simulation (bottom panel), which is less clear after including feedback (top panel). This lack of correlation implies that the mean halo transit times for both hot and cold gas are comparable. That is, there is no strong signature of the past thermal state of gas in its dynamical history, at least insofar as is measured by the halo transit time. In both simulations, notably, there is clearly a lower envelope. This is indicated in the top panel by the dashed black line, which shows the relation

$$\log\left(\frac{T_{\rm max}}{T_{\rm vir,acc}}\right) = \frac{(t_{\rm gal} - t_{\rm halo})}{2\,{\rm Gyr}} - 0.8. \tag{3.1}$$

We see that gas does not populate the region to the lower-right of this relation – that is, there is a minimum $T_{\rm max}$ reached which increases as a function of increasing halo transit time. Note that we have here stacked together haloes of all masses, but we see this relation also when we examine small bins in halo mass. This implies that there is at least some link between virial heating and the gas dynamics of accretion. We therefore want to understand where the maximum temperature of each tracer is reached – is the $T_{\rm max}$ event closely related to the virial crossing time, or does it occur on average with either a positive or negative relative lag.

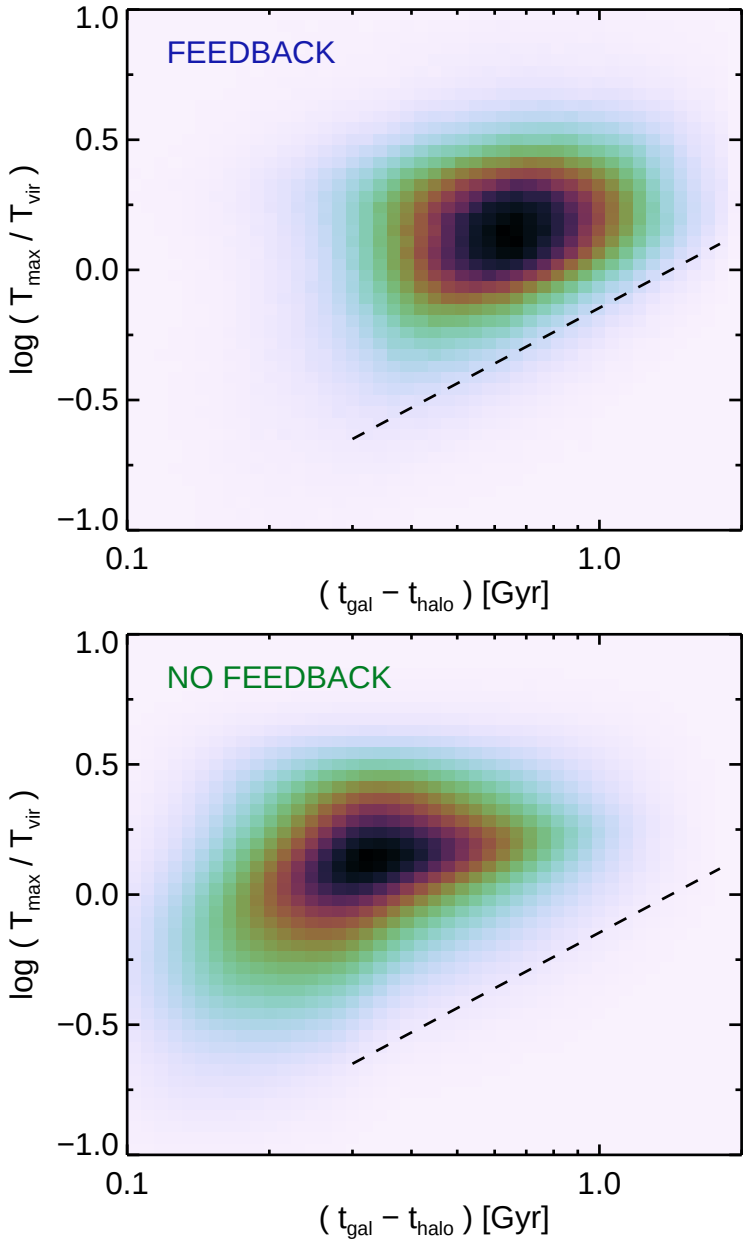

Figure 3.10 Correlation between $T_{\rm max}$ and "halo transit time" ($t_{\rm gal}$-$t_{\rm halo}$), where the colour scale indicates the mass distribution of tracers smoothly accreted by $z = 2$ in this plane. While the noFB run (bottom panel) may indicate a relation between maximum past temperature and halo transit time, this is less clear in the FB run (top panel). In both cases there is a lower envelope, approximately indicated by the dashed line in the top panel, indicative of a minimum $T_{\rm max}$ which increases with increasing transit time.

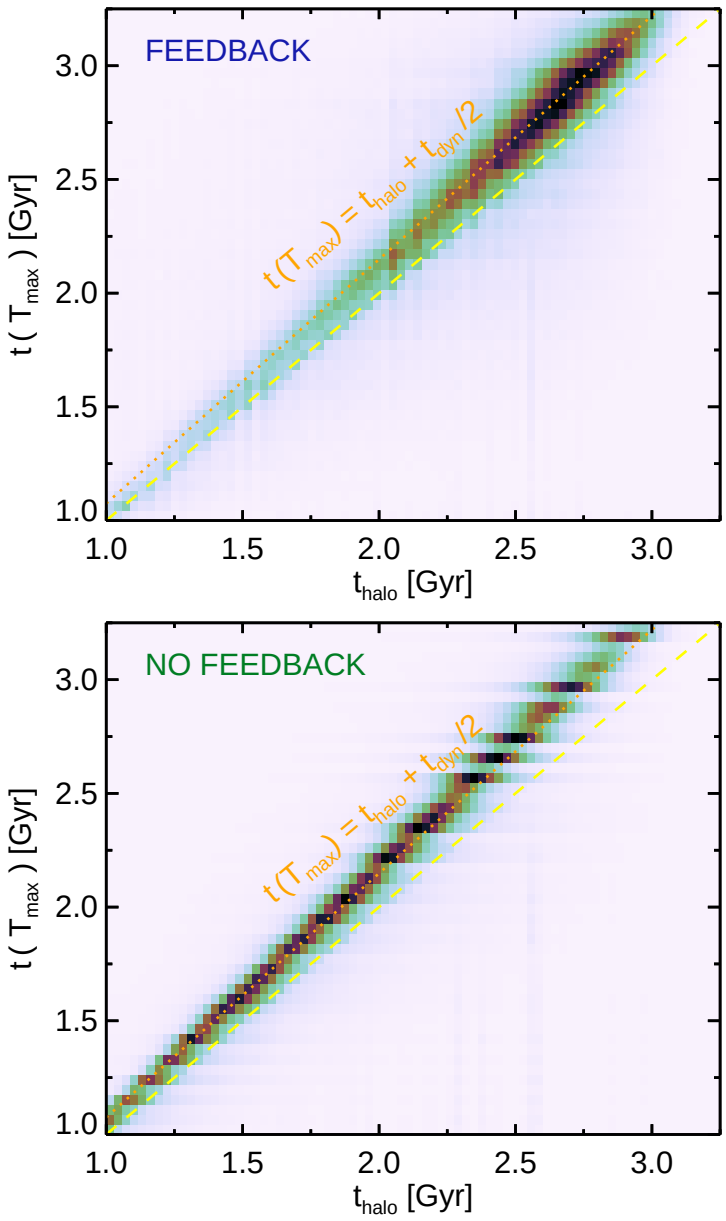

Figure 3.11 The tracer mass distribution in the plane of $t(T_{\rm max})$, the time when the maximum past temperature was obtained, and t_2, the time of first virial crossing. Both are given in terms of the age of the universe, where the yellow dashed line shows the 1-to-1 relation. The orange dotted line is offset by a constant factor, such that tracers on this relation would have reached their maximum temperature a time $t_{\rm dyn}/2$ after crossing the virial radius. Smooth accretion only, for tracers which have entered haloes by $z=2$.

In Fig. 3.11 we show the correlation between the time of $T_{\rm max}$ and the time of the first virial radius crossing, both in terms of the age of the universe, on a tracer by tracer basis. Only smooth accretion is included. The yellow dashed line shows the 1-to-1 line, which would imply that gas heats at $r_{\rm vir}$ to its maximum temperature, subsequently cooling in order to join the ISM of the galaxy. The dotted orange line shows a constant positive time offset of $t_{\rm dyn}/2$ later, where we take the redshift dependent

$$t_{\rm dyn} = r_{\rm vir}/v_{\rm circ}(r_{\rm vir}). \tag{3.2}$$

In both runs, these two lines bound the majority of accreted material, implying that gas reaches its maximum temperature shortly *after* first accretion into the halo, and never before. It is clear that the ratio $t(T_{\rm max})\,/\,t_{\rm halo}$ increases with time, but since $t_{\rm dyn}$ also evolves with redshift, becoming longer at later times, we can see that the characteristic lag time is $\simeq 0.5\,t_{\rm dyn}$, at least for all accretion which has occurred by $z=2$. Comparing the two panels, we conclude that the relation between heating and virial crossing for smoothly accreted gas is not strongly affected by the presence of our fiducial feedback model. This result holds also for accretion taking place by $z=1$ and $z=0$, although the mean lag time increases to $\simeq 1.0\,t_{\rm dym}$ for gas entering the halo at late times ($z<1$).

3.6 Discussion

3.6.1 The contribution of recycled gas

We have so far focused exclusively on the accretion of material of primordial, or cosmological, origin. We can contrast this evolution of $\dot{M}_{\rm gas}^{\rm prim}$ from Section 3.3 with the

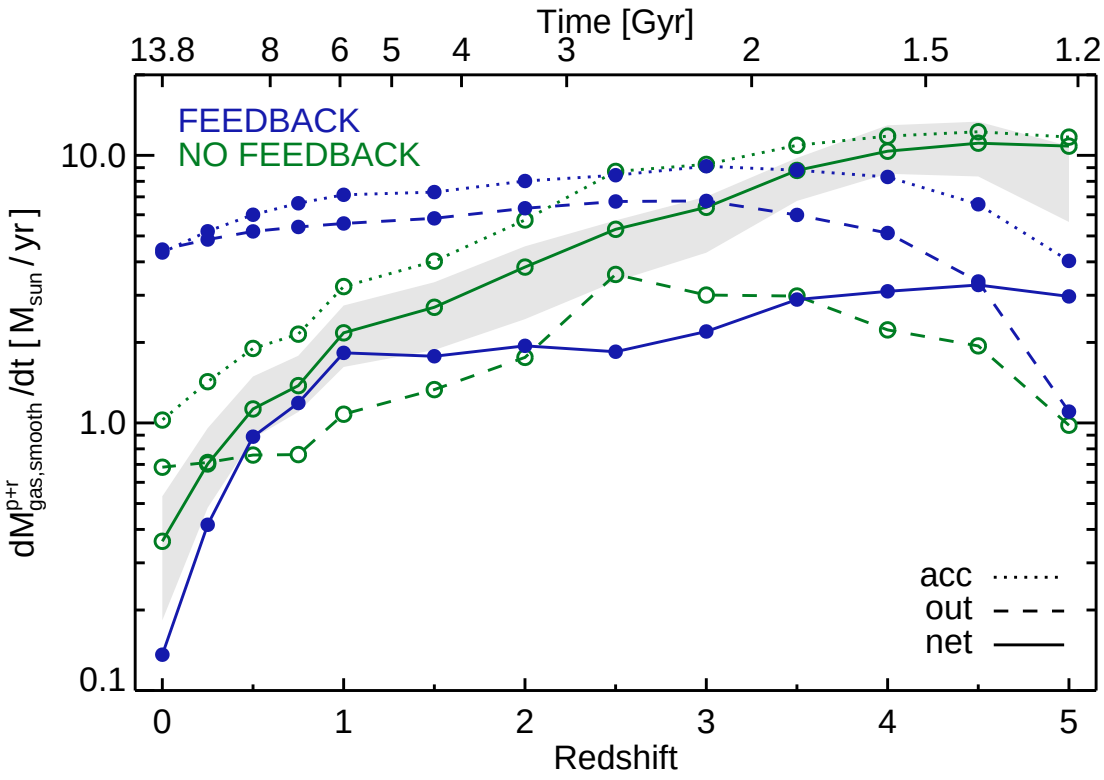

Figure 3.12 As in Fig. 3.2 but showing the combined smooth gas accretion rate, including both primordial and material previously cycled through the main progenitor branch. We include accretion onto central galaxies as a function of redshift. As before, we include only haloes in the mass range $11.3 < \log(M_{\rm halo,tot}/{\rm M}_\odot) < 11.4$, and consider separately the contribution from net inflow, outflow, and raw accretion.

fundamentally different net accretion rates obtained if we count tracers with *most recent* incorporation times, as opposed to *first* incorporation times, within the past 250 Myr. We denote this measurement $\dot{M}_{\rm gas}^{\rm p+r}$, where the combined contributions from primordial and material previously cycled through the MPB are two disjoint subsets of accretion which sum to the total instantaneous accretion rate. As a result, this measurement includes all gas motion across the boundary defining the galaxy, regardless of when that gas initially entered the galaxy. Note that by definition, tracers within the galaxy with outward crossings are a strict subset of tracers within the galaxy with inward crossings. They represent the mass component which cycles out of, and back into, the galaxy during the 250 Myr time interval.

Fig. 3.12 revisits the smooth accretion rate as a function of redshift, where the relationship between the FB and noFB runs is significantly altered when compared to Fig. 3.2. As expected, inflow and outflow rates are higher for both runs at low redshift. The scaling of inflow rates with redshift in the noFB run remains similar, with the net accretion rate declining monotonically from $z \simeq 3$ to $z = 0$. However, the inclusion of feedback maintains a nearly constant net accretion rate until $z = 1$. The smooth "p+r" accretion rate at high redshifts, $z > 4$, is similar to the primordial accretion only value, implying that the contribution of recycled gas is negligible at early times. However, recycling becomes more important towards low redshift, at a level sufficient to balance the decreasing inflow of primordial gas. Between $z = 1$ and $z = 0$ the raw inflow rates decrease by factor of ~ 2 in both simulations, pulling the net acquisition of smooth mode material down by a similar factor.

We can also reconsider the ratio of accretion with $T_{\rm max} < T_{\rm vir,acc}$ to the total rate, as in Fig. 3.3, but including the recycled component (not shown). In the feedback run, the

fraction of material with $T_{\rm max} < T_{\rm vir,acc}$ is larger at all redshifts. In particular, at $z = 0$, the ratio to the total is $\simeq 0.04$, 0.15, 0.25 and 0.4 for smooth, stripped, recycled and clumpy modes, respectively. The greater fraction of accretion with a cold temperature history reflects a simple time preference – that recycled gas on average enters the halo earlier than non-recycled gas, combined with a cold fraction which increases towards higher redshift.

The other interesting difference we find when including the recycled component in our analysis appears in the fractional contribution of each accretion mode to the total net rate, as in Fig. 3.5. In this case, the conclusions for hot material with $T_{\rm max} > T_{\rm vir,acc}$ are all qualitatively unchanged, whereby feedback still suppresses the relative importance of smoothly accreted gas by a similar amount. However, for colder material ($T_{\rm max} < T_{\rm vir,acc}$) the fraction contributed by each of the smooth, clumpy and stripped modes converges to similar values for the FB and noFB runs. The reason remains somewhat unclear, and may be coincidental. As expected, the recycled mode also becomes more significant, since we now include recycling within the main progenitor branch. In this case, the contribution is monotonically increasing towards low redshift, increasing from $z = 5$ to $z = 0$ from 0.3 to 1.0 (cold) and from 0.6 to 1.0 (hot). That is, by redshift zero, essentially all gas being accreted by galaxies in this mass regime has been previously recycled at least once.

An in-depth consideration of the recycled mode is beyond the scope of this paper, and remains a topic for future work. Of particular interest is the recycling timescale – that is, the amount of time between ejection in a wind and re-incorporation into a galaxy. This quantity is particularly important in semi-analytic models, for instance to fix the overly early build-up of low mass galaxies (Henriques et al., 2014). Further, our

analysis methodology, which revolves around recording only one radial crossing time per direction per radius per tracer, implicitly assumes that gas does not cycle through the same radius more than once over the time interval (250 Myr). Oppenheimer et al. (2010) found, for instance, a recycling time of $\sim 1\,\mathrm{Gyr}$ at $z = 1$ for haloes of this mass. Also of interest is the prevalence of multiple recyclings, or, the number of times a baryon mass element has belonged to a wind in its history. Similarly, the contribution of satellite recycling and satellite winds in general as a contributor to both the intergalactic medium and the circumgalactic medium of larger host haloes.

However, the process of recycling involves the complex wind physics associated with feedback as well as the difficult numerical problem of multiphase gas interaction (Agertz et al., 2013; Sarkar et al., 2014). More faithful modelling of stellar feedback processes and/or higher resolution than is currently available in full cosmological volumes is likely required before the process of recycling can be better understood (Hopkins et al., 2014; Aumer et al., 2013; Stinson et al., 2013; Faucher-Giguere et al., 2014; Agertz & Kravtsov, 2014; Ceverino et al., 2014). In more massive haloes, treatments of AGN feedback will also need to be improved (Wurster & Thacker, 2013; Costa et al., 2014; Bachmann et al., 2014; Li & Bryan, 2014), while the contribution from other outflow mechanisms such as cosmic rays (Booth et al., 2013; Hanasz et al., 2013; Salem & Bryan, 2014) is potentially also critical. Clearly, the baryon cycle in its entirety cannot yet be satisfactorily understood given the current level of sophistication of feedback modelling in cosmological simulations.

3.6.2 Implications for a critical halo transition mass

In all semi-analytic models of galaxy formation (e.g. White & Frenk, 1991; Croton et al., 2006; Bower et al., 2006, and more recent efforts), a foundational theoretical consideration is the cooling of hot halo gas. This essentially governs the growth of the cold mass reservoir of the galaxy, and so the subsequent star formation and stellar mass growth. A common approach is to let the hot halo gas with $t_{\rm cool} < t_{\rm dyn}$ or $t_{\rm cool} < t_{\rm H}$ cool onto the central galaxy.

Furthermore, they generally implement a critical halo transition mass, below which hot halo cooling proceeds in a "rapid" mode, and above which cooling proceeds in a "slow" mode, where the gas fraction eligible to cool drops to zero. This difference in the behaviour of quasi-static hot halo gas between more and less massive dark matter haloes is motivated largely from the classic, analytical arguments of Silk (1977); Rees & Ostriker (1977); White & Rees (1978). Analytic arguments supported by one-dimensional, spherically symmetric calculations, cast in terms of whether or not a stable virial shock can exist (Birnboim & Dekel, 2003a; Dekel & Birnboim, 2006) have also identified a transition point in halo mass. They find that haloes below $\sim 10^{11}$ $\rm M_{\odot}$, independent of redshift, cannot provide enough post-shock pressure to support a virial shock against losses from radiative cooling.

In Section 3.5 we found that for both simulations, with and without feedback, the time taken by baryons to transit from the virial radius to the central galaxy is independent of halo mass (see also Brook et al., 2014). This holds true, at the very least, for haloes with 10^{10} $\rm M_{\odot}$ $< \rm M_{halo} <$ 10^{12} $\rm M_{\odot}$. This indicates that, in contrast to analytic expectations, we see no strong evidence in the simulations for a sharp halo transition

mass, above and below which the process of gas accretion proceeds in a fundamentally different manner. It will be important to validate this result in the future by extending the dynamic range of this analysis to more massive haloes. In particular, good sampling is available up to 10^{14} $M_\odot$ in the $\simeq (100 \text{ Mpc})^3$ volume of Illustris.

The agreement between theoretical cooling models and hydrodynamic simulations has been an occasional subject of interest over the past decade (Yoshida et al., 2002; Viola et al., 2008; Lu et al., 2011; Monaco et al., 2014), where the general conclusion has been that the cooling models implemented in SAMs provide an acceptable match to simulations. However, these studies have all used the smoothed particle hydrodynamics (SPH) technique, and (as Monaco et al. (2014) points out) in Nelson et al. (2013) we found that numerical inaccuracies in "classic" SPH strongly affect the processes of gas accretion in cosmological simulations, particularly failing to accurately capture the cooling properties of hot haloes. The very fact, then, that older SPH simulations agree with classic hot halo cooling theory points to a looming disagreement between that theory and modern, cosmological simulations. Understanding in what ways these simple theoretical arguments agree with numerical calculations (see also Dekel et al., 2013), and how the picture of hot halo cooling in semi-analytical models can potentially be improved upon, remains an open, important question for future work.

3.6.3 Comparison to previous studies

As pointed out in the introduction, reconciling previous conclusions is complicated by a variety of different feedback implementations, numerical methods, halo mass regimes, statistical samplings, and interpretations. We compare our findings with those works

which specifically focused on the impact of feedback on cosmological gas accretion.

In particular, Oppenheimer et al. (2010) considered a few variations of kinetic wind schemes similar to the one presented herein, with constant velocities or a momentum-driven scaling, and decomposed star formation rates into three distinct modes: hot, cold and wind (recycled). Contrasting recycled versus non-recycled, they concluded that accretion is dominated by recycled gas below $z<1$ for $\mathrm{M_{halo}} > 10^{11}$ $\mathrm{M_{\odot}}$, neglecting AGN effects. This is qualitatively consistent with our findings and Figure 3.5 for primordial accretion. They further claimed that outflows efficiently suppress both cold and hot modes, where these are defined with the standard T_{max} criterion, excluding all gas ever belonging to a wind. Our definition of primordial is different, in that it allows for recycling in satellites prior to incorporation into the MBP. Given this caveat in the comparison, our findings in Fig 3.1 are also qualitatively consistent with this conclusion.

Faucher-Giguère et al. (2011) also considered kinetic wind models, with various constant velocities and mass loading factors, and measured accretion rates separated into cold and hot modes based on instantaneous mass fluxes through radial shells together with instantaneous gas temperatures. In this respect the analysis differs significantly from Lagrangian definitions of accretion and thermal history, and so is not quantitatively comparable. It furthermore prohibits the identification of recycled material as in Oppenheimer et al. (2010) and this work. They focused only on flux through the virial shell (accretion onto the halo as opposed to the galaxy), and showed galaxy fluxes only for simulations with no winds. Nonetheless, neglecting AGN effects, they found that the net cold gas accretion rate onto low-mass haloes was suppressed, while the effect in high-mass haloes was negligible. Specifically, for the halo mass range considered here, the decrease is between zero and a factor of a few, relatively independent of redshift,

which is tension with the changing balance of hot versus cold accretion onto halos found in our Figure 3.4, and likely a result of the different wind velocity parametrisation.

van de Voort et al. (2011b) considered many simulations with different feedback physics. They found that while the gas accretion rates onto haloes was relatively robust against the presence of feedback, the rates onto galaxies themselves depended sensitively on stellar winds as well as metal-line cooling. In particular, at our halo mass scale at $z = 2$, there was negligible change in the total smooth accretion rate onto galaxies between their reference and no feedback models, although this is a sensitive function of mass. However, they also found an order of magnitude decrease in total smooth rates introducing either a density dependent wind scaling or thermal AGN feedback. The balance of hot and cold smooth accretion onto galaxies at both $z = 2$ and $z = 0$ was found to be fairly insensitive to feedback, in qualitative agreement with our findings. Their AGN model did not alter the cold fraction of smooth accretion onto haloes at $z = 2$, in disagreement with our Figure 3.4. As an extension, van de Voort et al. (2011a) focused on the impact of AGN feedback on inflow and found that it preferentially prevented hot mode gas, with high maximum past temperature, from cooling from the halo onto the galaxy. That is, AGN feedback reduces hot halo accretion more than cold accretion, based on analysis of all resolved haloes, although this is presumably an important effect only in massive haloes above $\simeq 10^{12}$ $M_{\odot}$. We do not find this differential effect at fixed halo mass, although in this work we have focused on lower mass systems not dominated by AGN effects.

Woods et al. (2014) includes a combined delayed cooling supernova and early stellar feedback model, a numerically distinct mechanism for generation of galactic-scale winds from those just discussed. They found strong recycling, as we do. Contrasting weaker

to stronger stellar feedback, they concluded that the total gas accretion rates did not change with strong feedback, while the balance between cold and hot components did. In particular, that the cold component became more important. Comparing to our results for smooth accretion, this result is in qualitative disagreement. However, this work did not differentiate the merger contribution, and concludes that more cold gas is available for accretion in strong feedback runs because more cold gas is present in satellite haloes, not due to any change in accretion processes. This could be interpreted as tentative agreement with our conclusion that the fractional contribution of clumpy accretion to the net cold rate is larger in the run with feedback (upper right panel of our Figure 3.5).

Finally, Übler et al. (2014) implements a hybrid thermal/kinetic stellar feedback scheme, considering five haloes slightly more massive than those discussed herein. They found that strong outflows generate substantially higher raw accretion rates, and that recycled material dominates galactic gas accretion at $z<1$, in good qualitative agreement with our results. They also found that the strong feedback case actually increased the rate of "first" (our "primordial") gas accretion, without differentiating between thermal history or different modes. The cause is likely then related to the merger contribution, as above, with similar conclusions.

However, all six studies discussed this far have been conducted with the "classical" density formulation of smoothed particle hydrodynamics (SPH, see Springel et al., 2005a). They therefore agree in their "no feedback" cases with Kereš et al. (2009), finding that the cold mode dominates the accretion rates, and/or the star formation rates, of galaxies of all masses, particularly so at the higher redshifts of $z=3$ or $z=2$. However, we have previously shown that this result is incorrect and a consequence of numerical issues with the hydrodynamical method of SPH (Nelson et al., 2013). Moreover, as typically

employed in galaxy formation simulations, SPH is not strictly numerically convergent (Zhu et al., 2014). As a result, it is somewhat unclear how to interpret any subsequent conclusions based on the inclusion of further baryonic feedback processes.

Dubois et al. (2013) investigated AGN feedback at high redshift ($z = 6$) in a single massive halo (10^{15} $M_\odot$ at $z = 0$), and found that large-scale hot superwinds could morphologically disturb cold filaments and quench cold diffuse accretion. This is a substantially different mass scale than the one probed here and we cannot make any direct comparison.

It has been pointed out that details of the wind interaction, travel extent, IGM enrichment, as well as its recycling properties, the balance between inflow and outflow fluxes, and the timescale of reincorporation may all be sensitive to physical model inputs as well as numerical details (Oppenheimer et al., 2010). Furthermore, current models designed to incorporate feedback and other physical effects from AGN are diverse and relatively unsophisticated, at least as applied in cosmological volumes. Finally, different analysis methodologies can make straightforward comparison between simulations impossible. While the studies discussed above seem to agree on several qualitative conclusions, it would seem premature to claim any strong agreement between simulations as to the impact of feedback on cosmological gas accretion. To do so would require identical analysis techniques applied to either large volumes with statistically robust galaxy populations, or identical initial conditions of individual haloes, with equivalent physical models and numerical implementations for non-feedback physics, including stellar evolution, metals, cooling, star formation, ISM pressurisation, hydrodynamics, and gravity. All of the studies reviewed here differ in one or more of these aspects, motivating a more controlled investigation in order to resolve this question.

3.7 Conclusions

In this paper we have compared two simulations, realisations of the same initial conditions evolved with the moving mesh code AREPO, using the same methods for both gravity and hydrodynamics. One, which we label as "no feedback", includes only the simplest baryonic physics – radiative cooling of a primordial gas, and star formation. The other, which we label as "feedback", implements the full physics model of the Illustris simulation project, with all models and parameters unchanged. Most importantly, the latter includes metals and metal-line cooling from enriched gas, stellar feedback resulting in galactic-scale winds, as well as feedback from supermassive black holes. This model has been shown to successfully reproduce a number of key stellar observables, in particular the $z = 0$ stellar to halo mass relation, and the star formation rate density as a function of redshift, over the full halo mass range we consider here. Our aim was to understand how such a comprehensive set of feedback processes, implemented in a full cosmological volume, affects the mechanisms by which galaxies acquire their baryonic material. Focusing on the accretion of gas by galaxies over cosmic time, we have arrived at four principal conclusions:

1. We first consider the accretion rate of material contributing to the net growth of galaxies. We find that the presence of feedback strongly suppresses "smooth mode" gas – originating directly from the intergalactic medium, without prior incorporation into a satellite galaxy – at all redshifts. For accretion onto galaxies in $\sim 10^{11.5}$ $M_\odot$ haloes, the rate of smooth accretion is reduced by a factor of ~ 10 by $z = 1$, increasingly so towards $z = 0$. Furthermore, we find that this suppression is

independent of the temperature history of newly acquired gas, implying that star formation driven galactic winds have marginal impact on the thermal evolution of smoothly accreting material, and that the presence of winds does not preferentially prevent material as a function of its past virial heating, or lack thereof. In addition to suppressing the net rates, feedback also reduces the raw inflow rates of smooth accretion by a factor of $\sim$2, regardless of redshift.

2. We examine the spatial distribution, temperature, and dynamics of gas in haloes sufficiently massive to exhibit both inflowing streams of gas at the virial radius and strong feedback driven outflows arising from the central galaxy. Feedback populates the inner halo ($r/r_{\rm vir} < 0.5$) with a large mass of cold gas with high covering fraction, substantially altering the temperature and velocity structure of halo gas at these small radii. At the virial radius, however, gas inflow is largely unaffected by the introduction of our fiducial feedback model – for example, we find no notable difference in the temperature distribution of inflowing material, nor the inwards radial mass flux, across $r_{\rm vir}$. The spherical covering fraction of inflowing gas at $0.25\,r_{\rm vir}$ decreases substantially, at $z = 2$ from more than 80% to less than 50%, while the rates of both inflow and outflow increase, indicative of recycling across this boundary.

3. Comparing the relative contribution of different accretion modes – smooth, clumpy (merger), stripped and recycled – we find that the fraction of the total net accretion contributed by smooth accretion is lower in the simulation with feedback, by roughly a factor of two across all redshifts, and particularly so for $T_{\rm max}\,/\,T_{\rm vir,acc} < 1$ material at $z < 1$, which is suppressed in the feedback run to a negligible level. Gas which has been recycled through a wind phase prior to its incorporation into a

central galaxy makes up a large fraction of total accretion. For gas entering the galaxy for the first time, this fraction is >50% at $z<1$, and between 10% - 50% at $z>2$.

4. As a measure of the timescale of accretion, we calculate the time difference between the first (highest redshift) virial radius crossing and the first incorporation into the galaxy for accreting gas. For smooth accretion, the distribution of these "halo transit times" is uni-modal and broad. The mean transit time increases in the presence of feedback by a factor of $\simeq 2$ - 3, but independent of halo mass. Furthermore, the distribution of these times is *also* independent of halo mass, at least from 10^{10} $M_{\odot}$ to 10^{12} $M_{\odot}$. This holds true in both runs, with and without feedback, indicating that the timescale of accretion through the halo does not exhibit any sharp transition point in halo mass, above and below which the process of gas accretion proceeds in a fundamentally different manner. The full implications for theory concerning the cooling of hot halo gas, such as those commonly used in semi-analytical models of galaxy formation, remain an important and intriguing direction for future work.

Chapter 4

Zooming in on accretion - I. The structure of halo gas

D. Nelson, S. Genel, A. Pillepich, M. Vogelsberger, V. Springel, L. Hernquist, submitted to *Monthly Notices of the Royal Astronomical Society*, 2015

Abstract

We study the properties of gas in and around 10^{12} $M_\odot$ haloes at $z = 2$ using a suite of high-resolution cosmological hydrodynamic 'zoom' simulations. We quantify the thermal and dynamical structure of these gaseous reservoirs in terms of their mean radial distributions and angular variability along different sightlines. With each halo simulated at three levels of increasing resolution, the highest reaching a baryon mass resolution of

∼10,000 solar masses, we study the interaction of filamentary inflow and the quasi-static hot halo atmosphere. We highlight the discrepancy between the spatial resolution available in the halo gas as opposed to within the galaxy itself, and find that stream morphologies become increasingly complex at higher resolution, with large coherent flows revealing density and temperature structure at progressively smaller scales. Moreover, multiple gas components co-exist at the same radius within the halo, making radially averaged analyses misleading. This is particularly true where the hot, quasi-static, high entropy halo atmosphere interacts with cold, rapidly inflowing, low entropy accretion. We investigate the process of gas virialization and identify different regimes for the heating of gas as it accretes from the intergalactic medium. Haloes at this mass have a well-defined virial shock, associated with a sharp jump in temperature and entropy at $\gtrsim 1.25\, r_{\rm vir}$. The presence, radius, and radial width of this boundary feature, however, vary not only from halo to halo, but also as a function of angular direction, covering roughly $\sim 85\%$ of the 4π sphere. Our findings are relevant for the proper interpretation of observations pertaining to the circumgalactic medium, including evidence for large amounts of cold gas surrounding massive haloes at intermediate redshifts.

4.1 Introduction

Initially following the gravitational collapse of a dark matter overdensity, gaseous haloes subsequently grow through the accretion of baryons from the intergalactic medium (IGM). Their evolving structure across cosmic time has been the subject of theoretical as well as observational interest for several decades. As the transitional state between the diffuse IGM and the star-forming interstellar medium of galaxies, these gas reservoirs

regulate the stellar growth of forming galaxies. Understanding not only the structure of halo gas, but also its origin and subsequent evolution, is therefore essential for any comprehensive theory of galaxy formation.

Although the accretion of gas will follow that of dark matter, the presence of additional physical processes including hydrodynamical forces and radiative cooling imply additional complexity for the acquisition of baryons. In the classic picture, gas accreting from the intergalactic medium will shock heat to the virial temperature of the halo. If the radiative cooling timescale is sufficiently long, it will then form a hot, pressure supported atmosphere in approximate equilibrium (Rees & Ostriker, 1977; Silk, 1977; White & Rees, 1978). Cooling can proceed, delivering gas into the halo centre (White & Frenk, 1991). For sufficiently low mass haloes, this timescale will be short enough that a 'stable virial shock' cannot develop, and gas accretion from the IGM will proceed as rapidly as dynamically allowed (Birnboim & Dekel, 2003a). However, this theoretical foundation is largely based on a one-dimensional picture, while dark matter haloes and their gaseous counterparts are decidedly not spherically symmetric. Numerical hydrodynamical simulations have indicated that galaxies can acquire their gas in a fundamentally different manner (Katz et al., 2003; Abadi et al., 2003; Kereš et al., 2005; Ocvirk et al., 2008). In particular, finding that coherent, filamentary inflows can fuel star formation in a central galaxy while avoiding shock heating to the virial temperature. Such streams arise naturally from the topology of large-scale structure, particularly at high redshifts of $z \gtrsim 2$ (Dekel et al., 2009; Agertz et al., 2009; Kereš et al., 2009; Danovich et al., 2012).

Gas inflow has been studied in the context of many key questions in galaxy formation, most fundamental of which is perhaps its link to star formation (Oppenheimer

et al., 2010; Gabor & Bournaud, 2014; Sánchez Almeida et al., 2014) and the process of quenching (Birnboim & Dekel, 2003b; Birnboim et al., 2007; Gabor & Davé, 2012; Feldmann & Mayer, 2015; Aragon-Calvo et al., 2014). The accretion of cosmological gas will leave a fundamental imprint on the thermal and dynamical properties of the quasi-static halo gas, as will outflows from energetic feedback processes in galaxies (Putman et al., 2012). By studying the properties of this halo gas we can investigate the interplay of inflows and outflows with metals (e.g. Shen et al., 2013; Hummels et al., 2013; Ford et al., 2014) as well as with neutral hydrogen (e.g. Faucher-Giguère et al., 2011; Faucher-Giguere et al., 2014; Fumagalli et al., 2014).

The state of gas in and around galaxy haloes has also received significant observational scrutiny in the local universe as well as at $z \sim 2$, near the peak of the cosmic star formation rate. At these high redshifts both hydrogen and metals are accessible as absorption signatures in sightlines towards background objects, enabling a probe of the interaction between galaxies and the IGM in the vicinity at this period of rapid stellar growth. Efforts include the 'quasars probing quasars' series (Hennawi et al., 2006; Prochaska et al., 2013, 2014), the Keck Baryonic Structure Survey (Steidel et al., 2010; Rudie et al., 2012, 2013; Turner et al., 2014b), and other $2 \lesssim z \lesssim 3$ studies (Simcoe et al., 2006; Pieri et al., 2014; Rubin et al., 2014; Crighton et al., 2015). Collectively they probe the covering fractions, radial profiles, and kinematics of HI and many metal ions including those of oxygen, carbon, neon, silicon, and magnesium.

Perhaps the most puzzling discovery which remains at the present unreconciled with theory is the presence of a large amount of cold ($\sim 10^4$ K), metal-enriched gas widely distributed at least out to the virial radius of massive ($10^{12} \sim 10^{12.5}$ $M_\odot$) haloes. For instance, Prochaska et al. (2014) find a covering fraction approaching unity for strong C

II absorption at a hundred physical kiloparsecs or more from the halo centre. The origin of this gas and the process by which it is either maintained or replenished within the hot halo remains uncertain. Consequently, this regime provides a powerful test-bed for current hydrodynamic simulations of galaxy formation. Not only in terms of the physical modelling of feedback and the resultant galactic-scale outflows (Muratov et al., 2015), but more fundamentally in terms of our ability to resolve the gas-dynamical processes and the spatial scales relevant for the physics of cosmological gas accretion.

This paper investigates the thermal and dynamical structure of halo gas in eight simulated $\simeq 10^{12}$ $M_\odot$ haloes at $z=2$. In Section 4.2 we describe the simulation technique and analysis methodology. Section 4.3 addresses the issue of resolution for halo gas and presents a visual overview of the systems. Section 4.4 considers the radially averaged gas properties, while Section 4.5 expands this analysis to explore the angular variability of halo structure. We discuss our results in the context of observations of the gas content of haloes in Section 4.6 and summarize our conclusions in Section 5.7.

4.2 Methods

4.2.1 Initial Conditions

All simulations evolve initial conditions which are a random realization of a WMAP-9 consistent cosmology ($\Omega_{\Lambda,0} = 0.736$, $\Omega_{m,0} = 0.264$, $\Omega_{b,0} = 0.0441$, $\sigma_8 = 0.805$, $n_s = 0.967$ and $h = 0.712$). We use the MUSIC code (Hahn & Abel, 2011, v1.5, r375) to generate multi-mass 'zoom' ICs, under the 2LPT approximation and with a tabulated transfer function from CAMB (Lewis et al., 2000). First, we evolve a low resolution, dark matter

Table 4.1: General characteristics of our three resolution levels, L9, L10, and L11. First, the effective resolution of an equivalent uniform box. Next, the mean number of high resolution gas elements, number of timesteps, baryonic mass resolution, and dark matter mass resolution. The Plummer equivalent comoving gravitational softening lengths, and their physical values at $z = 2$. The minimum gas cell spatial size in physical parsecs at $z = 2$, and the mean gas cell spatial size in the halo, between $0.5\,r_{\rm vir}$ and $1.0\,r_{\rm vir}$, in physical kiloparsecs at $z = 2$.

Res	$N_{\rm part}^{\rm eff}$	$N_{\rm part}^{\rm HR}$	Δt	$m_{\rm baryon}$ [$M_\odot$]	$m_{\rm DM}$ [$M_\odot$]	$\epsilon_{\rm grav}^{\rm co}$ [pc]	$\epsilon_{\rm grav}^{z=2}$ [pc]	$r_{\rm cell}^{\rm min}$ [pc]	$r_{\rm cell}^{\rm halo}$ [kpc]
L9	512^3	800,000	80,000	1.0 x 10^6	5.1 x 10^6	1430	480	31	2.7
L10	1024^3	7,000,000	260,000	1.3 x 10^5	6.4 x 10^5	715	240	11	1.6
L11	2048^3	64,000,000	870,000	1.6 x 10^4	8.0 x 10^4	357	120	3.3	0.8

only uniform periodic box of side-length $20h^{-1}$ Mpc $\simeq$ 28.6 Mpc with 128^3 particles ('L7', where LN=2^N), from a starting redshift of $z = 99$ down to $z = 2$. At this redshift, there are 20 haloes with total mass between $10^{11.8}\,M_\odot$ and $10^{12.4}\,M_\odot$ from which we choose eight at random to re-simulate at higher resolution. We do not select for any additional properties – e.g., merger history or environment. All particles within some factor of the virial radius of each selected halo (ranging from $3.6\,r_{\rm vir}$ to $7.0\,r_{\rm vir}$ in all cases, see Oñorbe et al. (2014) for relevant considerations) are identified at $z = 2$. We take $r_{\rm vir} = r_{\rm 200,crit}$ the radius enclosing a mean overdensity 200 times the critical density. This factor was chosen by trial and error with evaluation of contamination levels in low resolution test runs. The convex hull of the $z = 99$ positions of all selected particles is then used to define the high resolution refinement region.

For each halo, new initial conditions are generated for each of L9, L10, and L11, corresponding to 512^3, 1024^3, and 2048^3 total particles if the parent box were to be

simulated at a uniform resolution. We note that the mass resolution of L11 is between Aquarius levels '3' and '4' (of e.g. Marinacci et al., 2014a; Scannapieco et al., 2012), while the resolution of L9 is approximately equal to the resolution in modern, large volume cosmological simulations. Baryons are included by splitting each dark matter particle according to the cosmological baryon fraction, into one DM particle and one gas cell, such that the centre of mass position and velocity are preserved. Therefore we do not consider a separate transfer function for the baryonic component. The files required to generate our initial conditions, including the CAMB transfer function, noise seeds, and convex hull point sets, are made publicly available online[1]. The fundamental characteristics of each resolution level are given in Table 4.1, while the physical properties and numerical details for each of the eight haloes are detailed in Table 4.2.

4.2.2 Simulation Code and Physics

We employ the AREPO code (Springel, 2010a, r25505) to solve the coupled equations of ideal continuum hydrodynamics and self-gravity. An unstructured, moving, Voronoi tessellation of the domain provides the spatial discretization for Godunov's method with a directionally un-split MUSCL-Hancock scheme (van Leer, 1977) and an exact Riemann solver, obtaining second order accuracy in space. Since we allow the mesh generating sites to move, with a velocity equal to the local fluid velocity field modulated by corrections required to maintain the regularity of the mesh, this numerical approach would be classified as an Arbitrary Lagrangian-Eulerian (ALE) scheme. Gravitational accelerations are computed using the Tree-PM approach, where long-range forces are

[1] `http://www.illustris-project.org/files/Nelson15b_ICs.zip`

Table 4.2: Details on the eight simulated haloes: the total halo mass and (physical) virial radius at $z = 2$ from the parent box. The radius of the enclosing sphere at $z = 2$ used to define the Lagrangian region, in terms of $r_{\rm vir}$, and the number of high resolution elements, for each of dark matter and gas. Both are listed for the L11 level only. The minimum radius from the halo centre reached by (i) contaminating low resolution dark matter particles and (ii) Monte Carlo tracers originating in low resolution gas cells, in units of $r_{\rm vir}$. The total number of timesteps to reach $z=2$.

Halo	$M_{\rm halo}^{\rm par}$	$r_{\rm vir}^{\rm par}$	$r_{\rm HR}^{\rm L11}$	$N_{\rm HR}^{\rm L11}$	$r_{\rm LR}^{\rm min}$	$r_{\rm LR}^{\rm min}$	Δt
#	[log $\rm M_\odot$]	[kpc]	[$r_{\rm vir}$]	[10^6]	$[r_{\rm vir}]_{\rm dm}$	$[r_{\rm vir}]_{\rm tr}$	#
h0	12.1	114	3.6	70.0	1.77	2.16	829714
h1	12.1	104	4.8	66.7	2.12	2.75	701681
h2	11.9	92	6.0	24.2	2.83	3.02	955189
h3	11.9	96	7.0	33.9	2.74	3.23	812983
h4	12.0	103	6.0	68.4	2.13	2.89	861224
h5	12.0	103	4.2	59.9	1.04	1.04	931088
h6	12.1	97	4.8	74.4	1.32	1.59	980918
h7	11.9	94	4.4	52.8	0.94	1.93	866242

calculated with a Fourier particle-mesh method, medium-range forces with a hierarchical tree algorithm (Barnes & Hut, 1986), and short-range forces with direct summation (as in Springel, 2005). A local, predictor-corrector type, hierarchical time stepping method obtains second order accuracy in time. Numerical parameters tangential to our current investigation – for example, related to mesh regularization or gravitational force accuracy – are detailed in Springel (2010a) and Vogelsberger et al. (2012).

We include a redshift-dependent, spatially uniform, ionizing UV background field (Faucher-Giguère et al., 2009). Gas loses internal energy from optically thin radiative

cooling assuming a primordial H/He ratio (Katz et al., 1996). The production of metals and metal line cooling contributions are not included.Star formation and the associated ISM pressurisation from unresolved supernovae are included with an effective equation of state modelling the ISM as a two-phase medium, following Springel & Hernquist (2003). Gas elements are stochastically converted into star particles when the local gas density exceeds a threshold value of $n_{\rm H} = 0.13$ cm^{-3}. Furthermore, there is no resolved stellar feedback that would drive galactic-scale winds, nor any treatment of black holes or their associated feedback. All of the simulations considered in this work disregard the possible effects of radiative transfer, magnetic fields, and cosmic rays. The set of implemented physics is essentially identical to the 'moving mesh cosmology' paper series as described in Nelson et al. (2013).

We identify dark matter haloes and their gravitationally bound substructures using the SUBFIND algorithm (Springel et al., 2001a; Dolag et al., 2009) which is applied on top of a friends-of-friends cluster identification. We refer to the most massive substructure in each FoF group as the halo itself. Merger trees are constructed using the SUBLINK code (Rodriguez-Gomez et al., 2015) to link haloes to their progenitors and descendants at different points in time.

All runs include the Monte Carlo tracer particle technique (Genel et al., 2013) in order to follow the evolving properties of gas elements over time, with five tracers per initial gas cell. This is a probabilistic method, where tracers are exchanged between parents based explicitly on the corresponding mass fluxes. By locating a subset of their unique IDs at each snapshot we can, by reference to their parents at that snapshot, reconstruct their spatial trajectory or thermodynamic history.

In the current work we make limited use of the assembly and accretion histories of the haloes, as determined with the tracer particles and merger trees, respectively. Specifically, we verify the correspondence between haloes at different resolution levels in order to comment on the convergence properties of our simulations. We reserve as future work – the second paper in this series – a quantitative analysis of the rates and modes of accretion and the impact of merging substructures.

4.3 Resolution Considerations and Visual Inspection

In the vast majority of galaxy formation simulations, spatial resolution is naturally adaptive and follows the hierarchical clustering of structure formation in ΛCDM. This is true for the dark matter, where the Vlassov-Poisson equations are solved with a Monte-Carlo approach. The N-body method is typically used (for notable exceptions see Yoshikawa et al., 2013; Hahn & Angulo, 2015). This natural adaptivity also holds for the gaseous component in particle methods like SPH, adaptive grid methods with the typical density-based refinement criteria (for notable exceptions see Iapichino et al., 2008; Rosdahl & Blaizot, 2012), and the moving-mesh code used in this work. The resolution is therefore better in collapsed structures than it is in under-dense regions, and maximal in galaxies themselves. The gaseous haloes surrounding galaxies are poorly resolved in comparison, to the extent that modern cosmological simulations (e.g. Khandai et al., 2014; Dubois et al., 2014; Vogelsberger et al., 2014b; Schaye et al., 2015) which just barely resolve the internal structure of individual galaxies may be insufficiently capturing gas-dynamical processes in the halo.

We begin by quantifying the spatial resolution of gas in the halo. Figure 4.1 shows

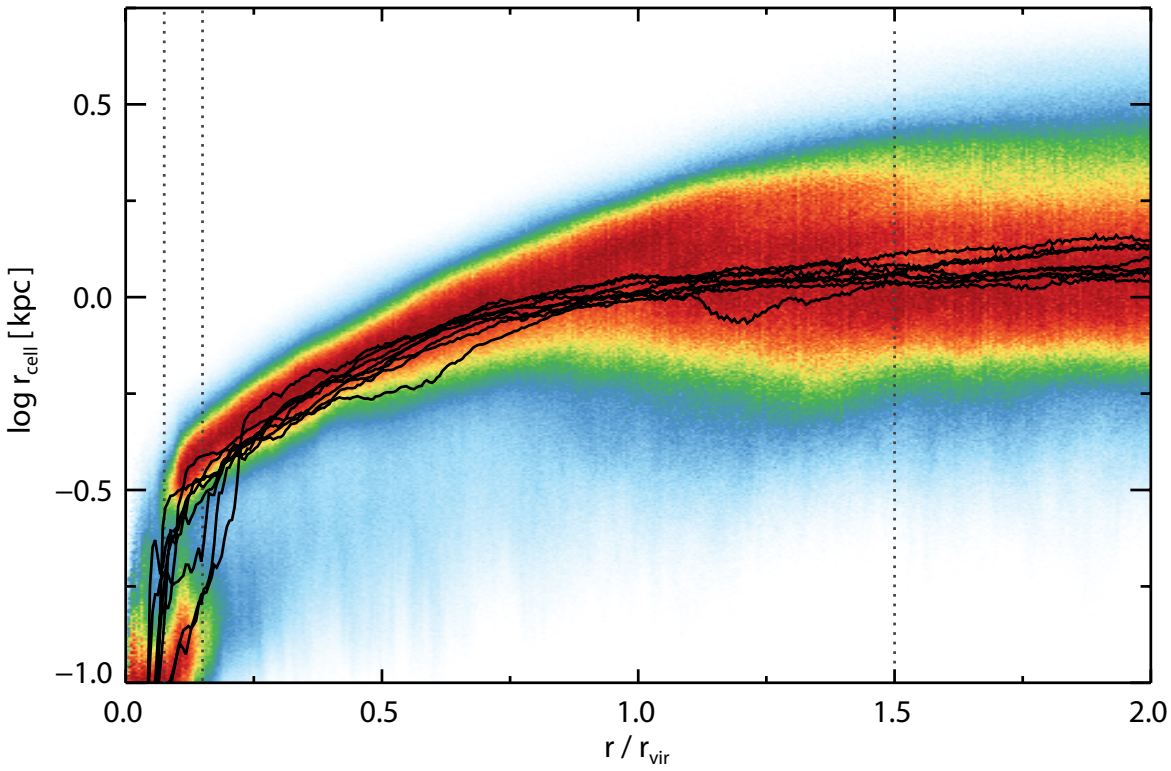

Figure 4.1 The mass-weighted distribution of gas spatial resolution as a function of radius, for all eight halos stacked together at L11 resolution. As a proxy for spatial resolution, we show the gas cell sphere-equivalent radii $r_{\rm cell}$. Each radial bin is normalized independently, such that the colour mapping reaches its maximum intensity at each radius, independent of the radial distribution of gas mass. The dark vertical lines indicate $1.5\,r_{\rm vir}$, $0.15\,r_{\rm vir}$ and $0.07\,r_{\rm vir}$, the last two bounding the radial range where a density bimodality arises due to the central galaxy. Substructures are excised.

the mass-weighted two dimensional distribution of cell size $r_{\rm cell}$ as a function of radius. As a proxy for the irregular shape of Voronoi polyhedra, we take the sphere-equivalent radius $r_{\rm cell} = (3V_{\rm cell}/4\pi)^{1/3}$ where $V_{\rm cell}$ is the exact Voronoi cell volume. The mesh regularization scheme ensures that this value is approximately equal to the actual geometrical distance between the generating point and each face centre. All haloes are stacked together at the highest (L11) resolution level, while black lines show the median relation for each halo separately.

Beyond $r/r_{\rm vir} > 1.25$ the hydrodynamic resolution is nearly constant, but with a broad distribution spanning from $\sim 500\,{\rm pc}$ to $\sim 3\,{\rm kpc}$ (physical), depending on what level of overdensity the gas cell resides in. In the halo region, $0.25 < r/r_{\rm vir} < 1.0$ the mean resolution scales from $\sim 400\,{\rm pc}$ up to $\sim 1\,{\rm kpc}$. The inner halo, typically bounded by $r \lesssim 0.15 r_{\rm vir}$ sees the mean cell size decrease rapidly under $\sim 100\,{\rm pc}$ as dense gas features associated with the central galaxy begin to dominate. We return to the implications of the relatively coarse halo resolution in the discussion.

In Figure 4.2 we show a visual example of the resolution issue and a first look at the impact of higher hydrodynamical resolution in the halo regime. One halo (h0) is included at the three increasing zoom levels considered in this work, L9, L10, and L11, increasing from left to right. On the top row we plot a slice of the Voronoi mesh used to evolve the gas component.[2] The slice is centred on the central galaxy, which lies at the origin, and to focus on details we include only the upper-right quadrant. The outer red circle marks the virial radius of 114 physical kpc at $z = 2$. At all three resolution

[2]Note that the intersection of a plane with the three-dimensional Voronoi tessellation does not in general produce a two-dimensional Voronoi graph. Therefore the shapes of individual cells, as shown, do not reflect the actual regularity of the computational mesh.

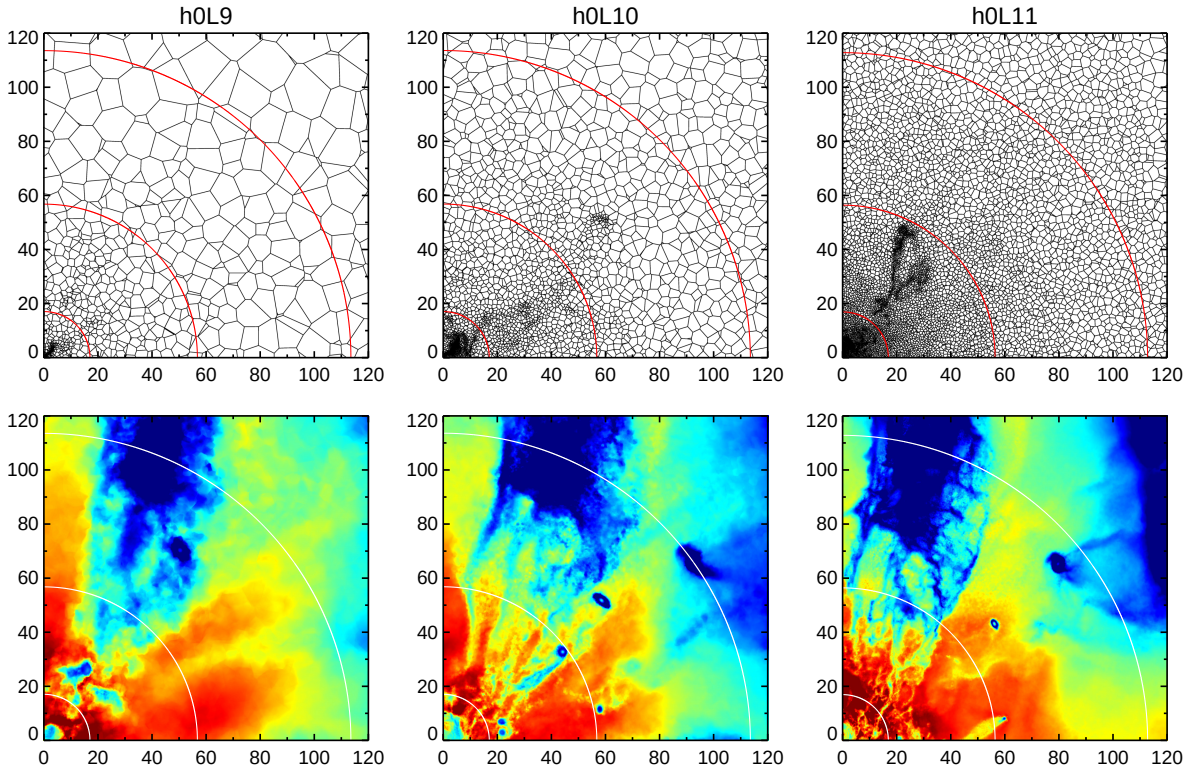

Figure 4.2 Slice through the Voronoi mesh showing the underlying spatial discretization (top) and a projection through the linearly reconstructed, mass-weighted gas temperature field (bottom) with $\sim 10^4\,\mathrm{K}$ as blue and $\sim 10^6\,\mathrm{K}$ as red. Halo h0 is shown at $z = 2$. We include only the upper right quadrant, after centring the halo at the origin, in orthographic projection. The circles denote $\{0.15,0.5,1.0\}\,r_{\mathrm{vir}}$, the axes units are in physical kpc, and the projection depth is 238 kpc. The average gas cell size near the virial radius drops with increasing spatial resolution (from left to right), from ∼2.7 kpc to ∼0.8 kpc physical. Note that the exact state of gas in the halo, and the location of satellites, differs somewhat due to timing offsets between the three runs.

levels we see that the high-density gas near the origin is resolved at the sub-kpc level and impossible to separate at this scale, while at the virial radius individual gas elements are substantially larger. At L9, characteristic of modern cosmological simulations, the virial arc across a quadrant is only resolved by $\simeq 20$ cells, while the radial structure between $0.5\,r_{\rm vir}$ and $1.0\,r_{\rm vir}$ is only resolved by $\simeq 10$ cells. This is roughly consistent with a zeroth order estimate, assuming uniform density, for a number of cells per dimension of

$$N_{\rm cell,dim} \sim \left(\frac{(\Omega_b/\Omega_m) \times M_{\rm halo}}{m_{\rm cell}} \right)^{1/3} \simeq 50 \tag{4.1}$$

which are available to resolve the halo gas structure of a 10^{12} $M_\odot$ halo. At L11 this increases to $N_{\rm cell,dim} \simeq 250$, but the scaling is obviously slow in three-dimensions. In the discussion we consider a possible method for improving upon this resolution issue in the halo with future work.

Across the bottom row we show a projection of mass-weighted gas temperature, where the colormap extends from cold, $\sim 10^4\,$K gas as blue to hot, $\sim 10^6\,$K gas as red. Some interesting differences emerge as the gas resolution in the halo increases. The large cold inflow which begins to experience substantial heating at $(0.4-0.5) \times r_{\rm vir}$ shows some temperature inhomogeneities at L9, but is essentially a single coherent structure. At L10, this inflow is resolved into a number of smaller gas filaments by the time it crosses half the virial radius. At L11 the temperature structure becomes even more complex, throughout the halo, with small scale features emerging below the cell-size of the lower resolution runs. Clearly the morphology of accreting gas and its interaction with the quasi-static hot halo material depends on numerical resolution to some degree. In the following sections we further investigate this dependence by analysing the structure of halo gas at our highest available resolution.

4.3.1 Visual Inspection

We present a visual overview of the eight simulated haloes at $z = 2$, focusing on the virial scale of haloes 0-3 in Figure 4.3, and of haloes 4-7 in Figure 4.4, one halo per row. In each case, we show mass-weighted projections of gas density, temperature, entropy, and radial velocity in each column (left to right). All gas cells within a cube of side-length $3\,r_{\rm vir}$ are included, while the temperature of star forming gas on the effective equation of state (eEOS) is set to a constant value of 3000 K. The three concentric circles denote $\{0.15, 0.5, 1.0\} \times r_{\rm vir}$.

We have separated the eight systems into two groups based roughly on the morphology of their halo gas – the first four haloes (h0-3) are closer to equilibrium at $z = 2$ and the structure of their hot gas is roughly spherical, whereas the second four (h4-7) are significantly more disturbed. In the first case, the virial shock can be clearly seen as a sharp increase to higher temperature and entropy, typically at $\gtrsim 1.25 r_{\rm vir}$. However, the radius of the shock (see also Schaal & Springel, 2015) or its existence at all depends strongly on direction, which we explore further in Section 4.5. These boundaries are also associated with a decreased inflow velocity – that is, radial stalling – as well as increased gas density.

At $z = 2$ haloes at this mass scale of $\simeq 10^{12}$ $M_\odot$ typically reside at the intersection of multiple large-scale filaments (and/or sheets) of the cosmic web. This naturally leads to a significant amount of filamentary inflow across the virial radius. We see that these inflows cover only a fraction of the virial sphere. They can maintain coherency to at least half the virial radius while maintaining a strong overdensity with respect to the mean gas density at each radius. They experience significant heating, typically at $\sim 0.5 r_{\rm vir}$,

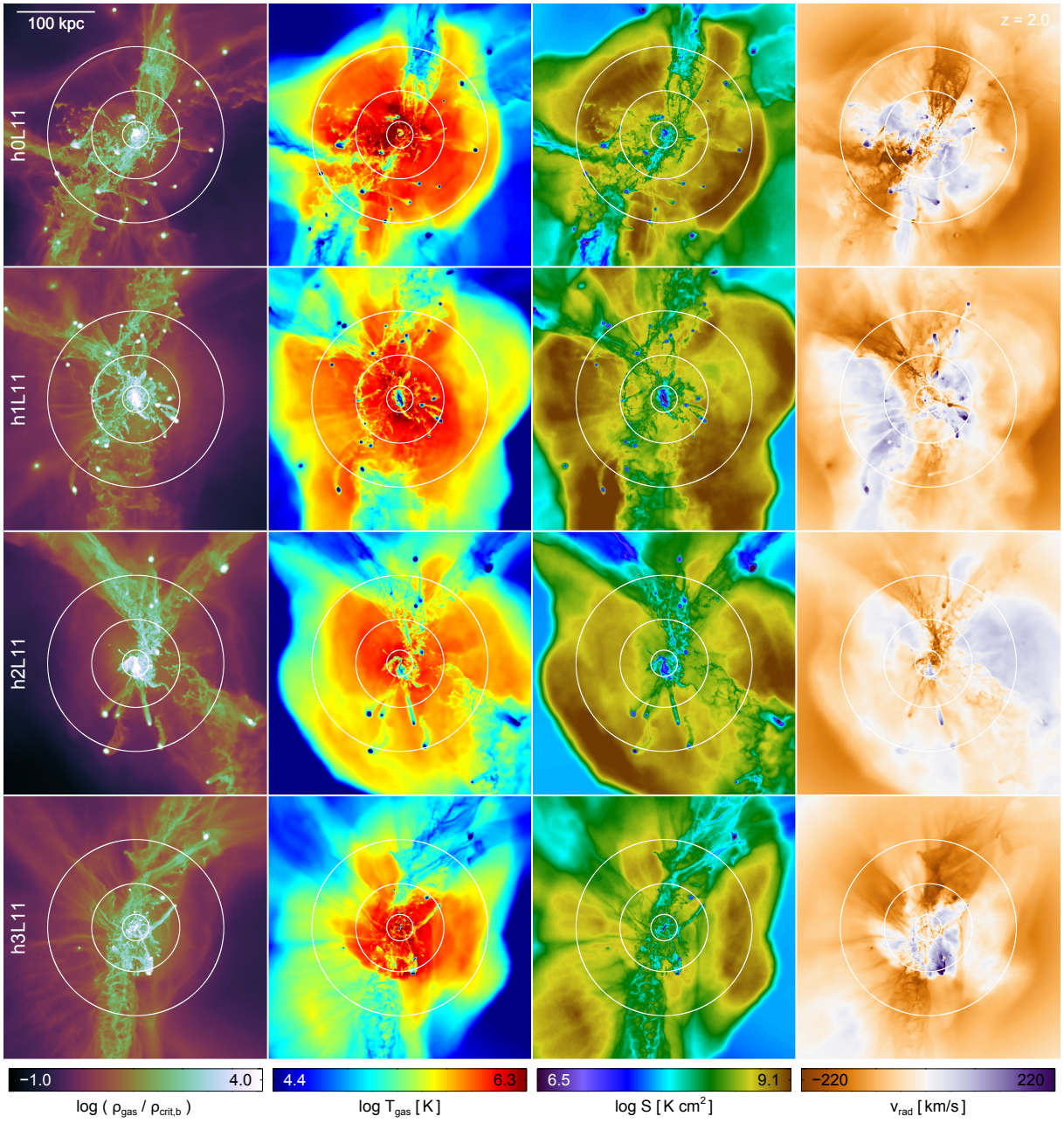

Figure 4.3 Mass-weighted projections of gas density, temperature, entropy, and radial velocity for the first four simulated haloes at $z=2$. In each case, all gas cells within a cube of side-length $3\,r_{\rm vir}$ are included, and distributed using the standard cubic spline kernel with $h = 2.5\,r_{\rm cell}$ in orthographic projection. The white circles denote $\{0.15,0.5,1.0\}\,r_{\rm vir}$. Gas density is normalized to the critical baryon density at $z=2$. The temperature of star forming gas on the eEOS is set to a constant value of 3000 K.

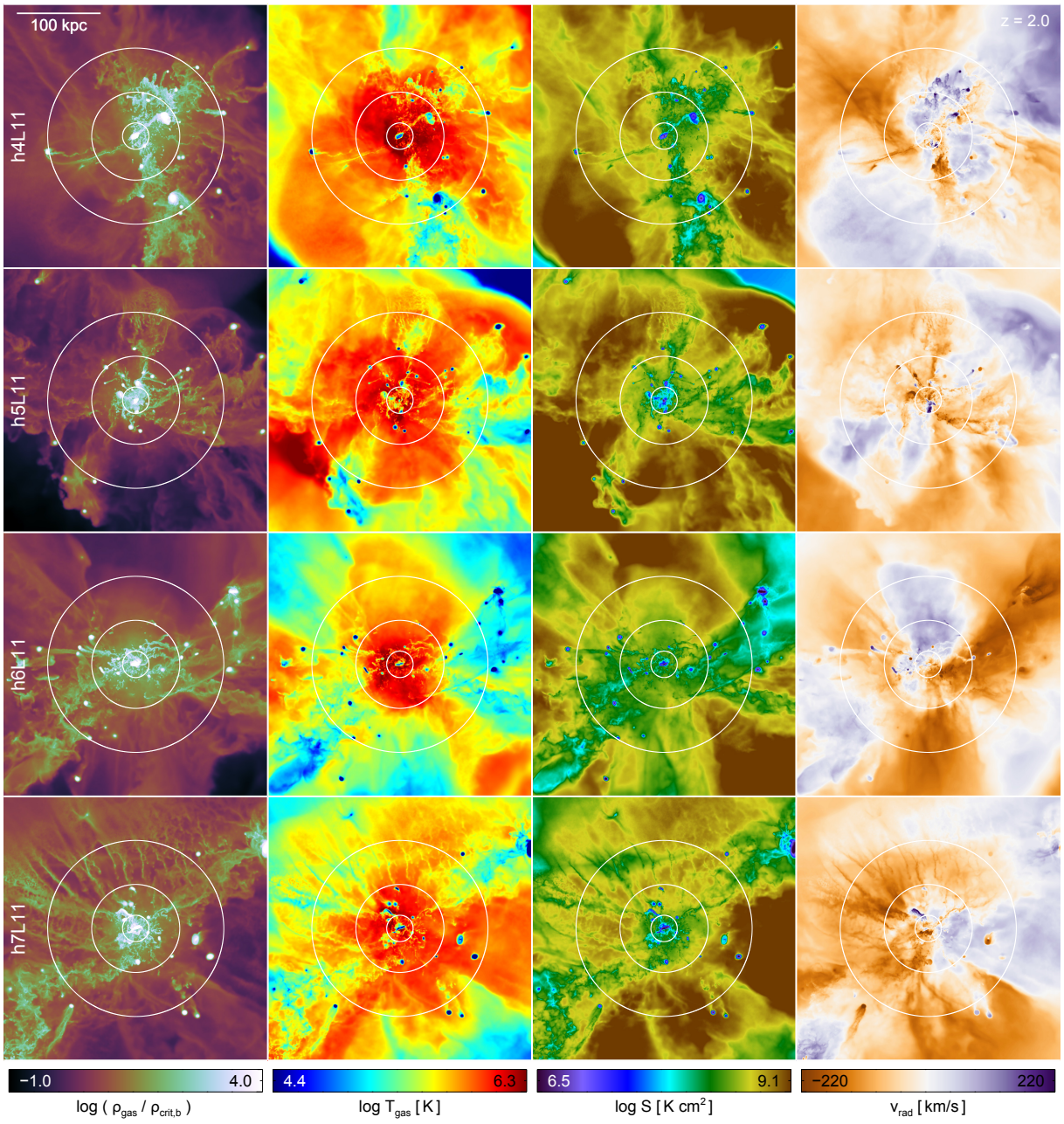

Figure 4.4 As in the previous figure, mass-weighted projections of gas density, temperature, entropy, and radial velocity at $z = 2$, but for the other four simulated haloes.

from $\lesssim 10^{4.5}\,\mathrm{K}$ to $\gtrsim 10^{6}\,\mathrm{K}$, reaching the peak temperature of the mean $T_{\mathrm{gas}}(r)$ in the inner halo. Their entropy can remain lower than that of the hot halo gas at these radii, $\sim 10^{8.0-8.5}\,\mathrm{K\,cm^2}$ compared to $\sim 10^{9}\,\mathrm{K\,cm^2}$. Although the mean radial velocity inside the halo is generally near equilibrium/zero (denoted by white in the colormap), we see that in a given direction this is rarely the case, finding instead either non-negligible inflow or outflow velocities.

Two haloes, h3 and h7, have a large number of small, prominent, inflowing streams in the radial range $0.5r_{\mathrm{vir}} \lesssim r \lesssim 1.25r_{\mathrm{vir}}$ (see density maps). They originate from the lower left direction in the case of h3, and from the top for h7. These filaments are not associated with large-scale structure, and have characteristics distinct from the much larger filaments which are associated with features at larger radii. In particular, although they obtain similarly high inflow velocity, their entropy and temperature are above that of the IGM, and their overdensity with respect to the radial mean is not as significant. Some of the other six haloes exhibit similar features at various points in the past, but they have disappeared by $z = 2$. In general, we observe that these features form between one and two times the virial radius semi-spontaneously, as in an instability, commonly triggered by perturbations either from substructure debris or the intersection of sheet-like structure in the cosmic web. They are marginally evident at L10 and absent at L9. We reserve an in-depth analysis of the formation of these features and whether or not their growth corresponds to a physical instability in the gas for future work.

We visualize two additional scales, the inner halo structure as well as the large-scale context, of the eight simulated haloes at $z = 2$. Figure 4.5 includes haloes 0-3 while Figure 4.6 includes haloes 4-7. The left two columns show the projected density of the gas and stellar components, the white circles indicating $0.05\,r_{\mathrm{vir}}$ and $0.15\,r_{\mathrm{vir}}$. The right

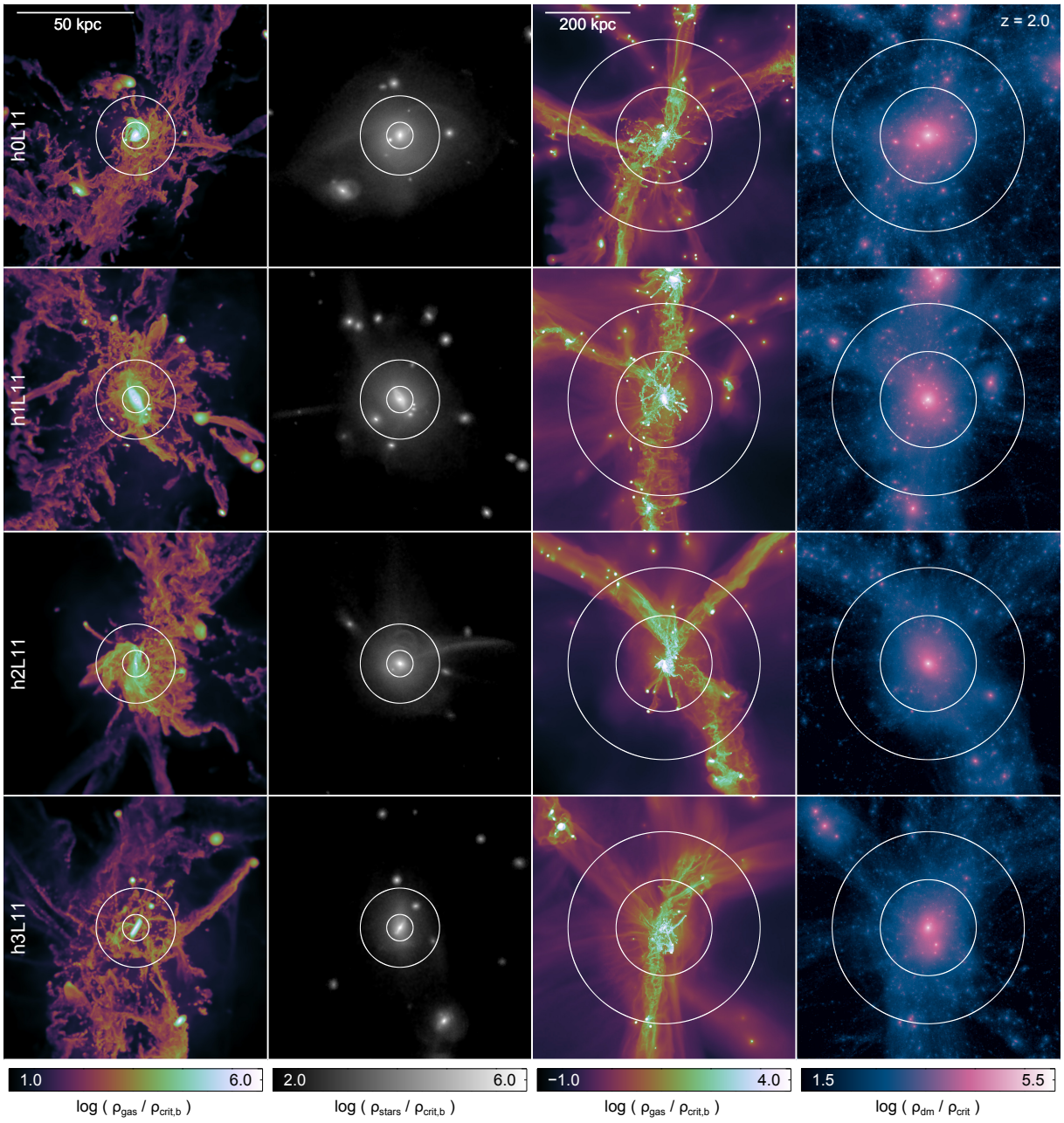

Figure 4.5 Mass-weighted projections of gas and stellar density temperature at small scales (left two columns), and projected gas density and dark matter density at large scales (right two columns). The first four simulated haloes are shown at $z = 2$. All gas cells or particles within a cube of side-length $1.0\,r_{\rm vir}$ (small scales) or $5.5\,r_{\rm vir}$ (large scales) are included, and distributed using the standard cubic spline kernel with $h = 2.5\,r_{\rm cell}$ (gas) or $h = r_{32,\rm ngb}$. The white circles denote $\{0.05,0.15\}\,r_{\rm vir}$ (left two columns) and $\{1,2\}\,r_{\rm vir}$ (right two columns). Densities are normalized to the critical (baryon) density at $z = 2$.

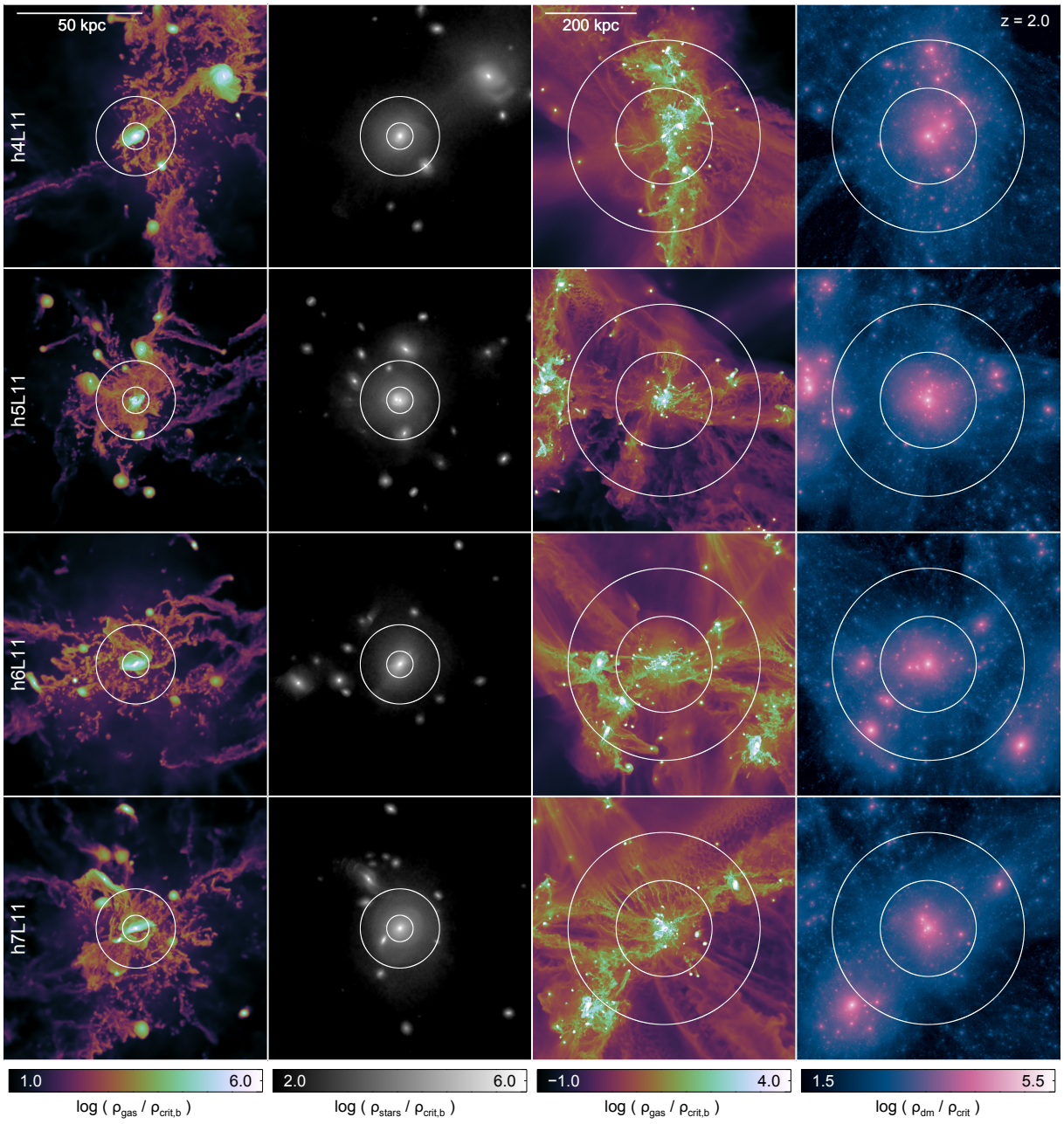

Figure 4.6 As in the previous figure, projections of gas and stellar density on small scales (left two columns), and gas density and dark matter density on large scales (right two columns) at $z = 2$, but for the other four simulated haloes.

two columns show the projected density of the gas and dark matter components, the white circles here indicating $1.0\,r_{\rm vir}$ and $2.0\,r_{\rm vir}$.

In the centre of the halo, $r/r_{\rm vir} \lesssim 0.3$, the gas component forms a complex morphology with multiple orbiting substructures, commonly associated with tails and gas streams from ram-pressure stripping, as well as tidally-induced bridges and spiral patterns. Rapidly varying velocity fields modulate inflowing as well as outflowing material. At $r/r_{\rm vir} \lesssim 0.05$ a large gas disc is a common feature – the final stage in the transformation from large-scale structure and the associated tidal torques to the formation of a galaxy. Here the coherent angular momentum of the halo as a whole must transition to the angular momentum of the forming disc through the complicated dynamics of this inner halo 'messy' region (Danovich et al., 2014). In the small-scale views of h2 we see tidal features in the collisionless stellar distribution (extending to the right from the centre), without any corresponding structure in the gas. In h3, we see a massive tidal tail in gas extending out to $\simeq 0.5 r_{\rm vir}$ (towards upper right of panel), with no corresponding structure in the stars. A recent post-pericenter passage of an ongoing major merger in h4 has induced tidal features in the large companion (upper right of figure) which is still on an outgoing orbit, and generated a long tidal tail, of which the inner half has reversed direction and is falling back on to the central galaxy. A recent merger in h5 has left a double nucleus in the stars, separating by roughly one kpc. In h7 we see a strong tidal warp in the gas and stars of the merging companion (to the upper left of the central), with several stellar shells from earlier minor mergers superimposed. In general, we can see that many of the narrow inflowing streams in the inner halo are tidal in nature, which highlights the importance of characterizing the origin of gas accretion features – this component of the accretion rate would have been characterized

as a 'stripped' component in the methodology of Nelson et al. (2015).

In the inner regions of the halos which we have identified as in equilibrium, the perturbations and gas structures generated by the dynamics of satellite galaxies (see Zavala et al., 2012) are superimposed on a gas background which is aligned at Mpc scales with large-scale gas features. For example, the broad inflow in h0 (towards the top of the image panel) with a width of $\sim 50\,\mathrm{kpc}$ at $0.5\,r_{\mathrm{vir}}$ is clearly associated with a gas overdensity extending out to at least $3\,r_{\mathrm{vir}}$. Likewise for the broad inflow in h2 (also, top of the image panel) which has coalesced at the intersection of two such cosmic web filaments. These inflows are not smooth features of constant density gas. They have a 'wavy' or rippled appearance, arising from gas density contrasts as high as a factor of $\sim 10^3$. By visual inspection it is clear that, at least at sufficient resolution, such inflowing filaments are in fact associations of multiple smaller structures with coherent kinematics and origin.

The second set of haloes (h4-7), which are in general more morphologically disturbed, do not show this clear association between small-scale and large-scale gas features. Their time evolution indicates that this is often a consequence of a major halo-halo merger where the two gas reservoirs collide and disrupt any radially coherent gas kinematics as the hot halo re-equilibrates. However, recent major mergers are not always seen, and the nearby environment and topology of large-scale structure also appear to play a significant role in the sphericity of halo gas, independent of assembly history. For example, we consider the number of major mergers with ratio $\eta \geq 1/3$ experienced by each halo while the stellar mass is greater than $10^{10.5}$ $\mathrm{M}_{\odot}$ (roughly $2 < z < 4$), as well as the maximum merger ratio of any merger, under the same constraint. In both cases, the mass ratio is determined at the time the less massive progenitor reaches its maximum stellar mass (see

Rodriguez-Gomez et al., 2015). We find that h0, h2, and h3 have had no major mergers exceeding this mass ratio, and indeed none with $\eta \geq 0.2$. The outlier is h1, which did have one major merger ($\eta = 0.8$). In the disturbed group, h4 has had four major mergers ($\eta_{\rm max} = 0.6$), h6 has had one ($\eta = 0.65$), and h7 has had three ($\eta_{\rm max} = 0.95$). The outlier is h5, which has had no major mergers satisfying this criterion, the largest mass ratio being $\eta = 0.2$.

4.4 The Physical State of Halo Gas

In this Section, we quantify the gas behaviours noted above, first by giving the spherically-averaged profiles of gas density, cell size, temperature, entropy, radial velocity, and specific angular momentum, and then by examining the pairwise correlation matrix of these same gas quantities.

4.4.1 Characteristic Halo Properties

In what follows we will normalize the distance of gas cells from the halo centre by the virial radius, as calculated for each halo at each resolution. For simplicity, we will in general not normalize the thermodynamic properties of gas since our halo mass range is so narrow. Instead, to provide a sense of reference, we calculate representative values for gas temperature, entropy, density, velocity, and angular momentum for a halo of $M_{\rm halo} \simeq 10^{12}$ M$_\odot$ at $z = 2$. Numerical values are given here. First, we take the virial temperature as

$$T_{\rm vir} = \frac{\mu m_p v_{\rm vir}^2}{2k_B} \simeq 10^{6.3}\,{\rm K} \tag{4.2}$$

following Barkana & Loeb (2001), where $\mu \simeq 0.6$ for a fully ionized, primordial gas. We take entropy as $S = P/\rho^\gamma$ where pressure is $P = (\gamma - 1)T\rho$. A virial entropy is then defined as

$$S_{\rm vir} = \frac{P_{\rm vir}}{\rho_{vir}^\gamma} = (\gamma - 1)T_{\rm vir}\rho_{\rm vir}^{(1-\gamma)} \simeq 10^{8.0}\,{\rm K\,cm^2} \tag{4.3}$$

where we have taken $\rho_{\rm vir} = 200(\Omega_b/\Omega_m)\rho_{\rm crit}(z)$, the baryon fraction multiplied by two hundred times the critical density of the universe at that redshift. Using the chosen overdensity criterion, the virial radius is then

$$r_{\rm vir} = \left(\frac{GM_{\rm halo}}{100H(z)^2}\right)^{1/3} \simeq 100\,{\rm kpc} \tag{4.4}$$

in physical units, while the virial velocity is

$$v_{\rm vir} = \sqrt{\frac{GM_{\rm halo}}{r_{\rm vir}}} \simeq 205\,{\rm km/s}. \tag{4.5}$$

Finally, for the specific angular momentum of the halo we take

$$j_{\rm vir} = \sqrt{2}\lambda v_{\rm vir} r_{\rm vir} \simeq 10^{3.0}\,{\rm kpc\,km/s} \tag{4.6}$$

using a spin parameter of $\lambda = 0.035$ (Barnes & Efstathiou, 1987; Bullock et al., 2001). These characteristic values for the temperature, density, entropy, radial velocity, and specific angular momentum of the halo gas are typically reached at some reasonable fraction of the virial radius, e.g. $\sim (0.1 - 0.5) \times r_{\rm vir}$. We use them below to provide a useful reference with which to identify halo gas in different physical states.

4.4.2 Radial Gas Profiles

In Figure 4.7 we show the median radial profiles of six gas quantities: temperature, density, radial velocity, cell size, entropy, and specific angular momentum. The main panel plots the profile for each of the eight haloes individually (different line colours) at the L11 resolution level. Light colours are used to indicate the four 'equilibrium' haloes, while darker colours indicate the four more 'disturbed' systems. The subpanels show the ratio of the radial profiles comparing the L9/L11 and L10/L11 runs, where the thick black lines indicate the average across haloes. We have excised all gravitationally bound substructures before constructing mean radial profiles, which would otherwise appear as a forest of small, dense, cold systems in these figures. A few observationally and theoretically motivated scalings with radius are included for reference (dotted and dashed lines).

For instance, going beyond an isothermal model for gas temperature, we can take a gas component tracing dark matter distributed according to a NFW profile. This leads to a scaling of $T_{\rm gas} \propto M_{\rm enc}(r)/r$ (e.g. Makino et al., 1998), which is still much too shallow (dotted line). Instead for $r/r_{\rm vir} \gtrsim 0.15$ we have approximately $T_{\rm gas} \propto r^{-1/2}$. Of all eight systems, h0 exhibits the sharpest temperature drop defining an outer boundary of the halo at $r \simeq 1.2 r_{\rm vir}$, while in most cases there is no sharp transition at or near the virial radius. In all cases the temperature drops precipitously in the inner halo, at 5%-10% of $r_{\rm vir}$. Note that we have assigned a gas temperature of 3000 K to gas on the star-forming equation of state, which would otherwise have an artificially high effective temperature. The $T_{\rm gas}(r)$ profiles are well converged with resolution. The mean across all eight haloes and averaged over $0.15 < r/r_{\rm vir} < 1.5$ is L9/L11 $\simeq 1.15$ and L10/L11 $\simeq 1.08$, while at

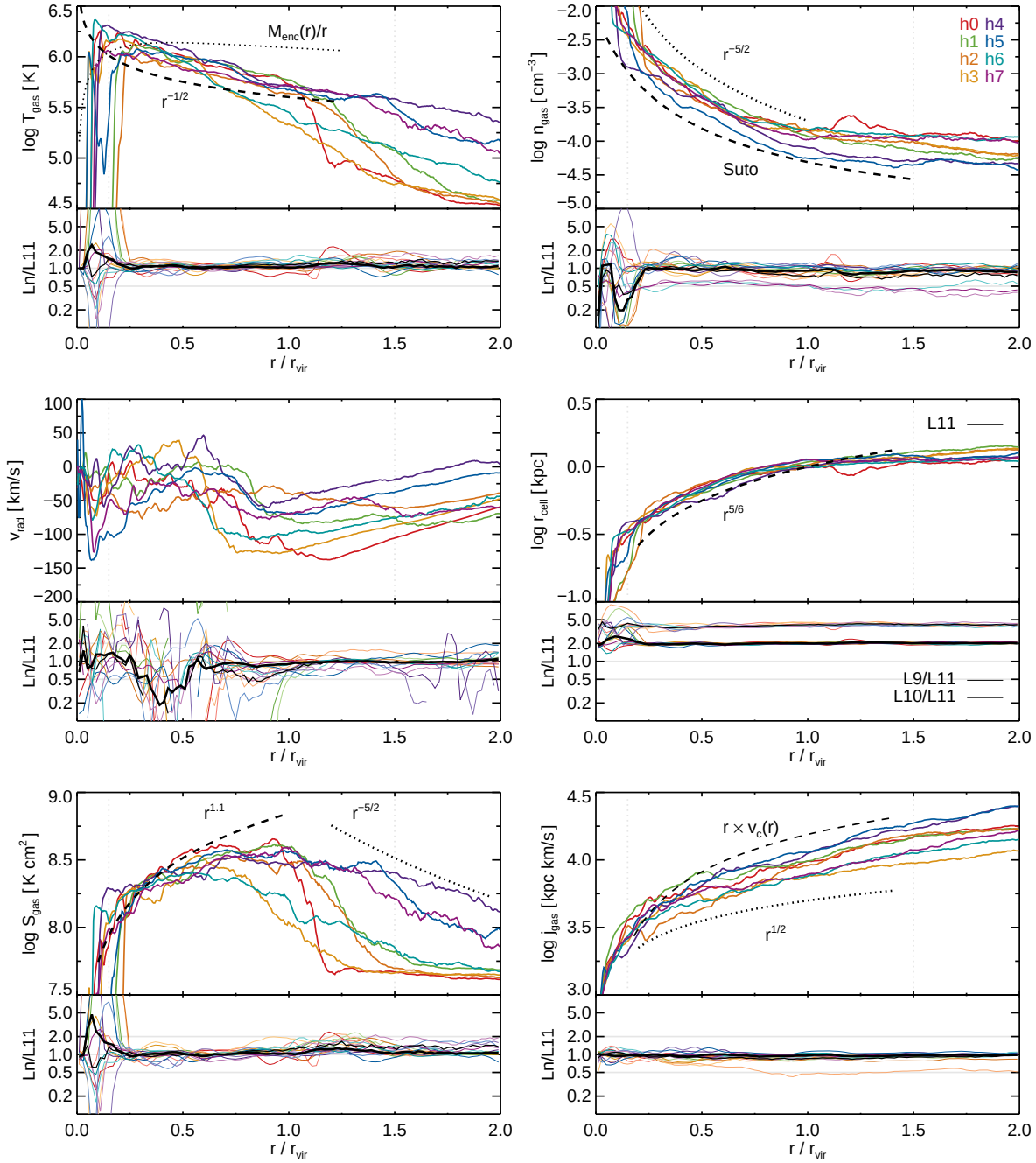

Figure 4.7 The median radial profiles of six quantities: gas temperature, density, radial velocity, cell size, entropy, and angular momentum. Each halo is shown separately (by colour). The main panels for each quantity show the highest resolution only, while the subpanels show the ratio of the two lower resolution runs to L11 (the mean ratios across all eight haloes are indicated with thick black lines). Various scalings with radius are provided for reference (dotted and dashed lines). Gas within $2r_{\rm vir}$ is included.

the largest radii the lower resolution runs have mean temperatures biased high by up to $\simeq 30\%$. Physics will dominate resolution – different models for galactic-scale stellar winds or AGN feedback can substantially modify the mean gas temperature profile inside as well as outside $r_{\rm vir}$ (Suresh et al., 2015).

A gas density scaling as $n_{\rm gas} \propto r^{-5/2}$ approximates the mean behaviour of the haloes well, as does a model for an isothermal gas in hydrostatic equilibrium (Suto et al., 1998, using $B = (n+1)B_p$ with $n = 7$, $B_p = 1$). However, the other panels make it clear that the gas is not in hydrostatic equilibrium, nor isothermal, implying this agreement may be coincidental. Convergence at lower resolutions is reasonable – over the same radial range we find L9/L11 $\simeq 0.8$ and L10/L11 $\simeq 0.9$, indicating that the mean gas densities are lower in the lower resolution runs. This is potentially a consequence of better resolving satellite galaxies and their interactions with the central host, which fills the halo volume with more high density gas cells. These are no longer instantaneously bound to the satellite, and so are not excluded. [3] The total halo gas mass actually decreases somewhat at higher resolution ($\sim 10\%$ or $\sim 5\%$ lower for L11 as compared to L9 or L10, respectively), which would otherwise modify gas densities in the opposite direction. Finally, the variable assembly histories lead to the large halo-to-halo scatter, while the sensitivity of merger states to temporal offsets driven by short dynamical time-scales leads to large scatter between resolution levels. As with temperature, however, the caveat is that radial density profiles will depend in detail on the implemented feedback models (Hummels et al., 2013).

[3]The filamentary feature in the upper right panel of Figure 4.2, extending from $\simeq 0.15\, r_{\rm vir}$ to $\simeq 0.5\, r_{\rm vir}$ is an example of such a tidal tail.

We see a similar scatter in the Hubble-corrected gas radial velocities. The haloes have different velocity structures, which commonly feature a slowly increasing inflow velocity down to $\sim 0.75\, r_{\rm vir}$, at which point there is a noticeable bump towards an equilibrium value of $v_{\rm rad} \simeq 0\,{\rm km/s}$. Until reaching the central galaxy, the radial velocity profiles are then roughly flat. In the following discussion of Figure 4.8, however, we show that the mean (or median) velocity profile is an exceedingly poor representation of the dynamics of halo gas. For example, in the mean the quasi-static halo rarely has $v_{\rm rad} > 0$ for $0.15 < r/r_{\rm vir} < 1.0$, while in actuality this value is driven down by the superposition of rapidly inflowing and gently outflowing components.

The gas spatial resolution, for which we again use the sphere-equivalent radii $r_{\rm cell}$ as a proxy, becomes better with decreasing radius as expected. Given our constant cell mass refinement criterion, it scales roughly as the cube-root of the density scaling. That is, $r_{\rm cell} \propto V_{\rm cell}^{1/3} = (m_{\rm cell}/\rho_{\rm cell})^{1/3} \propto \rho_{\rm cell}^{-1/3}$. At L9 and L10 the cell sizes are a factor of four and two larger, respectively, scaling with the mean inter-cell spacing as $r_{\rm cell} \propto L_{\rm box} N_{\rm cell,tot}^{-1/3}$.

We overplot the gas entropy profiles with the near-linear scaling $S_{\rm gas} \propto r^{1.1}$ often seen in x-ray observations of local clusters (e.g. George et al., 2009). Despite looking at the significantly more massive cluster PKS 0745-191, $\sim 10^{14}$ M$_\odot$ in the nearby universe ($z \simeq 0.1$), the inferred radial characteristics from George et al. (2009) are in reasonable agreement with the haloes simulated in this work. A similar level of qualitative agreement is evident at $z = 2$ for $\sim 10^{12}$ M$_\odot$ haloes in gas temperature, entropy, and density with previous cosmological simulations at L8-equivalent resolution and incorporating more realistic feedback physics (van de Voort & Schaye, 2012). In both cases, entropy rises to near the virial radius at which point it flattens and becomes roughly constant. While some haloes exhibit this gradual plateau (h4), the radial entropy profile of others drops

as r^{-2} or faster (h2) at some radius $\geq r_{\rm vir}$. There is a correspondence between the equilibrium state of the halo and its properties beyond the virial radius – the more disturbed systems (h4-7, darker line colours) generally have higher entropy, without any strong transition. The entropy profiles, averaged over all eight haloes, are well converged with respect to L9 and L10. [4]

The specific angular momentum scales roughly as $r^{1/2}$ or even more accurately as $j_{\rm gas} \propto r\, v_c(r)$ given the circular velocity $v_c(r) \propto (M_{\rm enc}(r)/r)^{1/2}$ for an NFW profile of this halo mass. There is broad uniformity among the eight haloes and good convergence in the lower resolution runs, with a small systematic bias towards lower angular momentum content in the halo, L9/L11 $\simeq 0.88$ and L10/L11 $\simeq 0.96$, again driven by high velocity tidal debris becoming better resolved at the highest resolution.

4.4.3 Distributions Beyond Radial Dependence

To avoid missing important features by averaging over multiple gas populations with distinct properties residing at similar radii, we examine the two dimensional distributions of these same gas quantities. Instead of only considering their radial dependence, in Figure 4.8 we show a full pairwise correlation matrix, every quantity plotted with respect to every other. Here we have stacked together all eight haloes at L11, normalizing only the radius of each gas cell with respect to $r_{\rm vir}$ of its parent halo. Despite the different thermal and dynamical structures of the haloes we verify that the key features

[4] In the inner halo, h1 and h2 have sharp drops in both $T_{\rm gas}$ and $S_{\rm gas}$ at a larger radius than typical, $\sim 0.25 r_{\rm vir}$. In these two cases, while the profiles are properly centred on the galaxy, the galaxy is not entirely centred within the halo. This leads to more cold, low entropy gas beyond the typical disc size.

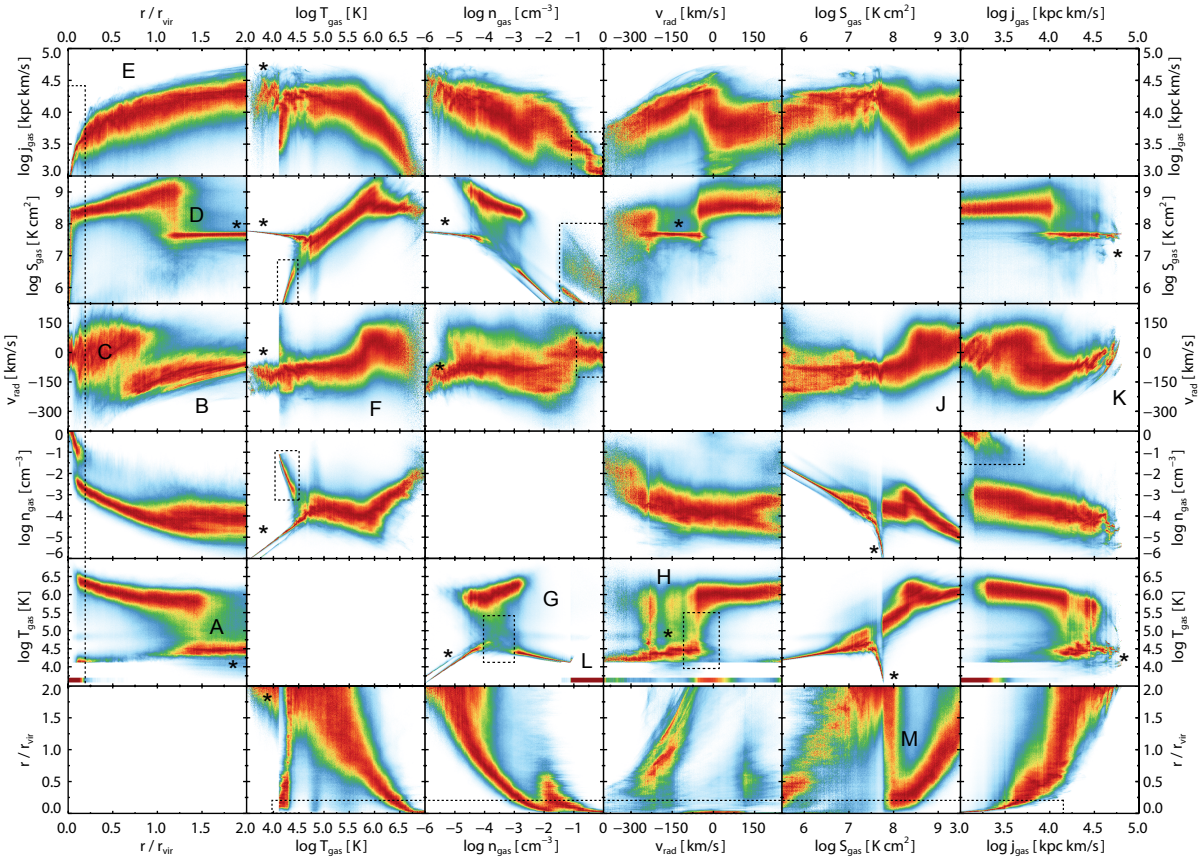

Figure 4.8 The mass-weighted, two dimensional correlation matrix between all pairwise combinations of six quantities: gas radius, temperature, density, radial velocity, entropy, and angular momentum. All eight halos are included and stacked. Each radial bin is normalized independently, such that the colour mapping reaches its maximum intensity at each x-value, independent of the radial distribution of gas mass. Therefore, panels symmetrically across the diagonal from one another are not simply transpositions, but include additional information. Substructures are excised.

seen individually in each are preserved in this collective view. To explore the various relations, we have separately normalized each panel 'column by column'. That is, for each value along the x-axis, the colour mapping of the corresponding one dimensional, vertical slice is independent, such that each slice extends from low values (white/blue) through intermediate values (green/yellow) to high values (orange/red). Therefore, there is no global indication of the distribution of mass within each panel (although very noisy regions indicate a poor sampling and therefore comparably little gas mass). In addition, the corresponding panels across the diagonal are not just transpositions of one another, but answer different questions. For example, the $T_{\rm gas}(r)$ panel in the first column indicates, for each radius, the dominant temperature(s) of gas at that radius, whereas the $r(T_{\rm gas})$ panel in the second column indicates, for each temperature, the dominant radius (or radii) of gas with that temperature. We describe some of the more interesting features seen in this matrix, which have been labelled with letters (A)-(M).

- (*) Denotes features which arise in the intergalactic medium and disappear if gas at $r > r_{\rm vir}$ is excluded. In particular, all the strong constant entropy features at $S_{\rm gas} \simeq 10^{7.7}\,{\rm K\,cm^2}$ in the second row from the top, and the near-discontinuities at this same value in the second column from the right.

- $\square$ Regions demarcated with dotted rectangles indicate features arising in the galactic disc or its vicinity, and disappear if gas at $r < 0.2 r_{\rm vir}$ is excluded. This includes all low j gas at high density, which is already above the threshold for star formation in the galaxy. It also includes low temperature gas which has a radial velocity near zero, as well as the thin horizontal feature of $v_{\rm rad}$ spanning $\sim 0 \pm 100\,{\rm km/s}$ near $r = 0$ corresponding to a rotating disc. Interestingly, the

prominent temperature bridge between the hot halo and galaxy at intermediate densities also largely disappears, indicating that a majority of this cooling occurs at $r < 0.2 r_{\rm vir}$ (see also G).

- (A) The virial shock seen as a sharp temperature jump from $T_{\rm IGM} \sim 10^{4.5}\,{\rm K}$ to $\sim 10^{6}\,{\rm K}$. The width of this transition varies greatly between the eight haloes, h0 being the narrowest at $\Delta r / r_{\rm vir} \simeq 0.1$, and as seen here in the stack broadened to $\Delta r / r_{\rm vir} \simeq 0.5$ or more. The radius where this transition occurs is typically between $1.25\,r_{\rm vir}$ and $1.5\,r_{\rm vir}$. After shocking, the temperature slowly increases with radius, following the mean radial profile of the halo, until it reaches $\sim 0.25 r_{\rm vir}$ where higher densities lead to accelerated radiative cooling, allowing gas to join the cold ISM phase.

- (B) The radial velocity profile of inflowing gas. This component speeds up as it flows in from large distances down to $\sim 0.5 r_{\rm vir}$ at which point it largely disappears. The mean velocity is approximately -75 km/s at $2\,r_{\rm vir}$, -150 km/s at $r_{\rm vir}$, and -225 km/s at $r_{\rm vir}/2$. For free-fall from rest at infinity to a point mass of 10^{12} $M_\odot$ we would expect a speed of -200 km/s at twice the virial radius. The actual value is less due the combination of the Hubble expansion and gas dynamics. Given this offset at the halo outskirts, the subsequent scaling is roughly consistent with $v_{\rm rad} \propto r^{-1/2}$ as expected from free-fall, at least down to $\sim 0.5\,r_{\rm vir}$, below which inflow no longer dominates by mass.

- (C) The radial velocity distribution of the hot halo gas. These two components overlap between $0.5 < r/r_{\rm vir} < 1.0$ as gas transitions from rapidly inflowing to quasi-static. It really is 'quasi', however, since a large mass of gas has positive

radial velocity. Gas with $v_{\rm rad} > 0$ arises from the dynamical formation of the halo atmosphere and associated splashback motion in the baryons (e.g. Wetzel & Nagai, 2014). In this stacked view the radius of this transition is notably interior to the virial shock, $\sim 0.75 r_{\rm vir}$ as compared to $\sim 1.25 r_{\rm vir}$. Investigating the haloes individually we conclude that this is radial offset is largely a misleading feature arising from the stacking of radially averaged velocity profiles. Instead, the radius of a strong jump in temperature and entropy also closely corresponds to a sudden decrease of inward velocity, as a result of the transfer of kinetic to thermal energy.

- (D) The virial shock seen in a sharp entropy jump from the nearly constant value in the IGM of $\sim 10^{7.7}\,{\rm K\,cm^2}$ to $\sim 10^{9}\,{\rm K\,cm^2}$ characteristic of the hot gas for haloes of 10^{12} $M_\odot$. The entropy increase occurs over a narrower radial range than the corresponding temperature increase, and at a slightly smaller radius with respect $r_{\rm vir}$. This may be indicative of pre-shock compressive heating, although we caution that this radial difference varies from halo to halo and as shown in the stack. For example, in our subsequent exploration of h0 in Figure 4.10 we see that there is a close correspondence between the radii of temperature and entropy jumps. After shocking, the entropy slowly declines with radius, following the mean radial profile of the halo.

- (E) The angular momentum distribution is uni-modal and scatters about its mean profile, with no evidence for multiple populations of gas having distinct $j_{\rm gas}$ at any radius.

- (F) The radial velocity depends somewhat on the temperature of the gas, with the hottest gas populating the high positive velocity tail, while cold gas is inflowing.

At all temperatures there is a continuous distribution of radial velocities, with an upturn at $\sim 10^6\,\mathrm{K}$, above which gas has a mean radial velocity consistent with zero, and below which the mean $v_{\rm rad}$ is always negative (see also Joung et al., 2012). The rapidly inflowing hot gas, discussed in (C), is the high velocity tail of the gas distribution at these temperatures and so sub-dominant by mass.

- (G) The usual 'phase-diagram' plotted for cosmological simulations. Cold, low density gas ($n_{\rm gas} \lesssim 10^{-4}\,\mathrm{cm}^{-3}$) in the IGM (denoted by a star) occupies the lower left corner, the tight relation with $T \propto n^{2/3}$ indicative of adiabatic compression. Shock-heated gas at intermediate densities is the only source for $T > 10^{5.5}\,\mathrm{K}$ gas. At higher densities ($n_{\rm gas} \gtrsim 10^{-3}\,\mathrm{cm}^{-3}$) strong cooling flows develop at small radii, after which gas approaches the effective temperature floor of $\sim 10^4\,\mathrm{K}$ until it reaches the star formation threshold at $n_{\rm H} = 0.13\ \mathrm{cm}^{-3}$.

- (H) The heating of inflow, seen here as the dominant temperature for gas at each radial velocity. The handful of strong vertical features are largely an artefact of the stacking. For a given halo, when we restrict this panel to include only gas with $0.2 < r/r_{\rm vir} < 1.0$ we find a clear correlation between deceleration and heating, where gas transitions from cold at -300 km/s to hot by -150 km/s. This is decidedly interior to the virial shock, and outside the galaxy.

- (J) The average radial velocity as a function of entropy, restricted to gas in $0.2 < r/r_{\rm vir} < 1.0$ (not shown) resolves nicely into two distinct populations, where material with $S_{\rm gas} \gtrsim 10^{8.5}\,\mathrm{K\,cm}^2$ has a mean $v_{\rm rad} = 0$ which is roughly constant. On the other hand, gas with lower entropy than this threshold has a mean $v_{\rm rad} \simeq -150\,\mathrm{km/s}$, with a gradual trend towards faster inflow with lower entropy.

- (K) Similarly, the average radial velocity as a function of angular momentum, restricted to gas in $0.2 < r/r_{\rm vir} < 1.0$ resolves nicely into two components. Gas with $j \lesssim 10^{4.25}\,{\rm kpc\,km/s}$ has zero mean radial velocity, whereas higher angular momentum gas has $v_{\rm rad} \simeq -v_{\rm vir}$ (this is not shown explicitly in the figure, which includes all gas). The two populations overlap for $10^{4.25} < j_{\rm gas} < 10^{4.5}$. Together with (J) this panel shows the distinct physical properties of the quasi-static and inflowing halo gas.

- (L) Star forming gas has a temperature on the effective equation of state, which is set to a low, constant value to differentiate it from other gas. This results in the constant temperature band at $\simeq 10^{3.5}\,{\rm K}$ across this entire row. This gas is restricted to small radii, high densities, low radial velocities, and low angular momenta.

- (M) Gas with high entropy, $S \gtrsim 10^{8}\,{\rm K\,cm^2}$ has a well defined relation to radius, following the mean halo profile. Below this threshold, if we exclude material in the IGM, gas at any given entropy is essentially distributed throughout the entire halo.

We have seen in several cases how the mean or median radial profiles fail to fully capture the full state and structure of halo gas. Clear examples are the distributions of temperature, entropy, and radial velocity as a function of radius. On the other hand, the distributions of density and angular momentum as a function of radius are comparatively well described by a median and scatter. By measuring average properties within radial bins we implicitly assume that the structure of halo gas is spherically symmetric. Even Figure 4.8 fails in this regard. For instance, the finite radial thickness of the temperature jump associated with the virial shock could either arise from (i) a spherically symmetric feature of that same thickness, or (ii) the superposition of many

thin shocks spread throughout the same radial range, different shock fronts existing at different radii depending on direction. Motivated by the latter possibility, we proceed to investigate and quantify any variation of the structure of halo gas as a function of angle on the sphere.

4.5 Angular Variability

As a first step away from radial averages, we consider the properties of gas in a broad radial shell encompassing the regime of interaction between quasi-static material and filamentary inflow. In Figure 4.9 we plot the distributions of temperature, density, radial velocity, and entropy, for all gas cells within $0.5r_{\rm vir} < r_{\rm gas} < 1.0r_{\rm vir}$. In particular, we examine how well the lower resolution runs reproduce the distributions of these quantities found in the highest resolution (L11) runs. In this radial range, we find that the gas of all simulated haloes has broadly similar properties, differing only in the details.

The temperature distribution (upper left) shows a broad peak centred roughly at $T_{\rm vir}$, slowly falling off towards lower temperatures and with a distinct low temperature peak at $\simeq 10^{4.25}\,{\rm K}$. While the fractional amount of gas at each temperature in the L9 and L10 runs can differ throughout this regime by up to a factor of two, depending on halo, the mean ratio of both L9/L11 and L10/L11 is consistent with unity from $10^{4.5}\,{\rm K}$ to $10^{6.5}\,{\rm K}$. At the highest temperatures $> 10^{6.5}\,{\rm K}$ the lower resolution runs have much larger deviations with respect to L11, but we attribute this primarily to poor sampling of the extreme tail of the distribution due only to a low number of available gas cells. In contrast, the PDF at the low temperature peak is systematically lower in the lower resolution runs. At $\simeq 10^{4.25}\,{\rm K}$ where there are still a large number of gas cells, the

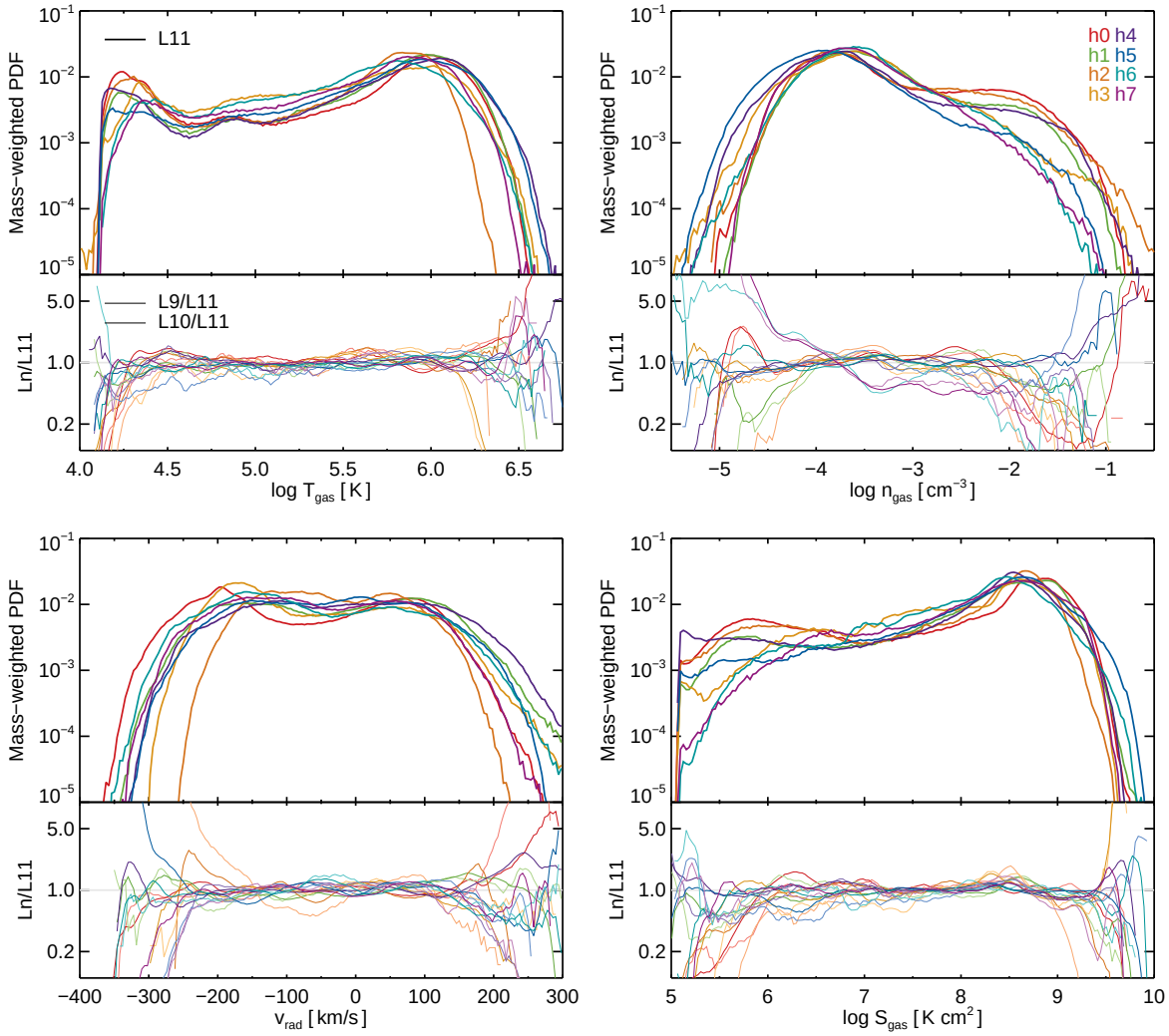

Figure 4.9 The mass-weighted PDF of instantaneous gas temperature, density, radial velocity, and entropy, restricted to the radial range $0.5r_{\rm vir} < r_{\rm gas} < 1.0r_{\rm vir}$. The upper panel shows all eight haloes at the highest resolution level (L11, different colours), while the subpanels below show the ratio of the two lowest resolution levels to the highest, L9/L11 (thin lines) and L10/L11 (medium lines). The vertical axes of the ratio subpanels are logarithmic from 0.1 to 10. For the most part, the eight haloes show similar structure and their lower resolution counterparts scatter about the L11 distributions (see text for details).

mean across all eight haloes is L9/L11 $\simeq 0.5$ and L10/L11 $\simeq 0.8$, indicating that there is a smaller fraction of the total gas mass at these temperatures at lower numerical resolution. Note that 10^4 K is effectively the temperature floor due to cooling in these simulations, and the mean $T_{\rm IGM}$ at redshift two is just above this value.

Gas density shows an even larger variation on a halo to halo basis – some have a strong second peak at higher densities, while this feature is largely absent for the less relaxed haloes. The ratio with respect to the lower resolution runs is consistent with unity for $\log n_{\rm gas} < -2.5\,{\rm cm}^{-3}$. Between this value and the star formation threshold of $\simeq 10^{-1.0}\,{\rm cm}^{-3}$, there is a decrease, at lower resolutions, in the fractional amount of gas at these densities. In this range, the mean ratio across all eight haloes is L9/L11 $\simeq 0.4$ and L10/L11 $\simeq 0.6$. As with temperature, we see that the density PDFs can disagree between resolutions by as much as a factor of ten, but only in the tails when the magnitude drops to low values of $\leq 10^{-4}$. The disagreement is similar with radial velocity, where the mean Ln/L11 ratios are consistent with unity for all values away from the extremes.

The distribution of $v_{\rm rad}$ itself is comprised of two broad components, roughly centred about zero, the positive and negative peaks indicative of outflow and inflow, respectively. The primary driver of inter-halo variation – for example, that h0 has smaller minimum and maximum velocities – appears to be differences in assembly history, particularly a recent merger with another massive halo, or lack thereof. We explore this further in the following section. Finally, gas entropy behaves similarly to temperature, where we find that low entropy gas is strongly sub-dominant by mass with respect to the high entropy hot halo material. The largest variation between haloes occurs for $\log S_{\rm gas} < 10^6\,{\rm K\,cm}^2$, down to the floor at $\simeq 10^5\,{\rm K\,cm}^2$. The halo-average PDF at L9 and L10 reveals a smaller fraction of gas at these lowest entropies when compared to the L11 run.

In general, we conclude that for any given halo, large variations between the three resolution levels can be seen, with deviations up to a factor of ten in the fractional amount of gas at a particular temperature, density, radial velocity, or entropy. These deviations occur mostly in the tails of the distributions which are poorly sampled at the lower resolution levels. Furthermore, the timing differences present in any single comparison likely influence some of the largest outliers. With respect to the average behaviour across all eight haloes, we find that the lower resolution runs have less low temperature, high density, and low entropy gas. By mass fraction with respect to the total gas mass in this radial range, the magnitude of the effect is ~ 2 (~ 1.5) for L9 (L10), although it is unclear if this gas population is necessarily converged for these haloes at L11.

4.5.1 Structure Along Radial Sightlines

To proceed, we consider sightlines originating outwards from the halo centre in different directions, which will have different radial structure. If the hot halo gas was triaxial, for example, we could anticipate that the radius of a strong virial shock would differ along its major and minor axes. We therefore measure the angular variability of the thermal and dynamical structure of the halo by casting many such sightlines from the centre of each halo. Along each of these 'radial rays' we sample the continuous fields of gas temperature, radial velocity, entropy, density, and angular momentum in fixed steps of Δr. Using this ensemble of rays we can then quantify the structure of halo gas without binning in spherically symmetric radial shells.

Ray directions are set using the HEALPIX discretisation of the sphere (Górski

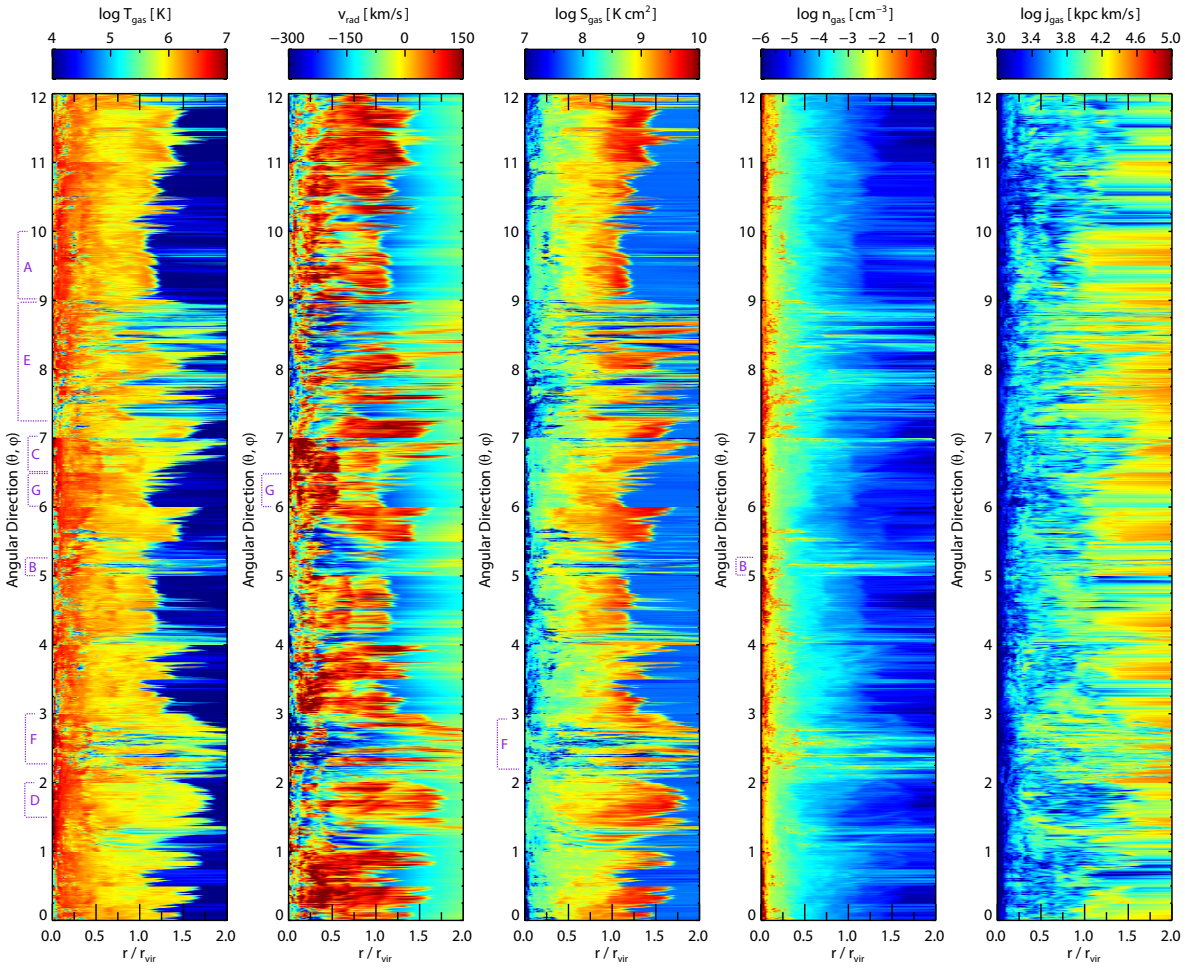

Figure 4.10 The angular variability of gas temperature, radial velocity, entropy, density, and angular momentum at a given radius. Each pixel along the vertical direction represents a single radial ray, which are equally spaced in angular separation. The radial variability of the virial shock is evident in the shifting boundary between dark blue and yellow/orange in the left panel. Penetration of low temperature gas to radii smaller than $r_{\rm vir}$ correlates with higher inflow velocities, lower entropy, and higher density, often extending out into the IGM at $r \geq 2r_{\rm vir}$. In this figure we include only a single halo, h0L11, with $N_{\rm side} = 16$, $N_{\rm rad} = 100$. All substructures other than the primary have been excised, thereby excluding satellites.

et al., 2005) into equal area pixels, which implies equal angular spacing of rays at all refinement levels. The total number of radial rays is $N_{\rm ray} = 12N_{\rm side}^2$, corresponding to an area subtended by each ray equal to $\Omega_{\rm ray} = 4\pi/12N_{\rm side}^2$ sr, or $\theta_{\rm ray} = (180^2/3\pi N_{\rm side}^2)^{1/2}$ deg. The sampling in the radial direction is linear and controlled by the parameter $N_{\rm rad}$ such that $\Delta r/r_{\rm vir} = 2.0/N_{\rm rad}$. Our fiducial parameters of $N_{\rm rad} = 400$, $N_{\rm side} = 64$ result in a sampling of $\Delta r \simeq 0.5\,{\rm kpc}$ and $\Delta\theta = 1\,{\rm deg}$. At each point, mean gas properties are estimated with a tophat kernel with adaptive size equal to the radius of the sphere enclosing the $N_{\rm ngb} = 20$ nearest gas neighbours.[5]

In Figure 4.10 we show a snapshot of the combined radial and angular structure of a single halo (h0L11) at $z = 2$. The twelve demarcated intervals along the y-axis denote the twelve base pixels of the HEALPIX discretisation. Within each base pixel the nested ordering scheme uses a hierarchical quad-tree to preserve adjacency, and the four sub-intervals delineate the top nodes of each such quad-tree, implying that structure seen in the vertical direction is spatially coherent within these major and minor intervals. We note that there is no mass weighting or, for that matter, any indication of the mass distribution within each of the five panels, since the sampling points are smoothly distributed throughout the volume regardless of the underlying gas cell distribution.

Focusing first on the temperature structure we clearly see a boundary separating the cold, intergalactic medium from hot, virialised gas. At this halo mass scale, the temperature increases by roughly two orders of magnitude, from $\simeq 10^4\,{\rm K}$ to $\simeq 10^6\,{\rm K}$

[5]This is close to the mean number of natural neighbours (or cell faces) of the evolved Voronoi mesh in a cosmological simulation (Vogelsberger et al., 2012). Therefore the spatial scale of this effective smoothing is approximately matched to the scale of the stencil used in gradient estimation for linear reconstruction of fluid quantities in the MUSCL-Hancock scheme of AREPO (Springel, 2010a).

(dark blue to yellow/orange). In some directions this heating is coherent and occurs at essentially uniform radius (for example, 9-10, marked 'A'), demarcating a clear 'virialization boundary'. However, across all sightlines we also observe a temperature jump of similar magnitude anywhere from $1.0 < r/r_{\rm vir} < 1.5$, and smaller jumps can occur out to twice the virial radius. Even so, from Figure 4.3 we know that this halo is one of the most spherically symmetric, with a noticeable transition in the state of gas just outside the virial radius. In directions where the gas temperature is warm and exceeds $\sim 10^5\,{\rm K}$ out to twice $r_{\rm vir}$, we find a correspondence to a baryonic overdensity, with respect to mean at that distance. This is indicative of large-scale gas filaments and the heating associated with their earlier collapse.

Interestingly, in directions where gas coherently penetrates to radii smaller than the virial radius and remains cold (e.g. B), this same correspondence to overdensities at larger distance remains. Qualitatively, the existence of an inflowing gas filament arising from the cosmic web suppresses a strong shock at the virialization boundary. This can arise from previous heating from filament formation outside of the halo, which increases the gas temperature to an intermediate state between $T_{\rm IGM}$ and $T_{\rm vir}$. There can nonetheless be a (smaller) temperature jump around the mean virialization boundary (C), although this is not always true and gas can gradually heat seemingly all the way to the peak of the halo temperature profile just exterior to the disc region (e.g. D). Alternatively, this shock suppression can also arise from a delay of strong heating to deeper within the halo, in which case there is less preheating evident at large distances. The radius of the strongest temperature jump then decreases, typically to $0.25r_{\rm vir} < r < 0.75r_{\rm vir}$ as in much of (E). There are essentially no sightlines along which the gas does not experience heating above $T_{\rm IGM}$ at some radius. Since the velocity field,

and so the streamlines of accreting gas, are not purely radial, this does not preclude the possibility that gas could avoid heating along its actual dynamical path. Such a trajectory would be curved in these panels. In some cases (F) we can largely rule out this behaviour, since over a large coherent solid angle no cold gas at $r < 0.5r_{\rm vir}$ is directly connected to cold gas at larger radii. In general, however, the analysis based on radial rays is limited in this respect.

As with density, we also see a strong correlation between temperature and both entropy and radial velocity. Strong temperature jumps just outside the virial radius are also evident as sharp transitions from the gradually increasing (negative) IGM inflow velocity, through equilibrium, to a generally (positive) outflow speed indicative of the virialised halo gas. In some directions (G) multiple components are clearly visible, where an inner halo is comprised of hotter, denser, and more rapidly expanding gas which has more recently bounced back under its own thermal pressure. Similarly, the entropy of the low-density IGM undergoes a sharp increase before declining towards the halo centre. On a sightline by sightline basis the radius of the entropy jump is nearly equal to that of the temperature and radial velocity jumps. Finally, the angular momentum of the gas is clearly the least affected by the virialization boundary. As the velocity field becomes increasingly complex towards smaller radii, the non-radial component contributing to the angular momentum does not have a clear correlation with any of the other local gas properties.

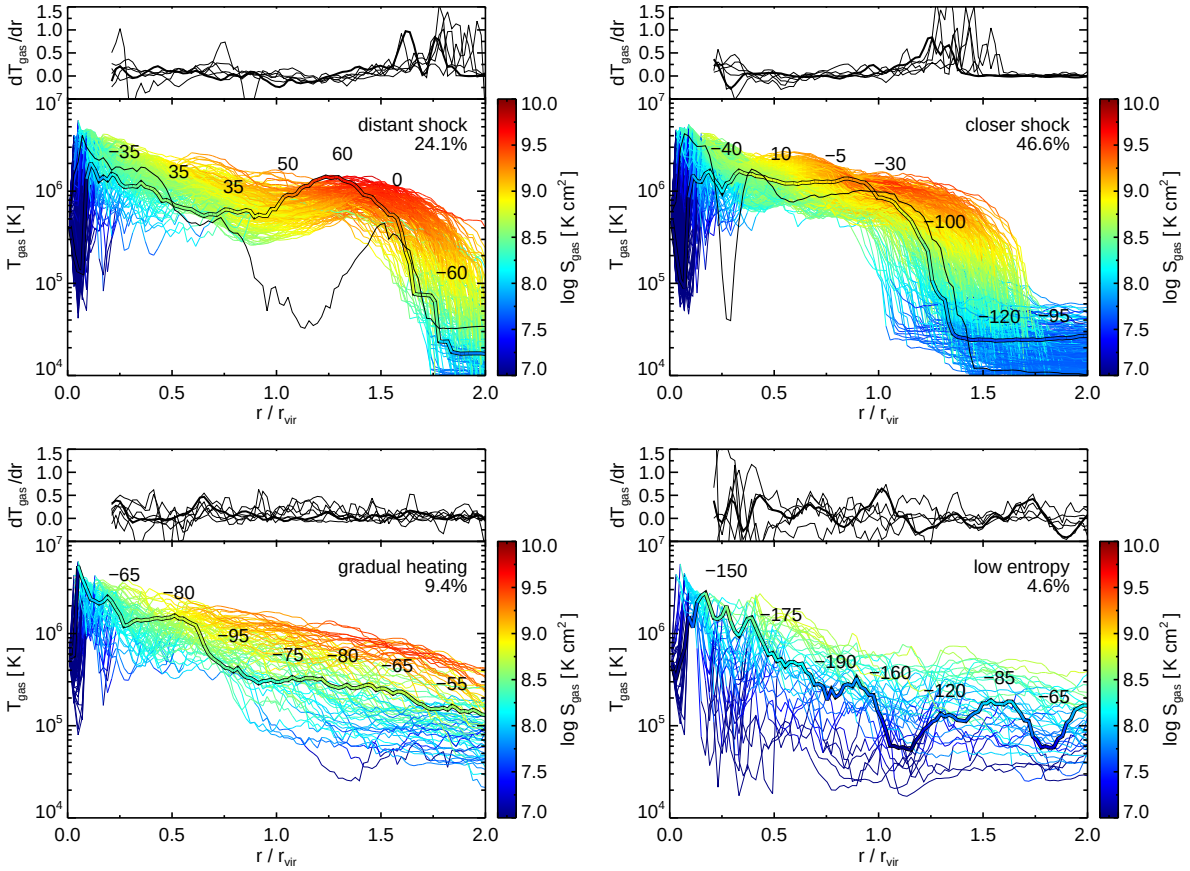

Figure 4.11 The temperature profiles of individual radial rays (four main panels), with colour indicating gas entropy. Each of the four panels includes a disjoint subset of the entire ray set, where the selection for each was chosen by visual inspection in order to find four types of sightlines, each with similar radial properties, which together cover the majority of behaviours. In each panel, all rays are shown as thin lines, while a single prototypical example is shown as a thick coloured line. The mean radial velocity profile for each ray type is indicated by the series of numbers (in units of km/s, negative denoting inflow). The smaller top insets show the derivative of gas temperature with respect to radius, in units of (log K)/(0.1 $r_{\rm vir}$). The percentage in each panel indicates the fraction of all rays satisfying the criteria, for this particular halo. Given our selections, some sightlines are excluded from all four types. The vast majority of these rays either fall just outside a selection, or experience a significant temperature dip within the halo, which is a signature of satellite debris or intersection with non-radial filamentary inflows. Two such examples are shown as thin black lines in the upper two panels. In this figure we include only one halo, h0L9, with $N_{\rm side} = 8$ and $N_{\rm rad} = 100$ for visual clarity.

4.5.2 Different Gas Heating Regimes

To understand the properties and the importance of sightlines exhibiting different radial behaviours, we would like to identify a few characteristic types. We separate rays into a small number of disjoint sets in Figure 4.11, each group ideally having a distinct radial behaviour. By visual classification we identify four such groups, which together encompass the majority of sightlines. In particular, we split rays into 'distant shock', 'closer shock', 'gradual heating', and 'low entropy' types, each of which is based on a quantitative selection applied across some radial range. Specifically, for the 'distant shock' type we require all of

- $\max(\mathrm{d}T/\mathrm{d}r|_{0.2<r<2.0}) > 0.25$
- $1.5 < r_{\max(\mathrm{d}T/\mathrm{d}r)} < 2.0$
- $\min(T_{\mathrm{gas}}|_{0.2<r<1.4}) > 2 \times 10^5\,\mathrm{K}$

where $\mathrm{d}T/\mathrm{d}r$ denotes the derivative of gas temperature with respect to radius, in units of $(\log \mathrm{K})/(0.1\,r_{\mathrm{vir}})$, $r_{\max(Q)}$ indicates the radius where the quantity Q reaches its maximum value, and $Q|_{r_1<r<r_2}$ indicates that a quantity Q is constrained only within the radial range between r_1 and r_2. For the 'closer shock' type we require

- $\max(\mathrm{d}T/\mathrm{d}r|_{0.2<r<2.0}) > 0.25$
- $0.8 < r_{\max(\mathrm{d}T/\mathrm{d}r)} < 1.5$
- $\min(T_{\mathrm{gas}}|_{0.2<r<0.7}) > 3 \times 10^5\,\mathrm{K}$.

That is, both must exhibit a strong temperature jump at either large or intermediate radii, while excluding rays which subsequently drop to low temperature at smaller radii

– an artefact of intersecting non-radial cold debris, as we subsequently discuss. Of the remaining rays not meeting the prior two conditions, we further require rays of a 'gradual heating' type to satisfy

- $\max(\mathrm{d}T/\mathrm{d}r|_{0.2<r<2.0}) < 0.8$
- $\min(T_{\rm gas}|_{0.2<r<0.75}) > 3 \times 10^5\,\mathrm{K}$.

Finally, the 'low entropy' type requires of any remaining rays

- $\max(S_{\rm gas}|_{0.2<r<2.0}) < 8 \times 10^8\,\mathrm{K\,cm^2}$.

In Figure 4.11 we show all four groupings as separate panels (for halo h0L9). The temperature is plotted as a function of radius, while colour indicates gas entropy. A single prototypical ray is included as a thick line, while all rays belonging to that type are shown underneath as thin lines. The mean radial velocity profile of all rays of that type, locally averaged in radius, is denoted by the series of numbers shown in each panel, in units of km/s (negative denoting inflow). The top inset above each panel plots the temperature derivative $\mathrm{d}T/\mathrm{d}r$, positive denoting increasing temperature with decreasing radius, for the prototypical ray (thick) and five other random sightlines of that type (thin).

We briefly describe the behaviour of each type. Sightlines experiencing a 'distant shock' undergo a jump in temperature from $\simeq 10^4\,\mathrm{K}$ to $\simeq 10^6\,\mathrm{K}$ over a radial range of $\sim 0.1 r_{\rm vir}$ (although, as in the example, multiple jumps can exist and be spread over a larger radial range). At this same radius the entropy also increases by approximately two orders of magnitude, from $\simeq 10^{7.5}\,\mathrm{K\,cm^2}$ to $\simeq 10^{9.5}\,\mathrm{K\,cm^2}$. The rapid inflow velocity decreases, reaching its maximum (positive) value near the radius of maximum

Table 4.3: The percentage of radial rays of each of the four types: 'distant shock', 'closer shock', 'gradual heating', and 'low entropy'. The mean fraction across all eight haloes is calculated separately for each resolution level. Since rays cover equal angular area, each fraction corresponds to the geometrical percentage of the sphere occupied by sightlines satisfying each criterion. The errors represent the standard deviation among the eight haloes.

type	L9	L10	L11
distant shock	23.6 ± 4.8	17.3 ± 4.9	12.5 ± 6.1
closer shock	22.5 ± 15.9	27.2 ± 14.0	30.4 ± 15.2
gradual heating	28.0 ± 15.2	14.5 ± 7.9	6.2 ± 3.9
low entropy	7.1 ± 2.7	7.5 ± 3.8	5.4 ± 2.9

temperature. Towards smaller radii, the ray temperature then follows the steadily increasing mean temperature profile of the halo, until reaching the centre. The 'closer shock' sightlines have the same behaviour – the radial distinction between close and distant shocks is arbitrary. In either case, the temperature profile can be non-monotonic if the shock occurs at sufficiently large radius, such that the shock temperature is greater than the local mean $T_{\rm gas}(r)$, and the gas can subsequently cool. Heating can be associated with one or multiple shocks (e.g. two, in the case of the prototypical example shown in the upper left panel). The maximum derivatives of temperature and entropy are $\simeq 2.0$ (in log) per $0.1\,r_{\rm vir}$, although the typical maximum along any given ray is roughly half as large.

The 'gradual heating' sightlines generally reach the same maximum temperature (in the inner halo) and entropy (near the virialization boundary), but do so without any sudden jumps. Here the inflow velocity is roughly constant and always negative, never approaching a quasi-static state, increasing towards smaller radii and with a mean

of approximately -75 km/s. The 'low entropy' rays are selected to have a maximum entropy of less than $8 \times 10^8\,\mathrm{K\,cm^2}$, although the exact value is arbitrary. In this case, the inflow velocity is always strongly negative, peaking in the inner halo, with a mean of approximately -135 km/s. The vast majority of these rays still reach high temperature, but at $\lesssim 0.25 r_{\rm vir}$ and possibly not until the disk-halo interface. At $r > 0.5 r_{\rm vir}$ we see that 'low entropy' sightlines also have systematically lower temperatures and higher inflow velocities than the other three types.

In Table 4.3 we include the fraction of each of these four ray types, calculated as the mean over all eight haloes. Although the balance between the two shock types shifts with resolution, the total ray fraction experiencing a strong shock remains at $\simeq 45\%$, with the temperature jumps moving somewhat inwards. The gradual heating fraction drops sharply at higher resolution levels, either because fluctuations become better resolved, or because the adopted cut on the temperature derivative is slightly too restrictive. The low entropy ray fraction remains fairly converged at $\simeq 6\%$.

We have verified that these values, and all the other quantitative results related to the radial rays, are well converged with the numerical parameters $N_{\rm rad} = 400$, $N_{\rm side} = 64$. Given our selections, approximately $\simeq 15\%$ of the sightlines remain excluded from all four types at L9, although this increases to $\simeq 45\%$ at L11. The vast majority of these excluded rays either (i) fall just outside a selection, or (ii) experience a significant temperature dip within the halo. By definition they do not satisfy the 'gradual heating' or 'low entropy' conditions. Two such examples are shown as thin black lines in the upper two panels. We find that the narrow dips are typically a signature of intersecting satellite debris or satellite outskirts, which were not calculated as locally gravitationally bound and so were not excised. Gas cells of this type fill a larger fraction of the halo

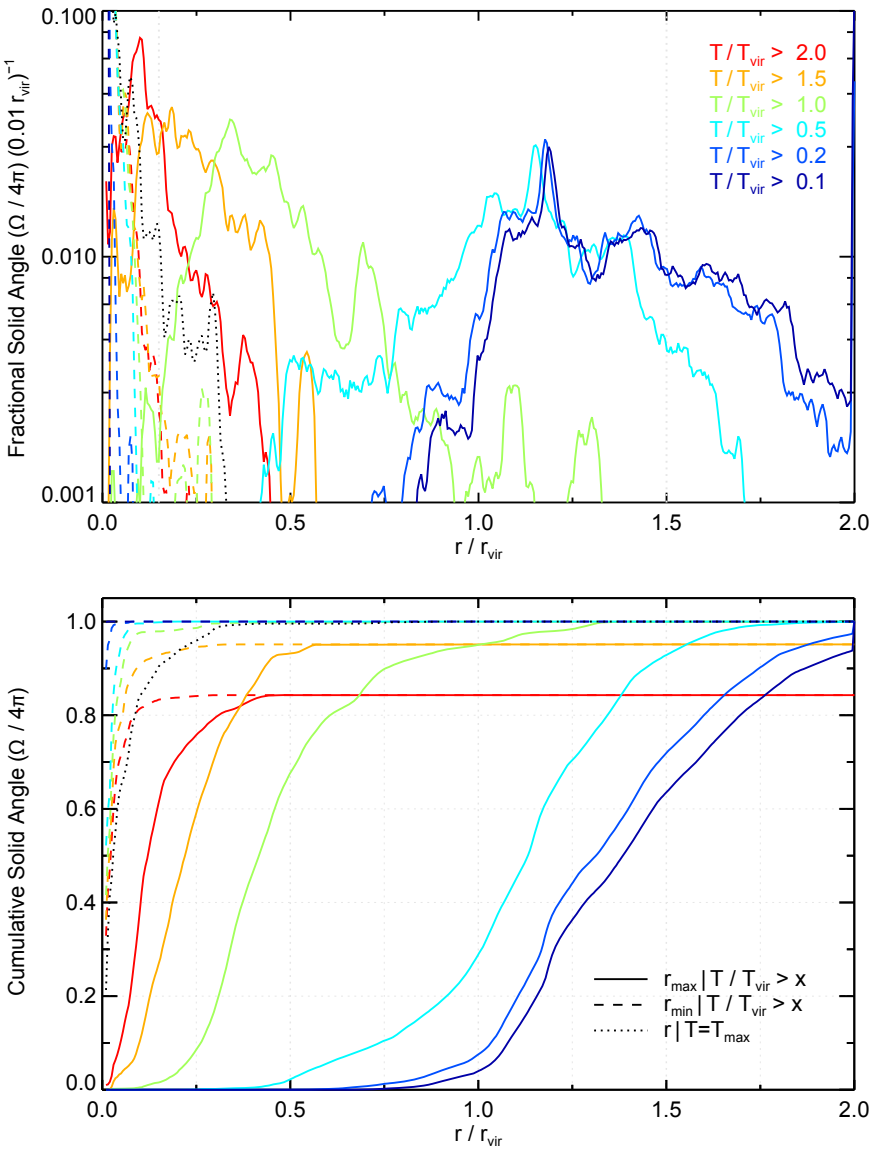

Figure 4.12 (Top) The histogram of radii satisfying three different types of criteria (different linestyles), computed separately for each radial ray. Since each subtends equal angle, this histogram is equal to the fractional solid angle of the whole sphere covered by rays satisfying each criteria, as a (differential) function of radius. The criteria are: (i) the maximum radius at which the temperature along the ray exceeds some fraction of $T_{\rm vir}$, (ii) likewise, but the minimum radius, and (iii) the radius at which each ray reaches its maximum temperature. For the first two conditions, we consider six different temperature thresholds (from blue to red). In this figure we include only one halo at the high resolution level, h0L11, with $N_{\rm side} = 64$ and $N_{\rm rad} = 400$. (Bottom) As in the upper panel, but here as a cumulative function of radius.

volume at higher resolution, leading to the increased percentage of rays excluded from all four types. Broader dips are typically intersections with large filamentary inflows which are not aligned in the radial direction. This exposes the main caveat of the above analysis – namely, that gas inflow with a tangential velocity component need not evolve according to the temperature, entropy, or velocity structure of any particular radial ray.

4.5.3 Quantifying The Asphericity of Temperature

In order to quantify the instantaneous halo temperature structure, we calculate differential histograms of radii satisfying a given criterion in Figure 4.12 (top panel). First, solid lines show the distribution of the maximum radius, for each radial ray, where the temperature exceeds some fraction of $T_{\rm vir}$ (different colours). To avoid washing out features by stacking, we include only one halo (h0L11). We see that for the three lowest temperature cuts, $(0.1 - 0.5) \times T_{\rm vir}$, the $r_{\rm max}$ values all show a strong peak at $\simeq 1.2\,r_{\rm vir}$, indicative of the virialization boundary. The distributions are broad, however. In the case of $T/T_{\rm vir} > 0.5$ the spread of maximum radii is roughly symmetric, extending from $0.5r_{\rm vir}$ out to $1.6r_{\rm vir}$. That is, although the majority of sightlines first exceed $0.5T_{\rm vir}$ at a well-defined surface sitting at $1.2r_{\rm vir}$, a non-negligible fraction first exceed this temperature already by $1.6r_{\rm vir}$, and an even larger fraction cross this threshold inside the virial radius, between $0.5 - 0.75r_{\rm vir}$.

The three higher temperature thresholds (green, orange, red) reflect the increasing mean radial temperature profile of the halo. The virial temperature is typically reached at $\simeq 0.5\,r_{\rm vir}$, while gas reaches twice $T_{\rm vir}$ just prior to the galaxy at $\simeq 0.1\,r_{\rm vir}$. The dashed lines show the minimum radius where each temperature threshold is reached. They are

uniformly peaked at the halo centre. Finally, the dotted line plots the distribution of the radius where each ray reaches its maximum temperature, which are also peaked in the halo centre. One of our principal goals is to identify gas shocking at the halo-IGM transition. We see that the peak of the radial distribution, for rays satisfying $r_{\rm max}(T/T_{\rm vir} > \{0.5, 0.2\})$, measures the most dominant radius for this transition. With respect to halo to halo variability, we find that the shapes and overall widths of each $r_{\rm max}$ distribution can vary substantially. For instance, h0 has a strong peak which is both the narrowest in radius and the most covering in angular fraction. Yet, it also has a prominent secondary peak of roughly half the geometrical importance at larger radius, $\Delta r \simeq 0.2r/r_{\rm vir}$ further from the halo centre. Even more extreme, h3 has two peaks of equal strength spaced $\Delta r \simeq 0.5r/r_{\rm vir}$ apart. Multiple radial peaks are a common occurrence.

The widths of the distributions also encode a measure of the spherical symmetry of any strong heating. Here it is useful to look at the cumulative histograms of radii satisfying these same temperature criteria, as shown in the bottom panel of Figure 4.12. The slope of this CDF indicates angular uniformity – presence of a Heaviside step function would indicate perfect spherical symmetry, while slopes approaching zero would indicate that the temperature transition takes place, depending on sightline, over widely disparate radii. The fact that several temperature thresholds plateau below unity indicate that only $\simeq$85% of rays reach $T > 2.0T_{\rm vir}$ (red line) and $\simeq$95% of rays reach $T > 1.5T_{\rm vir}$ (orange line), while all other temperature thresholds are exceeded along all sightlines. Any value for the cumulative solid angle can provide a good way to quantify the location of the mean virialization boundary. For instance, the radius by which 2π sr are covered by rays with maximum temperature exceeding half of $T_{\rm vir}$ is $\simeq 1.1r_{\rm vir}$.

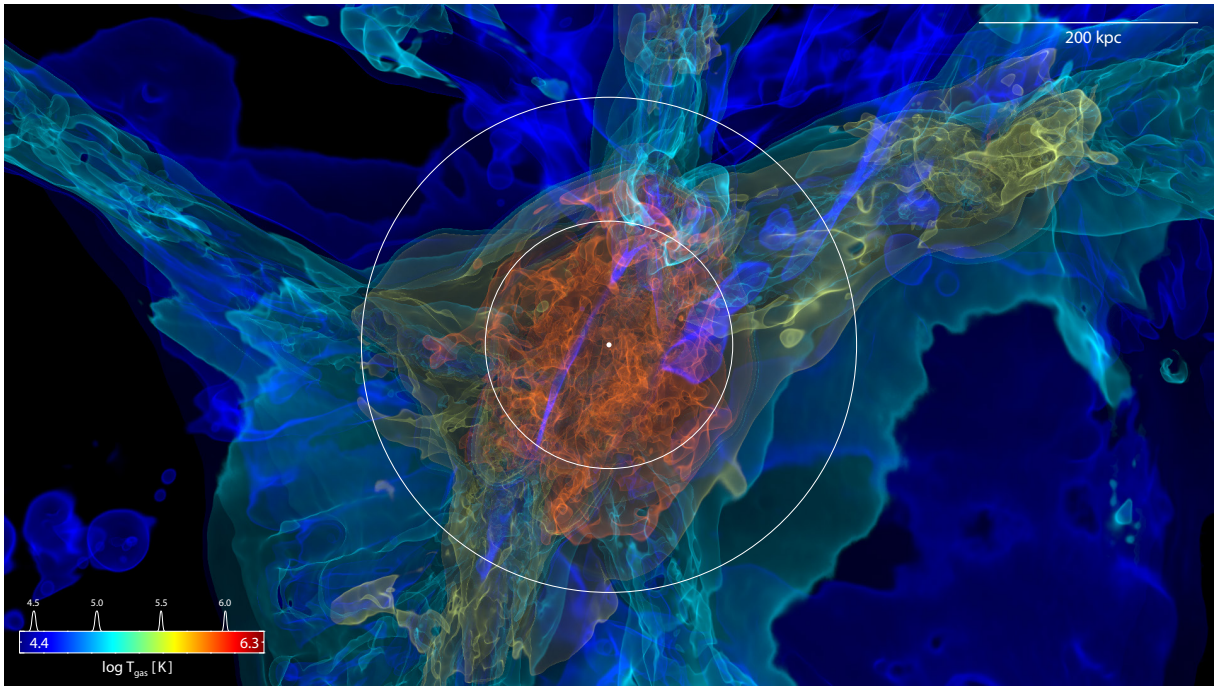

Figure 4.13 A volume rendering of the three-dimensional temperature structure around a single halo (h0L11) at $z = 2$. We use an orthographic projection with depth and height equal to $5.5\,r_{\rm vir}$ ($\simeq 620\,{\rm kpc}$), as in the zoomed out panels of Figure 4.5. The colortable and 4-gaussian transfer function are shown. A front-to-back ray tracing method is used, with no scattering and no absorption. Along each ray we sample with a constant step-size of 0.25 kpc. The temperature at each sample point is calculated using the standard cubic-spline SPH kernel interpolant with adaptive smoothing length h over the $N = 200$ nearest neighbours (the representation of the gas in terms of a Voronoi tessellation is not used). The white dot marks the halo centre, and the white circles denote one and two times the virial radius.

We conclude with a visual impression of the three-dimensional temperature structure in and around a single halo (h0L11) in Figure 4.13. A ray-traced volume rendering highlights five iso-temperature surfaces using the same blue-red colortable and physical bounds of $4.4 < \log T_{\rm gas}\,[{\rm K}] < 6.3$ as in the temperature projections of Figure 4.3. It is sampled with a transfer function comprised of four narrow alpha-channel gaussians at $\log T_{\rm gas}\,[{\rm K}] = \{4.5, 5.0, 5.5, 6.0\}$ with widths of $\sigma = \{6, 6, 3, 3\} \times 10^{-3}$ in log K. The scale of the image is $5.5\,r_{\rm vir}$ in both height and depth, as in the zoomed out panels of Figure 4.5. The dark blue surface straddles the mean $T_{\rm IGM}$ at $\simeq 2r_{\rm vir}$ and reveals a morphology at overdensities just below cosmic web filaments. The light blue surface at $10^5\,{\rm K}$ clearly outlines the structure of Mpc-scale, large-scale filaments, while the yellow surface at $10^{5.5}\,{\rm K}$ is restricted to the interior of warmer filaments. Both of these two iso-temperatures surfaces are closed. Further, they enclose the virialised halo itself – we can see how the filamentary shape of the $10^5\,{\rm K}$ surface at $r > 2r_{\rm vir}$ transitions smoothly into a quasi-spherical shape by opening up, wrapping around and encompassing the higher temperature shells. The mean radius of the $10^{5.5}\,{\rm K}$ surface is $\simeq 1.0\,r_{\rm vir}$, while that of the $10^6\,{\rm K}$ surface is $\simeq 0.5\,r_{\rm vir}$. The intersection of radial sightlines cast outwards in different directions with this complex temperature structure in and around the $10^{12}\ {\rm M}_\odot$ halo gives rise to the four distinct ray types described earlier.

4.6 Discussion

4.6.1 Observational Points

We have deferred a direct comparison against observations of the gas content of haloes to future work. Such a comparison deserves a careful treatment of the various steps required to make robust synthetic (or 'mock') observations from the simulations, for both hydrogen (e.g. Faucher-Giguère et al., 2011; Faucher-Giguere et al., 2014; Fumagalli et al., 2014; Bird et al., 2014) as well as metal signatures (e.g. Shen et al., 2013; Hummels et al., 2013; Ford et al., 2014; Suresh et al., 2015). As motivation and in preparation for the next work in this series, there are a number of insightful observations of the gas in and around haloes with masses of $\sim 10^{12}$ $M_\odot$ at $z \sim 2$ which prompt mention here.

Observations have studied the incidence and kinematics of HI absorption around haloes of this mass (e.g. the Lyman-break galaxies of Rudie et al., 2012, 2013). They found, for instance, a covering fraction for Lyman-limit systems within the virial radius of $\sim 30\%$. The hydrogen and metal content around lower mass damped Lyman-alpha systems (DLAs, $\sim 10^{11}$ $M_\odot$ haloes) at this redshift is roughly consistent with their 10^{12} $M_\odot$ counterparts (Rubin et al., 2014), where the covering fractions within 100 kpc of strong Si II absorption was found to be $\sim 20\%$, versus $\sim 60\%$ for strong C IV.

Around more massive 'quasar hosts' of $\sim 10^{12.5}$ $M_\odot$ the covering fractions of neutral hydrogen are larger (Prochaska et al., 2013). On the scale of the virial radius in such haloes the covering fractions of cold metal ions (e.g. C II and C IV in Prochaska et al., 2014) are even higher and may approach unity. The presence of cold gas at large radii

(Turner et al., 2014b) is a result which, in general, current hydrodynamical simulations have significant difficulty reproducing. On the other hand, hot gas as often probed by O VI at these redshifts (Turner et al., 2014a) is more readily reproduced given sufficiently energetic outflows (Ford et al., 2013; Suresh et al., 2015). In some cases there is strong evidence that the observed absorption results from high velocity galactic outflows (Crighton et al., 2015). Other observations imply that slower inflowing structures with a high degree of kinematic coherence are more likely (Rubin et al., 2014). In general, it is unclear if the cold gas seen in the haloes surrounding galaxies is *always* the consequence of one of (i) feedback-driven outflows (e.g. Marinacci et al., 2014b) or (ii) cosmological inflow, or if this depends on the halo mass under investigation. Either way, it is currently difficult, from the theoretical perspective, to explain how this cold gas is generated, distributed, and either maintained or replenished in the halo.

4.6.2 The Resolution Issue

Observations have also placed constraints on the physical size of the systems which give rise to metal absorption in gaseous haloes. Simcoe et al. (2006) estimated absorbers sizes around $z \sim 2.3$ galaxies ranging from the sub-parsec to the kiloparsec. This is consistent with the transverse scales of C IV absorbers derived in Rauch et al. (2001) from multi-sightline analysis of lensed QSOs at sub-kpc scales (see also Petitjean et al., 2000, based on Mg II). Based on photo-ionization modelling of C IV absorbers at similar redshifts, Schaye et al. (2007) derived typical densities of $n_{\rm H} \sim 10^{-3.5}\,{\rm cm}^{-3}$ with sizes of $\sim 100\,{\rm pc}$. Similarly, Crighton et al. (2015) observe metal absorption lines arising near a $z \simeq 2.5$ galaxy and derive a $(< 100 - 500)\,{\rm pc}$ size constraint, which they argue is a

ubiquitous size-scale for all low-ion halo gas in the circumgalactic environments of both low and high redshift galaxies. Pieri et al. (2014) conclude that a typical low-ionization absorber associated with the CGM of strong Lyman-alpha forest systems at $z \sim 2.5$ is $\sim 30\,\mathrm{pc}$ in size. Furthermore, the gas giving rise to various absorption lines may also have unexpected geometries with little resemblance to spherical clouds (Churchill et al., 2014).

Obtaining $\sim 100\,\mathrm{pc}$ hydrodynamic resolution at the virial radius by simply running higher resolution simulations of isolated haloes will be prohibitively expensive. Following the observed scaling of the gas resolution in the halo, we would obtain a mean $r_{\mathrm{cell}} \simeq 0.1\,\mathrm{kpc}$ physical at $z = 2$ with a L14-class simulation. That is, gas mass resolution approaching 30 $\mathrm{M}_{\odot}$, requiring $\gtrsim 10^{10}$ gas cells in a 10^{12} $\mathrm{M}_{\odot}$ halo. This is not a realistic goal for the near future. Even given such a simulation, the minimum size of resolved clouds would be a few times larger than the size of individual resolution elements.

On the other hand, numerical methods may enable novel ways to focus resolution within the halo regime. If we can avoid clustering the computational effort in the densest regions of space – within galaxies themselves – much higher spatial resolutions can be achieved in lower density media. In AMR simulations, the refinement criterion can be chosen as desired, on the gradient of the neutral hydrogen fraction for instance (Rosdahl & Blaizot, 2012), or with a progressively refined uniform Eulerian grid which is volume filling (Miniati, 2014). In our moving mesh simulations, we can enforce a target gas cell mass criterion of any type through adaptive refinement and de-refinement of Voronoi cells. In future work we will explore this possibility as a solution to the resolution issue, in addition to quantifying the rates and modes of accretion using the tracer particle information.

4.7 Conclusions

In this first paper of the series we present a suite of high-resolution cosmological hydrodynamic zoom-in simulations targeted at understanding the properties of halo gas and cosmological accretion in $\sim 10^{12}$ $M_\odot$ haloes at $z = 2$. Using the moving-mesh code AREPO we simulate each of eight haloes at three levels of increasing resolution, reaching a mean baryonic mass of ∼10,000 solar masses. We study the thermal and dynamical state of halo gas within $2\,r_{\rm vir}$ and in particular quantify density, temperature, entropy, angular momentum, and radial velocity in terms of their mean radial profiles and, more interestingly, in terms of their variation along radial sightlines in different directions.

In this mass regime, haloes typically reside at the intersection of one or more 'cosmic web' filaments arising from large-scale structure. The result is a significant amount of filamentary inflow across the virial radius.

- Investigating the interaction of this inflow and the quasi-static hot halo atmosphere, we find that stream morphologies become continuously more complex with better numerical resolution. In general, single coherent flows tend to resolve into multiple, narrower streams while producing density and temperature structure at smaller spatial scales. However, we point out that even at our highest resolution – comparable to the best zoom-in simulations that currently exist for haloes of this mass – the gas-dynamics in the circumgalactic regime are poorly resolved in comparison to within galaxies themselves.

- In general, the mean radial profiles of gas properties are well captured at our lowest resolution – equivalent to that currently available in full cosmological volume

simulations. However, gas which is stripped from infalling satellites, or which forms as the result of a strong tidal interaction, is less well converged, and populates the halo with additional high density gas in our highest resolution runs.

- Although average radial profiles are well converged, their direct interpretation can be misleading. Within the halo itself, $0.2 < r/r_{\rm vir} < 1.0$, there are clearly multiple gas components overlapping in radius and corresponding to quasi-static versus inflowing material. For example, although the mean radial velocity within the halo is near zero (equilibrium), this is not true for the majority of gas, which is actually either inflowing or outflowing.

- Examining the thermal and dynamical state of gas within $2\,r_{\rm vir}$, we clearly identify the existence of a strong virial shock. This 'virialization boundary' typically resides at $(1.25-1.5)\times r_{\rm vir}$ and is evident as a sharp increase in gas temperature, entropy, density, and a decrease in inwards radial velocity. Collimated inflows which remain cold past the virial radius experience significant heating, typically at $\sim 0.5 r_{\rm vir}$, from $\lesssim 10^{4.5}\,{\rm K}$ to $\gtrsim 10^{6}\,{\rm K}$.

- Although the mean radius of a strong virial shock may be $\simeq 1.25\,r_{\rm vir}$, our sightline analysis allows us to identify, in each halo, many radially-sharp virial shocks, each at some radius between $1.0\,r_{\rm vir}$ and $1.5\,r_{\rm vir}$, depending on angle on the virial sphere.

- Investigating the process of gas virialization, we identify different mechanisms responsible for the heating of gas in the circumgalactic regime. In addition to a single strong virial shock at $\lesssim r_{\rm vir}$, we find that gas can shock at much larger radii, $\lesssim 2\,r_{\rm vir}$, particularly in systems where the hot halo is far from equilibrium and can extend to these distances due to the dynamical response of a recent major merger.

We also see that gas can heat gradually, with temperature increasing slowly from its characteristic IGM value until reaching the mean halo temperature profile. Finally, we identify the existence of radial sightlines along which gas entropy remains always below the level characteristic of the hot halo. Even in this case, however, gas cannot avoid heating to $\sim T_{\rm vir}$ in the inner halo, $\sim 0.2\,r_{\rm vir}$, just prior to cooling onto the centrally forming galaxy. Both of these two cases – gradual heating and persistent low entropy – are sub-dominant, accounting geometrically for only $\sim 15\%$ of the total 4π, with a strong shock covering the other $\sim 85\%$.

- Finally, we conclude by assessing the ashpericity of halo gas and by measuring the location and width of the virialization boundary as demarcated by temperature. We find that the distributions of maximum radius $r_{\rm max}$, where the temperature exceeds e.g. $(0.1 - 0.5) \times T_{\rm vir}$ along each radial sightline, can be quite broad. In general, for the less disturbed systems, the majority of radial sightlines first surpasses $0.5\,T_{\rm vir}$ at a well-defined surface sitting at or just beyond $r_{\rm vir}$. However, a non-negligible number already exceeds this temperature at larger distance, and an even larger fraction crosses this threshold only inside the virial radius, between $0.5 - 0.75 r_{\rm vir}$. We propose that this geometrically-motivated analysis can be used to quantify the structure of gas heating as a result of the halo-IGM transition.

In order to serve as a benchmark for the realization of these eight haloes, we have intentionally not included existing models for energetic feedback due to star formation (in the form of a kinetic galactic-scale wind generation) and AGN (in the form of quasar and radio mode thermal energy input plus local radiative effects). We have deferred any direct comparisons with observations of the gas content of haloes. Both points remain

directions for future work, particularly in the sense of how simulations *without* galactic winds can reproduce observed hydrogen and metal signatures. The answer may depend sensitively on details of the numerics, including hydrodynamic resolution, particularly if the spatial scale of dense gas structures in galactic haloes is as small as commonly estimated.

Chapter 5

The Illustris Simulation: Public Data Release

D. Nelson, A. Pillepich, S. Genel, M. Vogelsberger, V. Springel, P. Torrey, V. Rodriguez-Gomez, D. Sijacki, G. Snyder, B. Griffen, F. Marinacci, L. Blecha, L. Sales, D. Xu, L. Hernquist,
submitted to *Astronomy and Computing*, 2015

Abstract

We present the full public release of all data from the Illustris simulation project. Illustris is a suite of large volume, cosmological hydrodynamical simulations run with the moving-mesh code Arepo and including a comprehensive set of physical models critical for following the formation and evolution of galaxies across cosmic time. Each

simulates a volume of $(106.5\ \mathrm{Mpc})^3$ and self-consistently evolves five different types of resolution elements from a starting redshift of $z = 127$ to the present day, $z = 0$. These components are: dark matter particles, gas cells, passive gas tracers, stars and stellar wind particles, and supermassive black holes. This data release includes the snapshots at all 136 available redshifts, halo and subhalo catalogs at each snapshot, and two distinct merger trees. Six primary realizations of the Illustris volume are released, including the flagship Illustris-1 run. These include three resolution levels with the fiducial "full" baryonic physics model, and a dark matter only analog for each. In addition, we provide four distinct, high time resolution, smaller volume "subboxes". The total data volume is ∼265 TB, including ∼800 full volume snapshots and ∼30,000 subbox snapshots. This paper describes the released data products as well as tools we have developed for their analysis. All data may be directly downloaded in its native HDF5 format. Additionally, we release a comprehensive, web-based API which allows programmatic access to search and data processing tasks. In both cases we provide example scripts and a getting-started guide in several languages: currently, IDL, Python, and Matlab. Finally, this paper addresses scientific issues relevant for the interpretation of the simulations, serves as a pointer to published and on-line documentation of the project, describes planned future additional data releases, and comments on technical aspects of the release.

5.1 Introduction

Our theoretical understanding of the origin and evolution of cosmic structure throughout the universe is increasingly propelled forward by large, numerical simulations. From humble beginnings (e.g. Press & Schechter, 1974; Davis et al., 1985), dark matter only

N-body simulations of pure gravitational dynamics have reached a state of maturity and extreme scale (e.g. Kim et al., 2011; Skillman et al., 2014). They form a foundation in our understanding of the ΛCDM cosmological model, including the nature of both dark matter and dark energy. Yet, such DM-only simulations have a fundamental limitation – they cannot provide any direct predictions for baryonic components of the universe: gas, stars, and black holes. While dark matter halo collapse forms the back bone of structure formation, the majority of observational astronomy is based on the properties of the baryons.

The natural successor to dark matter only N-body simulations are cosmological hydrodynamical simulations (e.g. Katz et al., 1992), which model the coupled evolution of dark matter and cosmic gas. Hydrodynamical simulations can also account for diverse phenomena such as the formation of stars, the growth of supermassive black holes, the energetic feedback processes arising from both populations, the production and distribution of heavy elements, and so forth. Modern efforts are now able to capture cosmological scales of $\gtrsim$100 Mpc, while simultaneously resolving the internal structure of individual galaxies at $\lesssim$1 kpc scales (Horizon-AGN: Dubois et al. (2014), MassiveBlack-II: Khandai et al. (2014), Illustris: Vogelsberger et al. (2014b), EAGLE: Schaye et al. (2015)). These simulations yield verifiable predictions or models for a wide range of interesting astrophysical problems including the spin alignment of galaxies on large scales (e.g. Hahn et al., 2010), the distribution of neutral hydrogen (e.g. Bird et al., 2014; Rahmati et al., 2015), or the impact of baryons on the structure of dark matter haloes (e.g. Schaller et al., 2014).

Observational data focused on the large-scale structure of the universe and the properties of galaxies across cosmic time also continue to increase. Surveys such as

SDSS (York et al., 2000), DEEP2 (Davis et al., 2003), CANDELS (Grogin et al., 2011), and 3D-HST (Brammer et al., 2012) provide local and high redshift measurements of the statistical properties of galaxy populations. Future instruments such as LSST (LSST Science Collaboration et al., 2009) and surveys such as DES (The Dark Energy Survey Collaboration, 2005) will provide increasingly precise observational constraints for theoretical models.

To confront theory and observation, the public dissemination of data from both sides is crucial. Efforts based on the availability of ubiquitous international networks began with the highly successful SDSS SkyServer (Szalay et al., 2000, 2002a), which addressed the problems of how remote users could mine data from large datasets (Gray et al., 2002; Szalay et al., 2002b). The approach, which continues to this day, is based on user written SQL queries executed against a large relational database system – query responses can be thought of as both search results and data extraction. Simple queries with near-instantaneous return, as well as long, queued job queries with results saved into temporary storage are supported.

The Millennium simulation (Springel et al., 2005b) public data release was the first large effort from the theoretical side. Modeled on the SDSS approach, the primary data products were described in terms of a "database", which users could search and extract data from using raw SQL queries (Lemson & Virgo Consortium, 2006). The focus is on the halo and subhalo catalogs, their merger trees, and various post-processed galaxy property catalogs computed with semi-analytical models. It has been continually extended with additional simulations, data products, and capabilities. The Millennium-II simulation (Boylan-Kolchin et al., 2009; Guo et al., 2011) was included, and the idea of the "virtual observatory" (VO) was realized with Overzier et al. (2013). These

efforts have occasionally implemented ideas for incorporating theory within the existing VO framework (Lemson & Zuther, 2009; Lemson et al., 2014). More generally, the Theoretical Astrophysical Observatory (TAO Bernyk et al., 2014) was also targeted at providing mock observations of simulated galaxy and galaxy survey data.

Other dark matter only simulations have adopted similar approaches. The Bolshoi and MultiDark simulations (Klypin et al., 2011) were released under a common database (Riebe et al., 2013), now called CosmoSim. The Dark Energy Universe Simulation (DEUS Rasera et al., 2010) data is available online, as are some data from the MICE simulations (Crocce et al., 2010) through the CosmoHub database. In contrast, the MassiveBlack-II (hydrodynamical) simulation (Khandai et al., 2014) made group catalogs available for direct download. Most recently, the Dark Sky simulation has likewise avoided the database and SQL query framework in favor of direct web access to binary data (Skillman et al., 2014).

In releasing the Illustris simulation data, we adopt a similar approach, offering direct online access to all snapshot, group catalog, merger tree, and supplementary data catalog files. In addition, we develop a web-based API which allows users to perform many common tasks without the need to download any full data files. These include searching over the group catalogs, extracting particle data from the snapshots, accessing individual merger trees, and requesting visualization and further data analysis functions. Extensive documentation and programmatic examples (in IDL, Python, and Matlab) are provided.

This paper is intended primarily as a guide for users of the Illustris simulation data. In Section 5.2 we give an overview of the simulations. Section 5.3 describes the

data products, and Section 5.4 discusses methods for data access. In Section 5.5 we present some scientific remarks and cautions, while in Section 5.6 we discuss community considerations including citation. In Section 5.7 we summarize. Appendix A discusses technical aspects of this data release, and Appendix B gives code examples for the API.

5.2 Description of the Simulations

The Illustris Project is a series of hydrodynamical simulations of a $(106.5\ \mathrm{Mpc})^3$ cosmological volume that follow the evolution of dark matter, cosmic gas, stars, and super massive black holes from a starting redshift of $z = 127$ to the present day, $z = 0$. It includes three runs at increasing resolution levels, Illustris-(1,2,3), where Illustris-1 is the flagship, highest-resolution box. Each has been simulated including a fiducial "full" baryonic physics model, as well as a dark-matter only analog, Illustris-(1,2,3)-Dark. Vogelsberger et al. (2014b,a); Genel et al. (2014); Sijacki et al. (2014) have presented the Illustris simulations and their galaxy and black hole populations, both at $z = 0$ as well as at higher redshifts. In what follows, we summarize the most relevant features.

In Table 5.1 we provide an overview of the specifications of the six Illustris runs, including the computational volume, gravitational softening lengths, and masses of the different particle/cell types, which collectively indicate the resolution and dynamic range achieved. To emphasize the variety of galaxy formation and evolution phenomena which can be addressed with the Illustris simulations, in Figure 5.1 we give the approximate number of a selection of interesting astrophysical objects that can be found in the simulated box, from dark-matter dominated halos at $z = 0$ to luminous active galactic nuclei (AGN) at higher redshifts.

Table 5.1: The most important numerical parameters for the six full volume runs. Gravitational softenings for all particle types other than DM are comoving kpc (with value equal to that of the DM) until $z = 1$ after which they are fixed to their $z = 1$ values, such that at $z = 0$ they have half the softening length as the DM. $m_{\rm baryon}$ is the "target gas mass" (i.e. only the mean mass). The number of gas cells equals the $N_{\rm GAS}$ value only in the initial conditions, the number will then drop as stars and black holes form. Moreover, the total number of baryonic particles (gas cells + star particles + wind particles + black holes) is also not conserved since gas cells can be refined/de-refined to keep their mass within a factor of 2 around $m_{\rm baryon}$. In contrast, the total number of tracers and dark matter particles are both conserved for the duration of the simulation. For all six runs, the volume arises from $L_{\rm box} = 75\,{\rm Mpc}/h$.

Run Name	Volume [Mpc^3]	$N_{\rm GAS}$	$N_{\rm TR}$	$N_{\rm DM}$	$\epsilon_{\rm baryon}$ [kpc]	ϵ_{DM} [kpc]	m_{baryon} [M$_\odot$]	m_{DM} [M$_\odot$]
Illustris-1	106.5^3	1820^3	1820^3	1820^3	0.7	1.4	1.6×10^6	6.3×10^6
Illustris-2	106.5^3	910^3	910^3	910^3	1.4	2.8	1.0×10^7	5.0×10^7
Illustris-3	106.5^3	455^3	455^3	455^3	2.8	5.7	8.0×10^8	4.0×10^8
Illustris-1-Dark	106.5^3	0	0	1820^3	-	1.4	-	7.6×10^6
Illustris-2-Dark	106.5^3	0	0	910^3	-	2.8	-	6.0×10^7
Illustris-3-Dark	106.5^3	0	0	455^3	-	5.7	-	4.8×10^8

A series of analyses based on the Illustris suite have already been performed. These include 1) comparisons to observations and studies of the impact of different feedback models on the distribution and content of gas on large scales, within halos and in the circumgalactic regime (Bird et al., 2014, 2015; Nelson et al., 2015; Suresh et al., 2015; Bogdan et al., 2015); 2) characterizations of the properties of galactic stellar halos (Pillepich et al., 2014), of the satellite populations across host masses (Sales et al., 2015), of the star formation histories (Sparre et al., 2015) and of the morphologies and angular-momentum build up of Illustris galaxies (Torrey et al., 2015; Snyder et al., 2015; Genel et al., 2015); 3) applications of shock finder algorithms (Schaal & Springel, 2015);

4) analyses on the formation of massive, compact galaxies at high redshifts (Wellons et al., 2015); 5) quantification of the galaxy merger rates (Rodriguez-Gomez et al., 2015), and 6) applications of post-processing radiative transfer algorithms in the study of cosmic reionization (Bauer et al., 2015).

5.2.1 Physical Models and Numerical Methods

All of the "full physics" Illustris runs contain the following physical components: (1) Primordial and metal-line radiative cooling in the presence of a redshift-dependent, spatially uniform, ionizing UV background field, with self-shielding corrections. (2) Stochastic star formation in dense gas. (3) Pressurization of the ISM due to unresolved supernovae using an effective equation of state model of a two-phase medium. (4) Stellar evolution with the associated mass loss (gas recycling) and chemical enrichment, taking into account SN Ia/II and AGB stars. (5) Galactic-scale outflows with an energy-driven, kinetic wind scheme. (6) Seeding and growth of supermassive black holes. (7) Feedback from AGN in both quasar and radio (bubble) modes, as well as modifications to the cooling curve of nearby gas due to radiation proximity effects. For complete details on the behavior, implementation, parameter selection, and validation of these physical models, see Vogelsberger et al. (2013), which describes the feedback models, and Torrey et al. (2014), which compares the model output with observations from $z = 0$ to $z = 3$.

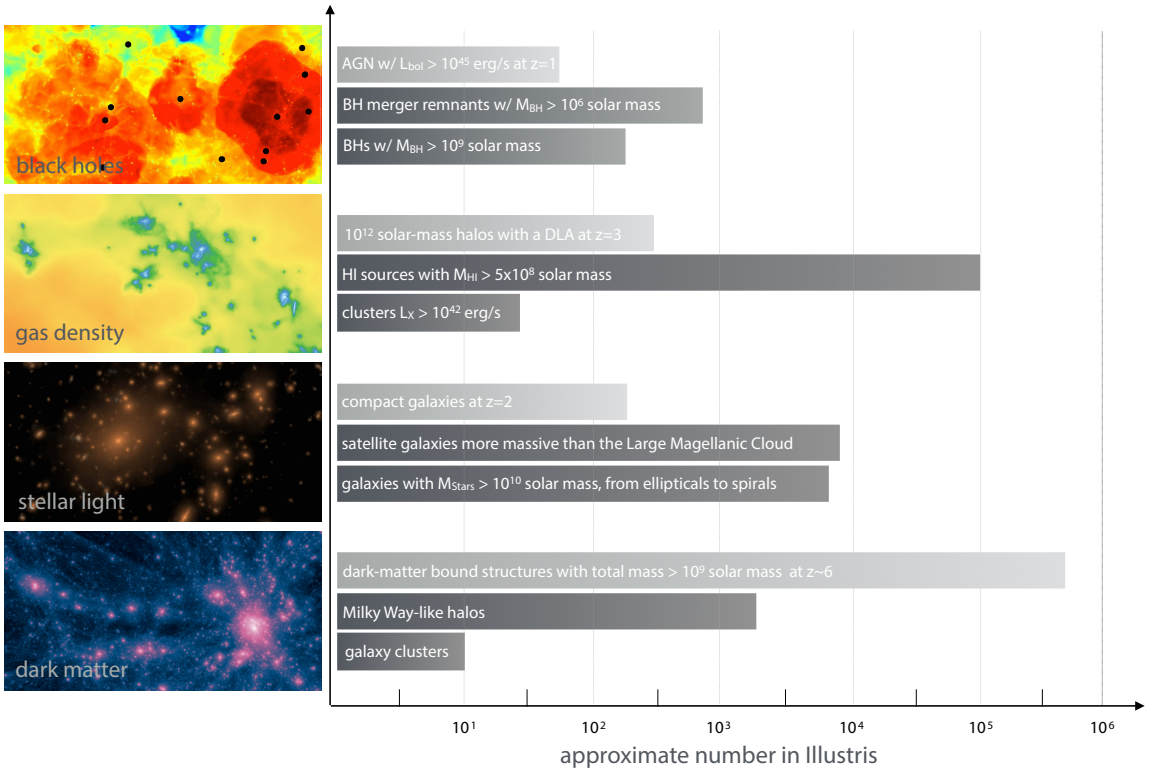

Figure 5.1 Overview of the variety of galaxy phenomena accessible in the Illustris simulations. A few classes of interesting objects are listed for each component: dark matter, stars, gas, and black holes. These are visualized on the left, showing DM density, stellar light, gas density and gas temperature maps, with black holes denoted as black dots. The approximate number present in the Illustris-1 volume is given (from bottom to top), for a) galaxy clusters at $z = 0$ with total mass $M_{200c} > 10^{14}\mathrm{M}_{\odot}$; b) Milky Way-like halos at $z = 0$ ($6 \times 10^{11} < M_{200c} < 2 \times 10^{12}\mathrm{M}_{\odot}$); c) gravitationally-bound objects (dark or luminous) resolved with more than a thousand particles at the end of the reionization epoch; d) galaxies at $z = 0$ with stellar mass exceeding $10^{10}\mathrm{M}_{\odot}$, including both centrals and satellites, from elliptical to disk morphologies; e) satellite galaxies at $z = 0$ more massive than the Large Magellanic Cloud (stellar mass $> 1.5 \times 10^{9}\mathrm{M}_{\odot}$), in any mass host; f) massive, compact galaxies at $z = 2$ according to the selection of Barro et al. (2013); g) clusters of galaxies at $z = 0$ emitting in the X-rays with luminosity exceeding 10^{42} erg/s; h) sources at $z = 0$ with neutral hydrogen mass exceeding $5 \times 10^{8}\mathrm{M}_{\odot}$; i) $10^{12}\mathrm{M}_{\odot}$ halos at $z = 3$ with at least a damped Lyman-alpha system (HI column density $> 10^{20.3}\mathrm{cm}^{-2}$) within 50kpc; j) black holes at $z = 0$ more massive than $10^{9}\mathrm{M}_{\odot}$; k) black-hole merger remnants at $z = 0$, i.e. sub grid black-hole binaries with $M_{\mathrm{BH}} > 10^{6}\mathrm{M}_{\odot}$ for each BH and 1 Gyr delay between the simulation BH merger time and the actual BH merger; l) AGNs at $z = 1$ with bolometric luminosity greater than 10^{45} erg/s.

The Illustris simulations employ the Arepo code (Springel, 2010a) which evolves the equations of continuum hydrodynamics coupled with self-gravity. The spatial discretization of the fluid is provided by an unstructured, moving, Voronoi tessellation. On the volumes defined by individual cells Godunov's method is employed, with a directionally unsplit MUSCL-Hancock scheme and an exact Riemann solver. The Voronoi mesh is generated from a set of control points which move with the local fluid velocity modulo mesh regularization corrections. Gravitational forces are computed using the Tree-PM approach, with long-range forces calculated with a Fourier particle-mesh method, and short-range forces with a hierarchical tree algorithm. The code is second order in space, and with hierarchical adaptive time-stepping, also second order in time. During the simulation we employ the Monte Carlo tracer particle scheme (Genel et al., 2013) to follow the Lagrangian evolution of baryons.

In terms of both physical models and numerical methods, the Illustris simulations rely on a substantial foundation of previous work. In Figure 5.2, we provide an abridged reference tree covering both the physical models and numerical methods. The papers along any given branch are essential for understanding the details and limitations of the data released here.

5.3 Data Products

In this data release we give public access to all 136 snapshots between redshift $z = 40$ and redshift zero of the Illustris cosmological volume. This is a periodic box of 106.5 Mpc per side, including up to five types of resolution elements (dark matter particles, gas cells, gas tracers, stellar and stellar wind particles, and black hole sinks). The same

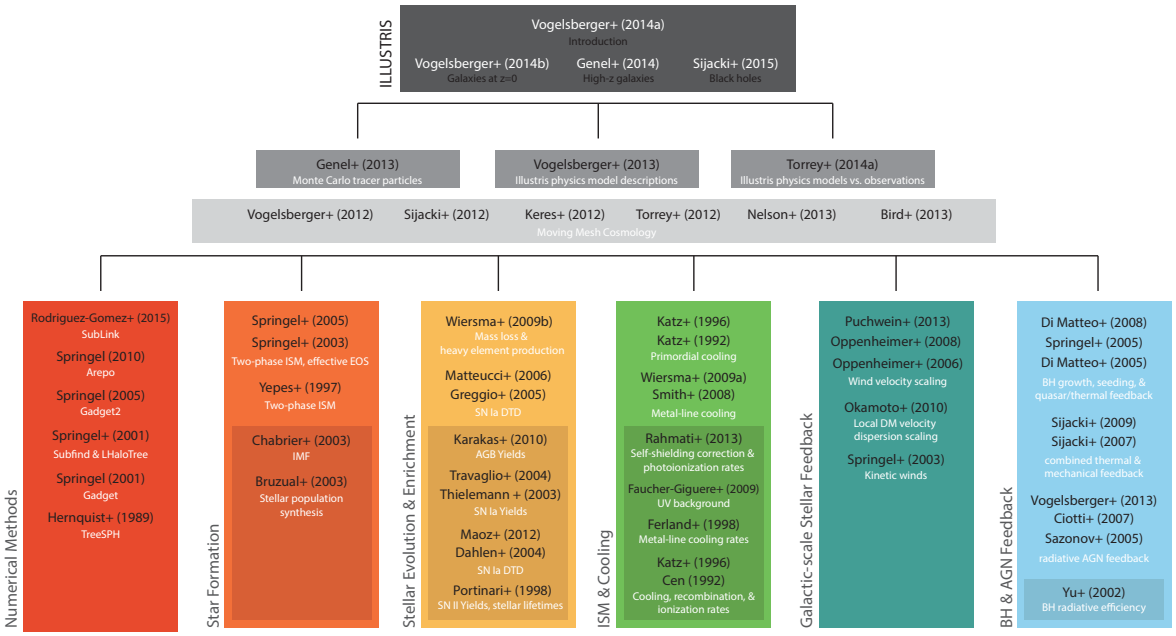

Figure 5.2 Reference tree for the major components of Illustris, including both numerical methods and physical models. Each paper links to its arXiv or ADS entry. We generally include both models and methods which were directly implemented in Illustris, while entries in the dark subboxes indicate model data inputs. The references are, for the second row: Genel et al. (2013); Vogelsberger et al. (2013); Torrey et al. (2014). The moving mesh cosmology series: Vogelsberger et al. (2012); Sijacki et al. (2012); Kereš et al. (2012); Torrey et al. (2012); Nelson et al. (2013); Bird et al. (2013). Numerical methods: Rodriguez-Gomez et al. (2015); Springel (2010a, 2005); Springel et al. (2001a,b); Hernquist & Katz (1989a). Star formation: Springel et al. (2005a); Springel & Hernquist (2003); Yepes et al. (1997); Chabrier (2003); Bruzual & Charlot (2003). Stellar evolution and enrichment: Wiersma et al. (2009b); Matteucci et al. (2006); Greggio (2005); Karakas (2010); Travaglio et al. (2004); Thielemann et al. (2003); Maoz et al. (2012); Dahlen et al. (2004); Portinari et al. (1998). ISM and cooling: Katz et al. (1992, 1996); Wiersma et al. (2009a); Smith et al. (2008); Rahmati et al. (2013); Faucher-Giguère et al. (2009); Ferland et al. (1998); Katz et al. (1996); Cen (1992). Galactic-scale stellar feedback: Puchwein & Springel (2013); Oppenheimer & Davé (2008, 2006); Okamoto et al. (2010); Springel & Hernquist (2003). BH and AGN feedback: Di Matteo et al. (2008); Springel et al. (2005a); Di Matteo et al. (2005); Sijacki et al. (2007, 2009); Vogelsberger et al. (2013); Ciotti & Ostriker (2007); Sazonov et al. (2005); Yu & Tremaine (2002).

volume is available at high (Illustris-1), intermediate (Illustris-2), and low (Illustris-3) resolution. For each resolution, realizations exist with our fiducial, full physics models

("Illustris"), as well as dark matter only analogs ("Illustris Dark"). For all six runs, at every snapshot, two types of group catalogs are provided: friends-of-friends (FoF) halo catalogs, and SUBFIND subhalo catalogs. In postprocessing, these catalogs are used to generate two distinct merger trees, which are both released: SUBLINK, and LHALOTREE. Finally, supplementary data catalogs are released for selected snapshots and runs. At present, these are focused on the stellar properties of Illustris-1 galaxies at $z = 0$, and include mock multi-band images, photometric non-parametric morphological estimates, circularities, angular momenta, and axis ratio measurements. All these data types are described below (snapshots, group catalogs, merger trees, and supplementary catalogs). In the near future we plan to release ROCKSTAR group catalogs and the associated CONSISTENT-TREES merger histories, together with expanded and new supplementary catalogs, with corresponding documentation.

5.3.1 Snapshots

Snapshot Organization

There are 136 snapshots stored for every run. These include all particles/cells in the whole volume. The full snapshot listings, spacings and redshifts can be found online. A partial listing is provided in Table 5.2. Every snapshot is stored in a series of "chunks", i.e. more manageable, smaller-size files. The number of chunks per snapshots is different for the different runs, and is given in Table 5.3.

Note that the snapshot data is **not** organized according to spatial position. Rather, particles within the snapshot files are sorted according to their group/subgroup

Table 5.2: Abridged snapshot list for all six runs. The output times correspond to the set of 128 output redshifts used by the Aquarius project (Springel et al., 2008), augmented by 8 additional saves at integer redshifts.

Snapshot	Scale factor	Redshift
0	0.020932	46.773
32	0.090937	9.9966
45	0.14264	6.0108
49	0.16678	4.9959
54	0.19968	4.0079
60	0.24949	3.0081
68	0.33311	2.002
85	0.50068	0.9973
103	0.66531	0.50305
135	1	0

Table 5.3: Details on the file organization for the six runs. In each case, N_f represents the number of files for each data type, while the provided sizes are the average for that data type. The approximate total data volume for each run is also listed.

Run	Total $N_{\rm DM}$	$N_{\rm f,snap}$	$N_{\rm f,groups}$	Snapshot Size	Groupcat Size	Data Volume
Illustris-3	94,196,375	32	2	22 GB	100 MB	3 TB
Illustris-3-Dark	94,196,375	8	2	3.2 GB	50 MB	0.4 TB
Illustris-2	753,571,000	256	4	176 GB	500 MB	24 TB
Illustris-2-Dark	753,571,000	32	4	26 GB	320 MB	3.5 TB
Illustris-1	6,028,568,000	512	8	1.5 TB	3.6 GB	204 TB
Illustris-1-Dark	6,028,568,000	128	8	203 GB	4 GB	28 TB

memberships, according to the FoF or SUBFIND algorithms. Within each particle type, the sort order is: GroupNumber, SubgroupNumber, BindingEnergy, where particles belonging to the group but not to any of its subgroups ("fuzz") are included after the

last subgroup. Figure 5.3 provides a schematic view of the particle organization within a snapshot, for *one particle type*. Note that the truncation of a snapshot in chunks is arbitrary, thus halos may happen to be stored across multiple, subsequent chunks. Similarly, the different particle types of a halo can be stored in different sets of chunks.

Snapshot Contents

Every HDF5 snapshot contains a "Header" and 5 additional "PartTypeX" groups, for the following particle types (the DM only runs have a single PartType1 group):

- PartType0 - GAS
- PartType1 - DM
- PartType2 - (unused)
- PartType3 - TRACERS
- PartType4 - STARS & WIND PARTICLES
- PartType5 - BLACK HOLES

The most important fields of the header are given in Table 5.4. The complete snapshot field listings, including dimensions, units and descriptions, are given for gas in Table 5.5, dark matter in 5.6, tracers in 5.7, stars in 5.8, and black holes in 5.9.

Table 5.4: Details of the Header group in the snapshot files.

Field	Dimensions	Units	Description
BoxSize	1	ckpc/h	Spatial extent of the periodic box (in comoving units).
MassTable	6	$10^{10}\mathrm{M}_\odot/h$	Masses of particle types which have a constant mass (only DM).
NumPart_ThisFile	6	-	Number of particles (of each type) included in this (sub-)file.
NumPart_Total	6	-	Total number of particles (of each type) across all (sub-)files of this snapshot, modulo 2^{32}.
NumPart_Total_HighWord	6	-	Total number of particles (of each type) across all (sub-)files of this snapshot, divided by 2^{32} and rounded downwards.
Omega0	1	-	The cosmological density parameter for matter.
OmegaLambda	1	-	The cosmological density parameter for the cosmological constant.
Redshift	1	-	The redshift corresponding to the current snapshot.
Time	1	-	The scale factor $a = 1/(1+z)$ corresponding to the current snapshot.
NumFilesPerSnapshot	1	-	Number of file chunks per snapshot.

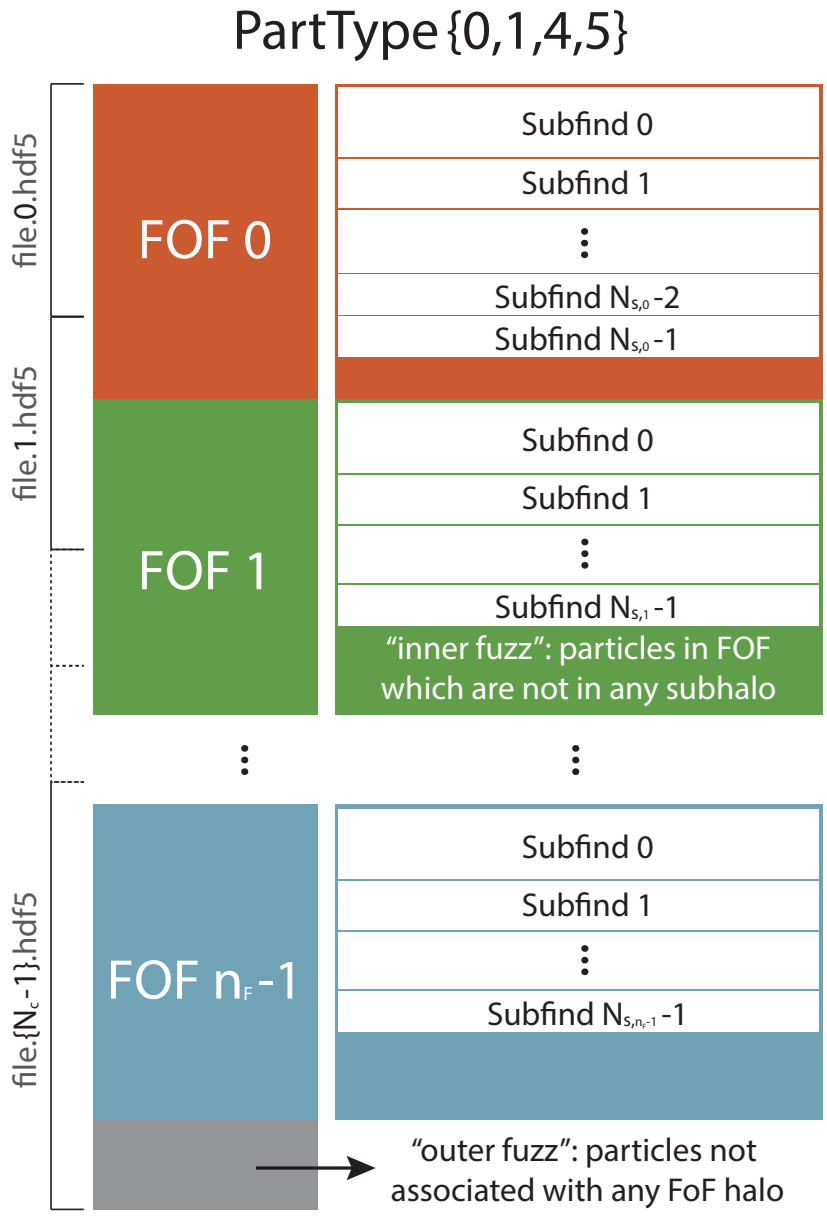

Figure 5.3 Schematic diagram of the organization of particle/cell data within the snapshots for a single particle type. Within a type, particle order is determined by a global sort of the following fields in this order: FoF group number, SUBFIND subhalo number, binding energy, nearest FoF group number. This implies that FOF halos are contiguous, although they can span file chunks. SUBFIND subhalos are only contiguous within a single group, being separated between groups by an "inner fuzz" of all FOF particles not bound to any subhalo. Here N_c indicates the number of file chunks, n_F the number of FOF groups, and $N_{S,j}$ the number of subhalos in j^{th} FoF group.

Table 5.5: Listing of all snapshot fields for gas (PartType0). The following fields are not listed for brevity, and are identical as in PartType4: NumTracers, ParticleIDs, Potential, SubfindDensity, SubfindHsml, SubfindVelDisp, Velocities.

Field	Dimensions	Units	Description
Coordinates	N,3	ckpc/h	Spatial position within the periodic box of size 75000 ckpc/h. Comoving coordinate.
Density	N	$\frac{10^{10}\mathrm{M}_\odot/h}{(\mathrm{ckpc}/h)^3}$	Comoving mass density of cell (calculated as mass/volume).
ElectronAbundance	N	-	Fractional electron number density with respect to the total hydrogen number density, so $n_e = \mathrm{ElectronAbundance} * \mathrm{n_H}$ where $n_H = X_H * \rho/m_p$. Use with caution for star-forming gas (see comment below for NeutralHydrogenAbundance).
GFM_ AGNRadiation	N	erg/s/cm^2	Bolometric intensity (physical units) at the position of this cell arising from the radiation fields of nearby AGN.
GFM_ CoolingRate	N	ergcm$^3/s$	The instantaneous net cooling rate experienced by this gas cell, in cgs units (e.g. $\Lambda_{\mathrm{net}}/n_H^2$).
GFM_ Metallicity	N	-	The ratio M_Z/M_{total} where M_Z is the total mass all metal elements (above He). To convert to solar metallicity, divide by 0.0127 (the primordial solar metallicity).
GFM_ WindDMVelDisp	N	km/s	Equal to SubfindVelDisp.
InternalEnergy	N	(km/s)^2	Internal (thermal) energy per unit mass for this gas cell.
Masses	N	$10^{10}\mathrm{M}_\odot/h$	Gas mass in this cell. Refinement/derefinement attempts to keep this fairly constant.
Neutral Hydrogen Abundance	N	-	Fraction of the hydrogen cell mass (or density) in neutral hydrogen, so $n_{H_0} = \mathrm{NeutralHydrogenAbundance} * \mathrm{n_H}$. (So note that $n_{H^+} = n_H - n_{H_0}$). Use with caution for star-forming gas, as the calculation is based on the 'effective' temperature of the equation of state, which is not a physical temperature.
SmoothingLength	N	ckpc/h	Twice the maximum radius of all Delaunay tetrahedra that have this cell at a vertex in comoving units (s_i from Springel et al. 2010).
StarFormationRate	N	$\mathrm{M}_\odot$/yr	Instantaneous star formation rate of this gas cell.
Volume	N	$1/(\mathrm{ckpc}/h)^3$	Comoving volume of the Voronoi gas cell.

Table 5.6: Listing of all snapshot fields for dark matter (PartType1).

Field	Dimensions	Units	Description
Coordinates	N,3	ckpc/h	Spatial position within the periodic box of size 75000 ckpc/h. Comoving coordinate.
ParticleIDs	N	-	The unique ID (uint64) of this DM particle. Constant for the duration of the simulation.
Potential	N	$(\mathrm{km/s})^2$	Gravitational potential energy.
SubfindDensity	N	$\frac{10^{10}\mathrm{M}_\odot/h}{(\mathrm{ckpc}/h)^3}$	The local total comoving mass density, estimated using the standard cubic-spline SPH kernel over all particles/cells within a radius of SubfindHsml.
SubfindHsml	N	ckpc/h	The comoving radius of the sphere centered on this particle enclosing the 64±1 nearest dark matter particles.
SubfindVelDisp	N	km/s	The 3D velocity dispersion of all dark matter particles within a radius of SubfindHsml.
Velocities	N,3	km$\sqrt{a}$/s	Spatial velocity. The peculiar velocity is obtained by multiplying this value by $\sqrt{a}$.

Table 5.7: Listing of all snapshot fields for tracer particles (PartType3).

Field	Dimensions	Units	Description
FluidQuantities	N,13	Various	Thirteen auxiliary quantities stored for each tracer with differing significance. See Tracer Quantities below.
ParentID	N	-	The unique ID (uint64) of the parent of this tracer. Could be a gas cell, star, wind phase cell, or BH.
TracerID	N	-	The unique ID (uint64) of this tracer. Constant for the duration of the simulation.

Table 5.8: Listing of all snapshot fields for stars (PartType4).

Field	Dimensions	Units	Description
Coordinates	N,3	ckpc/h	Spatial position within the periodic box of size 75000 ckpc/h. Comoving coordinate.
GFM_InitialMass	N	$10^{10}\mathrm{M}_\odot/h$	Mass of this star particle when it was formed (will subsequently decrease due to stellar evolution).
GFM_Metallicity	N	-	See entry under PartType0. Inherited from the gas cell spawning/converted into this star, at the time of birth.
GFM_Stellar FormationTime	N	-	The exact time (given as the scale factor) when this star was formed. **Note: The only differentiation between a real star ($>= 0$) and a wind phase gas cell (< 0) is the sign of this quantity.**
GFM_Stellar Photometrics	N,8	*mag*	Stellar magnitudes in eight bands: U, B, V, K, g, r, i, z. In detail, these are: Buser's X filter (Buser, 1978), where X=U,B3,V (Vega magnitudes), then IR K filter + Palomar 200 IR detectors + atmosphere.57 (Vega), then SDSS Camera X Response Function, airmass = 1.3 (June 2001), where X=g,r,i,z (AB magnitudes). They can be found in the filters.log file in the BC03 package[1]. The details on the four SDSS filters can be found in Stoughton et al. (2002), section 3.2.1.
Masses	N	$10^{10}\mathrm{M}_\odot/h$	Mass of this star or wind phase cell.
NumTracers	N	-	Number of child tracers belonging to this star/wind phase cell.
ParticleIDs	N	-	The unique ID (uint64) of this star/wind cell. Constant for the duration of the simulation.
Potential	N	$(\mathrm{km/s})^2$	Gravitational potential energy.
SubfindDensity	N	$\frac{10^{10}\mathrm{M}_\odot/h}{(\mathrm{ckpc}/h)^3}$	The local total comoving mass density, estimated using the standard cubic-spline SPH kernel over all particles/cells within a radius of SubfindHsml.
SubfindHsml	N	ckpc/h	The comoving radius of the sphere centered on this star particle enclosing the 64±1 nearest dark matter particles.
SubfindVelDisp	N	km/s	The 3D velocity dispersion of all dark matter particles within a radius of SubfindHsml.
Velocities	N,3	$\mathrm{km}\sqrt{a}/\mathrm{s}$	Spatial velocity. The peculiar velocity is obtained by multiplying this value by $\sqrt{a}$.

Table 5.9: Listing of all snapshot fields for black holes (PartType5). The following fields are not listed for brevity, and are identical as in PartType4: NumTracers, ParticleIDs, Potential, SubfindDensity, SubfindHsml, SubfindVelDisp, Velocities.

Field	Dimensions	Units	Description
BH_CumEgy Injection_QM	N	$\frac{10^{10}\mathrm{M}_\odot/h(\mathrm{ckpc}/h)^2}{(0.978\mathrm{Gyr}/h)^2}$	Cumulative amount of thermal AGN feedback energy injected into surrounding gas in the quasar mode.
BH_CumMass Growth_QM	N	$(10^{10}\mathrm{M}_\odot/h)$	Cumulative mass accreted onto the BH in the quasar mode.
BH_Density	N	$\frac{10^{10}\mathrm{M}_\odot/h}{(\mathrm{ckpc}/h)^3}$	Local comoving gas density averaged over the nearest neighbors of the BH.
BH_Hsml	N	ckpc/h	The comoving radius of the sphere enclosing the 64 nearest gas cells around the BH.
BH_Mass	N	$10^{10}\mathrm{M}_\odot/h$	Actual mass of the BH, does not include gas reservoir. Monotonically increases with time according to the accretion prescription, starting from the seed mass.
BH_Mass_bubbles	N	$10^{10}\mathrm{M}_\odot/h$	Accreted mass in current duty cycle for AGN radio mode bubble feedback. When this value reaches a critical fraction of BH_Mass_ini, the bubble energy is released.
BH_Mass_ini	N	$10^{10}\mathrm{M}_\odot/h$	BH mass at the start of the current duty cycle for AGN radio mode feedback, reset after each duty cycle. See BH_Mass_bubbles.
BH_Mdot	N	$\frac{10^{10}\mathrm{M}_\odot/h}{0.978\mathrm{Gyr}/h}$	The mass accretion rate onto the black hole, instantaneous.
BH_Pressure	N	$\frac{10^{10}\mathrm{M}_\odot/h}{(\mathrm{ckpc}/h)(0.978\mathrm{Gyr}/h)^2}$	Reference gas pressure (in comoving units) near the BH, defined as $(\gamma-1)\rho_{sfr}u_{eq}$, where ρ_{sfr} is the star-formation threshold and u_{eq} is BH_U (defined below).
BH_Progs	N	-	Total number of BHs that have merged into this BH.
BH_U	N	$(\mathrm{km/s})^2$	Thermal energy per unit mass in quasar-heated bubbles near the BH, assuming equilibrium between radiative cooling and thermal AGN heating near the BH. Defines BH_Pressure.
Coordinates	N,3	ckpc/h	Spatial position within the periodic box of size 75000 ckpc/h. Comoving coordinate.
HostHaloMass	N	$10^{10}\mathrm{M}_\odot/h$	Mass of FoF group that hosts the BH.
Masses	N	$10^{10}\mathrm{M}_\odot/h$	Total mass of the black hole particle. Includes the gas reservoir from which accretion is tracked onto the actual BH mass (see BH_Mass).

The general unit system is kpc/h for lengths, $10^{10}\mathrm{M}_\odot/h$ for masses, km/s for velocities. The frequently occurring $(10^{10}\mathrm{M}_\odot/h)/(0.978\mathrm{Gyr}/h)$ represents mass-over-time in this unit system, and multiplying by 10.22 converts to $\mathrm{M}_\odot$/yr. Comoving quantities can be converted in the corresponding physical ones by multiplying for the appropriate power of the scale factor a. For instance, to convert a length in physical units it is sufficient to multiply it by a, volumes need a factor a^3, densities a^{-3} and so on. Note that at redshift $z = 0$ the scale factor is $a = 1$, so that the numerical values of comoving quantities are the same as their physical counterparts.

Tracer Quantities

Each Monte Carlo tracer particle stores 13 auxiliary values. These are updated every timestep where the tracer parent is active. Many are reset to zero immediately after they are written out to a snapshot, such that their recording duration is precisely the time interval between two successive snapshots. Some are only relevant when the tracer resides within a parent of a specific particle type (e.g. gas or star). Table 5.10 describes these fields. Also note that tracers are exchanged (and can therefore change their parents) in the following ways:

- Gas -> Gas (finite volume fluxes, refinement, derefinement)
- Gas -> Stars (star formation, spawning new stars and converting cells into stars)
- Stars -> Gas (stellar mass return)
- Gas -> Wind (galactic scale stellar winds)
- Wind -> Gas (recoupling stellar wind)
- Gas -> BHs (black hole accretion)
- BHs -> BHs (black hole mergers)

Table 5.10: Listing of the thirteen auxiliary values stored by the tracer particles. The Reset column indicates whether or not this field is set to zero immediately after each snapshot is written.

Number	Name	Reset?	Units	Description
0	TMax	Y	Kelvin	The maximum past temperature of the parent gas cell, back to the previous snapshot. Only updated when parent is a gas cell.
1	TMax_Time	Y	-	Scale factor of the above TMax event. Only updated when parent is a gas cell.
2	TMax_Time_Rho	Y	$\frac{10^{10}\mathrm{M}_\odot/h}{(\mathrm{ckpc}/h)^3}$	Density of the parent gas cell when the most recent TMax was recorded. Only updated when parent is a gas cell.
3	RhoMax	Y	$\frac{10^{10}\mathrm{M}_\odot/h}{(\mathrm{ckpc}/h)^3}$	Maximum past density of the parent gas cell, back to the previous snapshot. Only updated when parent is a gas cell.
4	RhoMax_Time	Y	-	Scale factor of the above RhoMax event. Only updated when parent is a gas cell.
5	MachMax	Y	-	Maximum past mach number of the parent gas cell, as set in the Riemann solver. Only updated when parent is a gas cell.
6	EntMax	Y	$P/(\rho/a^3)^\gamma$	Maximum past entropy of the parent gas cell, back to the previous snapshot. Only updated when parent is a gas cell. Note slightly strange units, where P and ρ are pressure and density, as in the snapshots.
7	EntMax_Time	Y	-	Scale factor of the above EntMax event. Only updated when parent is a gas cell.
8	Last_Star_Time	N	-	Scale factor, set only when this tracer exchanges from a star/wind to a gas, or from a gas to a star/wind. These four cases respectively set LST = { a, -a, a+1, a+2 }.
9	Wind_Counter	N	int32	Integer counter, the number of moves from a gas cell to a wind particle.
10	Exchange_Counter	N	int32	Integer counter, the number of times this tracer is exchanged.
11	Exchange_Distance	N	ckpc/h	Cumulative sum of the spatial distance over which this tracer has moved due to Monte Carlo exchange between gas cells. In particular, the sum of the parent gas cell radii when either the originating parent or destination parent is of gas type.
12	Exchange_ Distance_Error	N	ckpc/h	Cumulative sum of $r_{\mathrm{cell}} \times (\sqrt{N_{\mathrm{exch}}} - \sqrt{N_{\mathrm{exch}} - 1})$, when either the originating or destination parent is of gas type.

Subboxes

Four separate "subbox" cutouts exist, for each full physics run. These are spatial cutouts of fixed comoving size and fixed comoving coordinates. They are output at each highest timestep, that is, their time resolution is significantly better than that of the main snapshots – see Table 5.11. This may be useful for some types of analysis or particular science questions, or for making movies. Two notes of caution: first, the time spacing of the subboxes is not uniform in scale factor or redshift, but scales with the time integration hierarchy of the simulation, and is thus variable, with some discrete factor of two jumps at several points during the simulations. Second, the subboxes, unlike the full box, are not periodic.

The four subboxes sample four different areas of the large box, roughly described by the environment column in Table 5.12. The particle fields are all identical to the main snapshots. However, the ordering differs. In particular, particles/cells in the subboxes are not ordered according to their group membership, as no group catalogs are available for these cutouts.

5.3.2 Group Catalogs

There is one group catalog associated with each snapshot, which includes both FoF and Subfind objects. The group files are split into a small number of sub-files, just as with the raw snapshots. Every HDF5 group catalog contains snapshot contains the following groups: Header, Group, Subhalo, Offsets. The IDs of the members of each group/subgroup are not stored in the group catalog files. Rather, particles/cells in the

Table 5.11: Details of the subbox snapshots. For each resolution level, from lowest to highest, the total number of subbox snapshots saved $N_{\rm snap}$. Each of the four subboxes has the same number of snapshots. The number of file pieces per snapshot N_c, and the approximate time resolution Δt at three redshifts: $z = 6$, $z = 2$, and $z = 0$.

Run	$N_{\rm snap}$	N_c	$\Delta t_{(z=6)}$	$\Delta t_{(z=2)}$	$\Delta t_{(z=0)}$
Illustris-3	1426	1	∼7 Myr	∼12 Myr	∼33 Myr
Illustris-2	2265	16	∼4 Myr	∼6 Myr	∼17 Myr
Illustris-1	3976	512	∼2 Myr	∼3 Myr	∼8 Myr

Table 5.12: Additional details of the subbox snapshots. For each subbox number, its physical environment, matter overdensity, center position, box size along each coordinate axis, and volume fraction with respect to the full box.

Subbox #	Environment	$\Omega_m^{\rm sub}$	(x_c, y_c, z_c)	$L_{\rm subbox}$	Volume Frac
0	One $\sim 5 \times 10^{13}{\rm M}_\odot$ halo	1.47	(9000, 17000, 63000)	7.5 cMpc/h	0.1%
1	Several $> 10^{12}{\rm M}_\odot$ halos	0.16	(43100, 53600, 60800)	8.0 cMpc/h	0.12%
2	Several $> 10^{12}{\rm M}_\odot$ halos	0.29	(37000, 43500, 67500)	5.0 cMpc/h	0.03%
3	Several $\sim 10^{12}{\rm M}_\odot$ halos	0.25	(64500, 51500, 39500)	5.0 cMpc/h	0.03%

snapshot files are ordered according to group membership. Each group contains its total length, allowing IDs and all other fields of member particles/cells to be accessed using an offset table type approach. This applies to subhalos as well, e.g. the subhalos belonging to group 0 are listed first.

In order to reduce confusion, we adopt the following terminology when referring to different types of objects. "Group", "FoF Group", and "FoF Halo" all refer to halos. "Subgroup", "Subhalo", and "Subfind Group" all refer to subhalos. The first (most massive) subgroup of each halo is the "Primary Subgroup" or "Central Subgroup". All other following subgroups within the same halo are "Secondary Subgroups", or "Satellite Subgroups".

FoF Groups. The Group fields are derived with a standard friends-of-friends (FoF) algorithm with linking length $b = 0.2$. The FoF algorithm is run on the dark matter particles, and the other types (gas, stars, BHs) are attached to the same groups as their nearest DM particle. The fields for the FoF halo catalog are described in Tables 5.13 and 5.14.

Subfind Groups. The Subhalo fields are derived with the SUBFIND algorithm, last described in Springel et al. (2005a). It has since undergone many modifications to add additional properties to each subhalo entry. Descriptions of all fields in this subhalo catalog are split across Tables 5.15, 5.16, and 5.17.

Header and Offsets. Table 5.18 describes the fields in the Header group, while Table 5.19 describes the fields in the Offsets group. Note that we simply store the offsets here, the relate to all types of data files and not solely to the group catalogs.

Table 5.13: Description of all fields in the FoF halo catalogs (Part I). All fields are float32 unless otherwise specified.

Field	Dimensions	Units	Description
GroupBHMass	N	$10^{10}\mathrm{M}_\odot/h$	Sum of the BH_Mass field of all black holes (type 5) in this group.
GroupBHMdot	N	$\frac{10^{10}\mathrm{M}_\odot/h}{(0.978\mathrm{Gyr}/h)}$	Sum of the BH_Mdot field of all black holes (type 5) in this group.
GroupCM	N,3	ckpc/h	Center of mass of the group, computed as the sum of the mass weighted relative coordinates of all particles/cells in the group, of all types. Comoving coordinate. (Available only for the Illustris-3 run)
GroupFirstSub	N	-	Index into the Subhalo table of the first/primary/most massive SUBFIND group within this FoF group (int32).
GroupGasMetallicity	N	-	Mass-weighted average metallicity (Mz/Mtot, where Z = any element above He) of all gas cells in this FOF group.
GroupLen	N	-	Integer counter of the total number of particles/cells of all types in this group (int32).
GroupLenType	N,6	-	Integer counter of the total number of particles/cells, split into the six different types, in this group. Note: Wind phase cells are counted as stars (type 4) for GroupLenType (int32).
GroupMass	N	$10^{10}\mathrm{M}_\odot/h$	Sum of the individual masses of every particle/cell, of all types, in this group.
GroupMassType	N,6	$10^{10}\mathrm{M}_\odot/h$	Sum of the individual masses of every particle/cell, split into the six different types, in this group. Note: Wind phase cells are counted as gas (type 0) for GroupMassType.
GroupNsubs	N	-	Count of the total number of SUBFIND groups within this FoF group (int32).
GroupPos	N,3	ckpc/h	Spatial position within the periodic box of size 75000 ckpc/h of the maximum bound particle. Comoving coordinate.
GroupSFR	N	$\mathrm{M}_\odot$/yr	Sum of the individual star formation rates of all gas cells in this group.
GroupStarMetallicity	N	-	Mass-weighted average metallicity (Mz/Mtot, where Z = any element above He) of all star particles in this FOF group.

Table 5.14: Description of all fields in the FoF halo catalogs (Part II). All fields are float32 unless otherwise specified.

Field	Dimensions	Units	Description
GroupVel	N,3	km/s/a	Velocity of the group, computed as the sum of the mass weighted velocities of all particles/cells in this group, of all types. The peculiar velocity is obtained by multiplying this value by $1/a$.
GroupWindMass	N	$10^{10}\mathrm{M}_\odot/h$	Sum of the individual masses of all wind phase gas cells (type 4, BirthTime <= 0) in this group.
Group_M_Crit200	N	$10^{10}\mathrm{M}_\odot/h$	Total mass of this group enclosed in a sphere whose mean density is 200 times the critical density of the Universe, at the time the halo is considered.
Group_M_Crit500	N	$10^{10}\mathrm{M}_\odot/h$	Likewise, but for 500 times the critical density of the Universe.
Group_M_Mean200	N	$10^{10}\mathrm{M}_\odot/h$	Likewise, but for 200 times the mean density of the Universe.
Group_M_TopHat200	N	$10^{10}\mathrm{M}_\odot/h$	Likewise, but for Δ_c times the critical density of the Universe, where Δ_c derives from the solution of the collapse of a spherical top-hat perturbation (fitting formula from Bryan & Norman (1998)). The subscript 200 can be ignored.
Group_R_Crit200	N	ckpc/h	Comoving radius of a sphere centered at the GroupPos of this Group whose mean density is 200 times the critical density of the Universe, at the time the halo is considered.
Group_R_Crit500	N	ckpc/h	Likewise, but for 500 times the critical density of the Universe.
Group_R_Mean200	N	ckpc/h	Likewise, but for 200 times the mean density of the Universe.
Group_R_TopHat200	N	ckpc/h	Likewise, but for Δ_c times the critical density of the Universe.

Table 5.15: Description of all fields in the SUBFIND subhalo catalogs (Part I). All fields are float32 unless otherwise specified.

Field	Dimensions	Units	Description
SubhaloBHMass	N	$10^{10}\mathrm{M}_\odot/h$	Sum of the masses of all black holes in this subhalo.
SubhaloBHMdot	N	$\frac{10^{10}\mathrm{M}_\odot/h}{0.978\mathrm{Gyr}/h}$	Sum of the instantaneous accretion rates $\dot{M}$ of all black holes in this subhalo.
SubhaloCM	N,3	ckpc/h	Comoving center of mass of the Subhalo, computed as the sum of the mass weighted relative coordinates of all particles/cells in the Subhalo, of all types.
SubhaloGasMetallicity	N	-	Mass-weighted average metallicity (Mz/Mtot, where Z = any element above He) of the gas cells bound to this Subhalo, but restricted to cells within twice the stellar half mass radius.
SubhaloGasMetallicityHalfRad	N	-	Same as SubhaloGasMetallicity, but restricted to cells within the stellar half mass radius.
SubhaloGasMetallicityMaxRad	N	-	Same as SubhaloGasMetallicity, but restricted to cells within the radius of V_{max}.
SubhaloGasMetallicitySfr	N	-	Mass-weighted average metallicity (Mz/Mtot, where Z = any element above He) of the gas cells bound to this Subhalo, but restricted to cells which are star forming.
SubhaloGasMetallicitySfrWeighted	N	-	Same as SubhaloGasMetallicitySfr, but weighted by the cell star-formation rate rather than the cell mass.
SubhaloGrNr	N	-	Index into the Group table of the FOF host/parent of this Subhalo (int32).
SubhaloHalfmassRad	N	ckpc/h	Comoving radius containing half of the total mass (SubhaloMass) of this Subhalo.
SubhaloHalfmassRadType	N,6	ckpc/h	Comoving radius containing half of the mass of this Subhalo split by Type (SubhaloMassType).
SubhaloIDMostbound	N	-	The ID of the particle with the smallest binding energy (could be any type, int64).
SubhaloLen	N	-	Total number of member particle/cells in this Subhalo, of all types (int32).
SubhaloLenType	N,6	-	Total number of member particle/cells in this Subhalo, separated by type (int32).
SubhaloMass	N	$10^{10}\mathrm{M}_\odot/h$	Total mass of all member particle/cells which are bound to this Subhalo, of all types.

Table 5.16: Description of all fields in the SUBFIND subhalo catalogs (Part II). All fields are float32 unless otherwise specified. Note that for all mass calculations by type, wind phase cells are counted as gas.

Field	Dimensions	Units	Description
SubhaloMassInHalfRad	N	$10^{10}\mathrm{M}_\odot/h$	Sum of masses of all particles/cells within the stellar half mass radius.
SubhaloMassInHalfRadType	N,6	$10^{10}\mathrm{M}_\odot/h$	Sum of masses of all particles/cells (split by type) within the stellar half mass radius.
SubhaloMassInMaxRad	N	$10^{10}\mathrm{M}_\odot/h$	Sum of masses of all particles/cells within the radius of V_{max}.
SubhaloMassInMaxRadType	N,6	$10^{10}\mathrm{M}_\odot/h$	Sum of masses of all particles/cells (split by type) within the radius of V_{max}.
SubhaloMassInRad	N	$10^{10}\mathrm{M}_\odot/h$	Sum of masses of all particles/cells within twice the stellar half mass radius.
SubhaloMassInRadType	N,6	$10^{10}\mathrm{M}_\odot/h$	Sum of masses of all particles/cells (split by type) within twice the stellar half mass radius.
SubhaloMassType	N,6	$10^{10}\mathrm{M}_\odot/h$	Total mass of all member particle/cells which are bound to this Subhalo, separated by type.
SubhaloParent	N	-	Index into the Subhalo table of the unique SUBFIND parent of this Subhalo (int32).
SubhaloPos	N,3	ckpc/h	Spatial position within the periodic box of size 75000 ckpc/h of the maximum bound particle. Comoving coordinate.
SubhaloSFR	N	$\mathrm{M}_\odot$/yr	Sum of the individual star formation rates of all gas cells in this subhalo.
SubhaloSFRinHalfRad	N	$\mathrm{M}_\odot$/yr	Same as SubhaloSFR, but restricted to cells within the stellar half mass radius.
SubhaloSFRinMaxRad	N	$\mathrm{M}_\odot$/yr	Same as SubhaloSFR, but restricted to cells within the radius of V_{max}.
SubhaloSFRinRad	N	$\mathrm{M}_\odot$/yr	Same as SubhaloSFR, but restricted to cells within twice the stellar half mass radius.
SubhaloSpin	N,3	(kpc/h)(km/s)	Total spin per axis, computed for each as the mass weighted sum of the relative coordinate times relative velocity of all member particles/cells.

Table 5.17: Description of all fields in the SUBFIND subhalo catalogs (Part III). All fields are float32 unless specified.

Field	Dimensions	Units	Description
SubhaloStarMetallicity	N	-	Mass-weighted average metallicity (Mz/Mtot, where Z = any element above He) of the star particles bound to this Subhalo, but restricted to stars within twice the stellar half mass radius.
SubhaloStarMetallicityHalfRad	N	-	Same as SubhaloStarMetallicity, but restricted to stars within the stellar half mass radius.
SubhaloStarMetallicityMaxRad	N	-	Same as SubhaloStarMetallicity, but restricted to stars within the radius of V_{max}.
SubhaloStellarPhotometrics	N,8	*mag*	Eight bands: U, B, V, K, g, r, i, z. Magnitudes based on the summed-up luminosities of all the stellar particles of the group. For details on the bands, see snapshot details.
SubhaloStellarPhotometrics MassInRad	N	$10^{10}\mathrm{M}_\odot/h$	Sum of the mass of the member stellar particles, but restricted to stars within the radius SubhaloStellarPhotometricsRad.
SubhaloStellarPhotometricsRad	N	ckpc/h	Radius at which the surface brightness profile (computed from all member stellar particles) drops below the limit of 20.7 mag arcsec^{-2} in the K band (in comoving units).
SubhaloVel	N,3	km/s	Peculiar velocity of the group, computed as the sum of the mass weighted velocities of all particles/cells in this group, of all types.
SubhaloVelDisp	N	km/s	One-dimensional velocity dispersion of all the member particles/cells (the 3D dispersion divided by $\sqrt{3}$).
SubhaloVmax	N	km/s	Maximum value of the spherically-averaged rotation curve.
SubhaloVmaxRad	N	kpc/h	Comoving radius of rotation curve maximum (where V_{max} is achieved).
SubhaloWindMass	N	$10^{10}\mathrm{M}_\odot/h$	Sum of masses of all wind-phase cells in this subhalo (with Type==4 and BirthTime<= 0).

Table 5.18: Description of all fields in the Header group of the group catalog files. Each header field is an attribute.

Field	Type	Description
SimulationName	string	e.g. 'Illustris-1' or 'Illustris-2-Dark'
SnapshotNumber	int	snapshot number (should be consistent with filename)
Ngroups_ThisFile	int	Number of groups within this file chunk.
Nsubgroups_ThisFile	int	Number of subgroups within this file chunk.
Ngroups_Total	int	Total number of groups for this snapshot.
Nsubgroups_Total	int	Total number of subgroups for this snapshot.
NumFiles	int	Total number of file chunks the group catalog is split between.
Num_ThisFile	int	Index of this file chunk (should be consistent with the filename).
Time	float	Scale factor of the snapshot corresponding to this group catalog.
Redshift	float	Redshift of the snapshot corresponding to this group catalog.
BoxSize	float	Side-length of the periodic volume in code units.
FileOffsets_Snap	$[N_c, 6]$ int array	The offset table (by type) for the snapshot files, giving the first particle index in each snap file chunk. Determines which files(s) a given offset+length will cover. A two-dimensional array, where the element (i, j) equals the cumulative sum (i.e. offset) of particles of type i in all snapshot file chunks prior to j.
FileOffsets_Group	$[N_c]$ int array	The offset table for groups in the group catalog files. A one-dimensional array, where the i^{th} element equals the first group number in the i^{th} groupcat file chunk.
FileOffsets_Subhalo	$[N_c]$ int array	The offset table for subhalos in the group catalog files. A one-dimensional array, where the i^{th} element equals the first subgroup number in the i^{th} groupcat file chunk.
FileOffsets_SubLink	$[N_c]$ int array	The offset table for trees in the SubLink files. A one-dimensional array, where the i^{th} element equals the first tree number in the i^{th} SubLink file chunk.

Table 5.19: Description of all fields in the Offsets group of the group catalog files. Note that all three LHaloTree or SubLink values equal -1 if that subhalo is not in the respective merger tree, which can occur if searching at a snapshot prior to $z = 0$. For the offsets, N_c indicates the number of file chunks (or pieces) over which that data product has been split.

Field	Dimensions	Description
Group_SnapByType	Ngroups_Total,6	The offset table for a given group number (by type), into the snapshot files. That is, the global particle index (across all snap file chunks) of the first particle of this group. A two-dimensional array, where the element (i, j) equals the cumulative sum (i.e. offset) of particles of type i in all groups prior to group number j.
Group_FuzzByType	Ngroups_Total,6	Offset into the "outer fuzz" (at the end of each snapshot file) for this group.
Subhalo_SnapByType	Nsubgroups_Total,6	The offset table for a given subhalo number (by type), into the snapshot files. That is, the global particle index (across all snap file chunks) of the first particle of this subhalo. A two-dimensional array, where the element (i, j) equals the cumulative sum (i.e. offset) of particles of type i in all subhalos prior to subhalo number j.
Subhalo_LHaloTreeFile	Nsubgroups_Total	The LHaloTree file number with the tree which contains this subhalo.
Subhalo_LHaloTreeNum	Nsubgroups_Total	The number of the tree within the above file within which this subhalo is located (e.g. TreeX).
Subhalo_LHaloTreeIndex	Nsubgroups_Total	The LHaloTree index within the above tree dataset at which this subhalo is located.
Subhalo_SublinkRowNum	Nsubgroups_Total	The SubLink global index of the location of this subhalo.
Subhalo_SublinkSubhaloID	Nsubgroups_Total	The SubLink ID of this subhalo.
Subhalo_SublinkLastProgenitorID	Nsubgroups_Total	The SubLink ID of the last progenitor of this tree (all the subhalos contained in the tree rooted in this subhalo are the ones with IDs between SubhaloID and LastProgenitorID).

5.3.3 Merger Trees

Merger trees have been created for the various Illustris simulations using SubLink (Rodriguez-Gomez et al., 2015), LHaloTree (Springel et al., 2005a), and Consistent-Trees (using Rockstar) (Behroozi et al., 2013). The LHaloTree are essentially identical to the primary trees of the Millennium simulations, Aquarius, and Phoenix, but in HDF5 format. In the population average sense the different merger trees give similar results. In more detail, the exact merger history or mass assembly history for any given halo may differ. For any given science goal, one type of tree may be more or less useful, and users are free to use whichever they prefer. These codes are all included in the Sussing Merger Trees comparison project (Srisawat et al., 2013).

Figure 5.4 shows a schematic of the structure of both the SubLink and LHaloTree merger trees. It is not necessary to understand the complete details of the trees to practically use them. In particular, the only critical links are the 'descendant' (black), 'first progenitor' (green), and 'next progenitor' (red) associations. These are shown for all tree nodes in the diagram. For their exact definitions, see Tables 5.20 and 5.21, the LHaloTree and SubLink tables. Walking back in time following along the main (most massive) progenitor branch consists of following the first progenitor links until they end (value equals -1). Similarly, walking forward in time along the descendants branch consists of following the descendant links until they end (value equals -1), which typically occurs at $z = 0$. The full progenitor history, and not just the main branch, requires following both the first and next progenitor links. In this way the user can identify all subhalos at a previous snapshot which have a common descendant. Examples of walking the tree are provided in the example scripts.

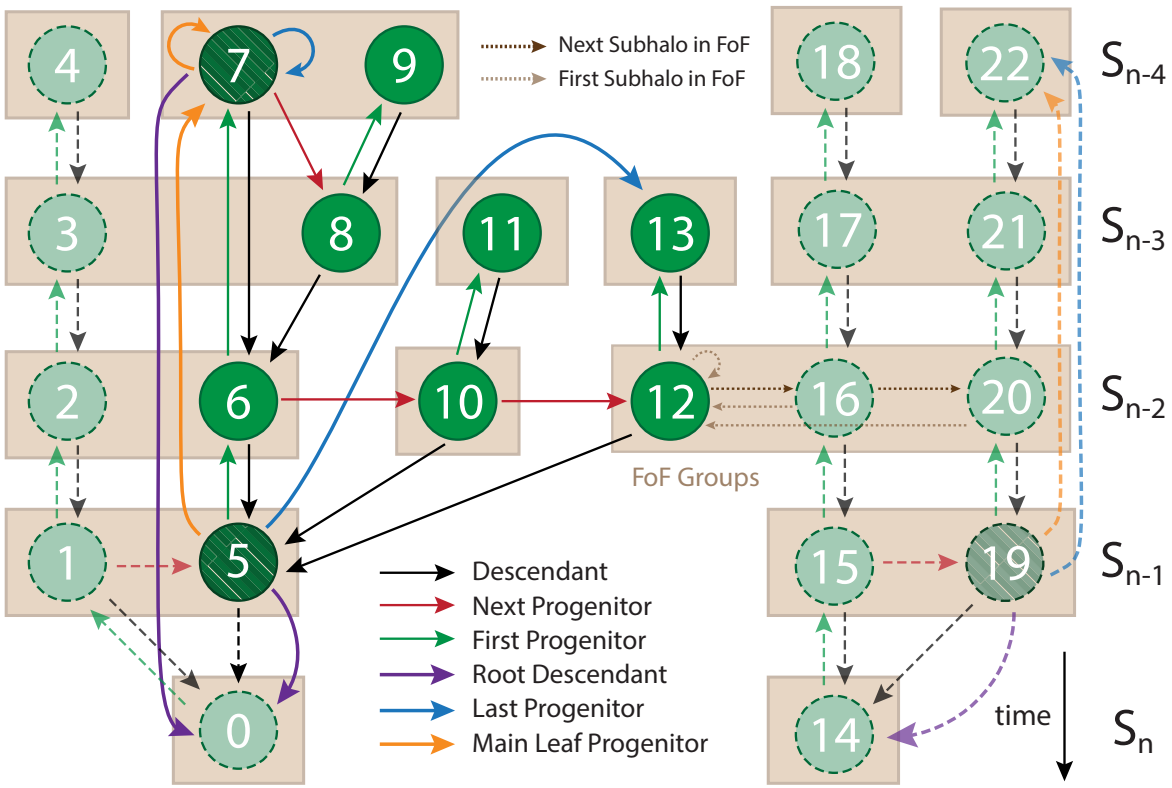

Figure 5.4 Schematic diagram of the merger tree structure for both SubLink and LHaloTree. Both algorithms connect subhalos (i.e., Subfind halos) across different snapshots in the simulation. Rows indicate discrete snapshots, with time increasing downwards towards redshift zero (the horizontal axis is arbitrary). Green circles represent subhalos (the nodes of the merger tree), while beige boxes indicate the grouping of the subhalos into their parent FoF groups. The most important links are for the descendant (black), first progenitor (green), and next progenitor (red), which are shown for all subhalos. The root descendant (purple), last progenitor (blue), and main leaf progenitor (orange) links exist only for the SubLink trees, and for simplicity these last three link types are shown only for subhalos 5, 7, and 19 (darker striped circles). For exact definitions of each link type, see the corresponding tables. For more information about this figure, consult the text.

The number inside each circle from the figure is the unique ID (within the whole simulation) of the corresponding subhalo, which is assigned in a depth-first fashion. Numbering also indicates the on-disk storage ordering for the SubLink trees, which adopt the approach of Lemson & Virgo Consortium (2006). For example, the main

progenitor branch (from 5-7 in the example) and the full progenitor tree (from 5-13 in the example) are both contiguous subsets of each merger tree field, whose location and size can be calculated using these links. The ordering within a single tree in the LHaloTree is not guaranteed to follow this scheme.

The 'root descendant' (purple), 'last progenitor' (blue), and 'main leaf progenitor' (orange) links exist only for the SubLink trees. For simplicity, these last three link types are shown only for nodes 5, 7, and 19 (darker striped circles). Using these links is optional, but allows efficient extraction of main progenitor branches, subtrees (i.e., the set containing a subhalo and "all" its progenitors), "forward" descendant branches, and other subsets of the tree. For their full definitions, see Table 5.20 with the SubLink details.

Each subhalo spans a "subtree" consisting of the subhalo itself and all its progenitors. As an example, the subhalos belonging to the subtree of subhalo 5 are shown in darker green in the figure. Other subhalos not belonging to this subtree are shown in lighter green, and their links are indicated with dashed arrows. In the SubLink trees, the subtree of any subhalo can be extracted easily using the 'last progenitor' pointer. As shown in the figure, since subhalo 13 is the 'last progenitor' of subhalo 5, the subtree of subhalo 5 consists of all subhalos with IDs between 5 and 13. Similarly, the main progenitor branch of any subhalo can be retrieved efficiently using the 'main leaf progenitor' link.

Both SubLink and LHaloTree contain the links 'first subhalo in FoF group' (light brown dotted arrow) and 'next subhalo in FoF group' (dark brown dotted arrow), which connect subhalos that belong to the same FoF group. The FoF groups do not

play a direct role in the construction of the merger tree. However, in SubLink, subhalos that belong to the same FoF group are also considered to be part of the same tree. As a result, two otherwise independent trees (based on the progenitor and descendant links) are considered to be the same tree if they are "connected" by a FoF group. This is exemplified in the figure by the FoF group containing subhalos 12, 16, and 20. This FoF group acts as a "bridge" between the left and right trees.

SubLink

SubLink constructs merger trees at the subhalo level. A unique descendant is assigned to each subhalo in three steps (see Rodriguez-Gomez et al. (2015)). First, descendant candidates are identified for each subhalo as those subhalos in the following snapshot that have common particles with the subhalo in question. Second, each of the descendant candidates is given a score based on a merit function that takes into account the binding energy rank of each particle. Third, the unique descendant of the subhalo in question is the descendant candidate with the highest score. Sometimes the halo finder does not detect a small subhalo that is passing through a larger structure, because the density contrast is not high enough. SubLink deals with this issue by allowing some subhalos to skip a snapshot when finding a descendant. Once all descendant connections have been made, the main progenitor of each subhalo is defined as the one with the "most massive history" behind it (following De Lucia & Blaizot (2007)).

The SubLink merger tree is one large data structure split across several sequential HDF5 files named `tree_extended.[fileNum].hdf5`, where [fileNum] goes from e.g. 0 to 9 for the Illustris-1 run. These files store the data on a per tree basis, and therefore

are completely independent from each other. More specifically, any two subhalos that are connected by any of the pointers described in the SubLink table are guaranteed to belong to the same tree, and, therefore, their data is found in the same file. Table 5.20 lists the fields which are present in each file. Note that in addition to the tree fields, all subhalo fields are also present, copied exactly from the Subfind catalogs. The advantage is that they are ordered in the same order as the tree structure. See the group catalog description for their units and descriptions. The Group_M_Crit200, Group_M_Mean200, and Group_M_Tophat200 fields are also present, but are FoF group quantities, such that all subhalos in the same FOF group will have the same value for these three fields.

Table 5.20: Listing of all fields and their descriptions for the SubLink merger trees.

Field	Type	Description
SubhaloID	int64	Unique identifier of this subhalo, assigned in a "depth-first" fashion (Lemson & Virgo Consortium, 2006). This value is contiguous within a single tree.
SubhaloIDRaw	int64	Unique identifier of this subhalo in raw format (= SnapNum$\times 10^{12}$ + SubfindID).
LastProgenitorID	int64	The SubhaloID of the last progenitor of the tree rooted at this subhalo. Since the SubhaloIDs are assigned in a "depth-first" fashion, all the subhalos contained in the tree rooted at this subhalo are the ones with SubhaloIDs between (and including) the SubhaloID and LastProgenitorID of this subhalo. For subhalos with no progenitors, LastProgenitorID == SubhaloID.
MainLeafProgenitorID	int64	The SubhaloID of the last progenitor along the main branch, i.e. the earliest progenitor following the FirstProgenitorID. For subhalos with no progenitors, MainLeafProgenitorID == SubhaloID.
RootDescendantID	int64	The SubhaloID of the latest subhalo that can be reached by following the DescendantID link, i.e. the root of this tree . For subhalos with no descendants, RootDescendantID == SubhaloID.
TreeID	int64	Unique identifier of the tree to which this subhalo belongs.
SnapNum	int16	The snapshot in which this subhalo is found.
FirstProgenitorID	int64	The SubhaloID of this subhalo's first progenitor. The first progenitor is the one with the "most massive history" behind it. For subhalos with no progenitors, FirstProgenitorID == -1.
NextProgenitorID	int64	The SubhaloID of the subhalo with the next most massive history with the same descendant as this subhalo. If there are no more subhalos with the same descendant, NextProgenitorID == -1.
DescendantID	int64	The SubhaloID of this subhalo's descendant. If it has none, DescendantID == -1.
FirstSubhaloInFOFGroupID	int64	The SubhaloID of the first subhalo (with the most massive history) from the same FOF group.
NextSubhaloInFOFGroupID	int64	The SubhaloID of the next subhalo (ordered by their mass history) from the same FOF group. If there are no more subhalos in the same FOF group, NextSubhaloInFOFGroupID == -1.
NumParticles	uint32	Number of particles in the current subhalo which were used in the merger tree to determine descendants (e.g. DM-only or stars + star-forming gas).
MassHistory	float32	Sum of the masses of all progenitors on the main branch (De Lucia & Blaizot, 2007), in $10^{10}\mathrm{M}_\odot/h$.
SubfindID	int32	Index of this subhalo in the Subfind group catalog.

LHaloTree

The LHaloTree algorithm is virtually identical to that used for the Millennium simulation, constructing trees based on subhalos instead of main halos, described fully in the supplementary information of Springel et al. (2005b). In short, to determine the appropriate descendant, the unique IDs that label each particle are tracked between outputs. For a given halo, the algorithm finds all halos in the subsequent output that contain some of its particles. These are then counted in a weighted fashion, giving higher weight to particles that are more tightly bound in the halo under consideration, and the one with the highest count is selected as the descendant. In this way, preference is given to tracking the fate of the inner parts of a structure, which may survive for a long time upon infall into a bigger halo, even though much of the mass in the outer parts can be quickly stripped. To allow for the possibility that halos may temporarily disappear for one snapshot, the process is repeated for Snapshot n to Snapshot $n + 2$. If either there is a descendant found in Snapshot $n + 2$ but none found in Snapshot $n + 1$, or, if the descendant in Snapshot $n + 1$ has several direct progenitors and the descendant in Snapshot $n + 2$ has only one, then a link is made that skips the intervening snapshot.

The LHaloTree merger tree is one large data structure split across several HDF5 files named `trees_sf1_135.[chunkNum].hdf5`, where [chunkNum] goes from e.g. 0 to 511 for the Illustris-1 run. Within each file there are a number of groups named "TreeX", where X corresponds to the FoF group number in the group catalogs at the final snapshot. However, note that the number X starts over at zero for each tree file chunk, so to recover the FoF group number you will have to add the sum of the number of trees in the previous tree file chunks. The pair (SubhaloNumber,SnapNum) provides

the indexing into the SUBFIND group catalog. The five other indices for each entry in a TreeX group index into that same group in the tree file. Table 5.21 describes the fields in the Header and TreeX groups. Note that in addition to the tree fields, the majority of subhalo fields are also present, copied exactly from the SUBFIND catalogs. The advantage is that they are ordered in the same order as the tree structure. See the group catalog description for their units and descriptions. The Group_M_Crit200, Group_M_Mean200, and Group_M_Tophat200 fields are also present, but since they are FoF group quantities, all subhalos from the same FOF group will have the same value for these three fields.

Table 5.21: Listing of all fields in the LHaloTree merger trees.

Field	Dimensions	Description
		Header Groups
Redshifts	{N_snap}	List of redshifts of the snapshots used to create this merger tree.
TotNsubhalos	{N_snap}	The number of Subfind groups in the group catalog, for each snapshot used to make this tree.
TreeNHalos	{N_halos}	The size of {N} for each TreeX group in this file, e.g. the total number of halos in that group.
FirstSnapshotNr	1	First snapshot number used to make these merger trees (should be 0).
LastSnapshotNr	1	Last snapshot number used to make these merger trees (should be 135).
SnapSkipFac	1	Snapshot stride when making these merger trees (should be 1).
NtreesPerFile	1	The size of {N_halos} for this file, can be used to calculate the offset to map a FoF group number to a TreeX group name (made to be roughly equal across chunks).
NhalosPerFile	1	The total number of tree members (subhalos) in this file. Equals the sum of all elements of TreeNHalos.
ParticleMass	1	The dark matter particle mass used to make these merger trees, in units of $10^{10}\mathrm{M}_\odot/h$.
		TreeX Groups
SubhaloNumber	(N)	The ID of this subhalo, indexing the Subfind group catalog at SnapNum.
Descendant	(N)	The index of the subhalo's descendant within the tree, if any (-1 otherwise). Indexes this TreeX group.
FirstProgenitor	(N)	The index of the subhalo's first progenitor within the merger tree, if any (-1 otherwise). The first progenitor is defined as the most massive one. (-1 if none) Indexes this TreeX group.
NextProgenitor	(N)	The index of the next subhalo from the same snapshot which shares the same descendant, if any (-1 if this is the last). Indexes this TreeX group.
FirstHaloInFOFGroup	(N)	The index of the most massive subhalo in the same FOF group. Indexes this TreeX group.
NextHaloInFOFGroup	(N)	The index of the next subhalo in the same FOF group (-1 if none). Indexes this TreeX group.
FileNr	(N)	File number in which the subhalo is found. Redundant, i.e. for a given [chunkNum] file, this array will be constant and equal to [chunkNum].
SnapNum	(N)	The snapshot in which this subhalo was found.

5.3.4 Supplementary Data Catalogs

The following additional data products have been computed in post-processing, based on the raw simulation outputs. They are either already available, and now unified under the Illustris data release and made available through the API, or are now made available. In the current effort we focus exclusively on additional properties derived for Illustris-1 galaxies, exclusively at $z = 0$ and above a stellar mass limit of $M_\star \gtrsim 10^9 \mathrm{M}_\odot$.

Stellar Mocks: Multi-band Images and SEDs

A catalog of synthetic stellar images and integrated spectra of galaxies in Illustris-1 at $z = 0$, produced using the radiative transfer code SUNRISE. For complete details on this data product, see Torrey et al. (2015) where it was first described and made available. For all galaxies with stellar masses $M_\star > 10^{10}\mathrm{M}_\odot$ ($\sim 10^4$ star particles and above), both integrated SEDs and spatially resolved photometric maps in 36 broadband filters are computed. There are approximately 7000 galaxies above this limit. For all galaxies with smaller stellar masses, down to 500 star particles, only integrated SEDs are calculated. The 36 bands include GALEX, SDSS, IRAC, Johnson, 2MASS, ACS, and preliminary NIRCAM filters. Note that this is the only data product which is in a format other than HDF5 (namely, FITS). However, the API provides extractions of individual bands and viewing angles in HDF5 format, as well as SEDs in text format, if requested. Finally, we have developed the Python code SUNPY[2] to add observational realism and make figures based on the raw stellar mock image FITS files.

[2] http://github.com/ptorrey/sunpy

5.3.5 Photometric Non-Parametric Stellar Morphologies

A catalog of photometric non-parametric morphologies of Illustris-1 galaxies at $z = 0$. This is meant to replicate automated diagnostics of galaxy stellar structure commonly used observationally, and is calculated by first adding observational realism to the idealized 'stellar mock' images from Torrey et al. (2015), then measuring $(G_{\rm ini}, M_{20}, C, r_P, r_E)$ statistics in four bands, rest-frame u, g, i, and H, each from four directions. For full details on the calculation of each value, see Table 5.22 and Snyder et al. (2015) (following Lotz et al., 2004). This data is available for essentially all subhalos with $M_\star > 10^{9.7}{\rm M}_\odot$ at $z = 0$ in Illustris-1. Treating each viewing direction independently, values have been computed for a uniform set of 42531 sources per filter.

Table 5.22: Details of the supplementary data catalog: Photometric Non-Parametric Stellar Morphologies. The four bands which replace band_name are: gSDSS, iSDSS, uSDSS, and hWFC3 (WFC3-IR/F160W). The four camera views are indexed 0, 1, 2, and 3, and the SubfindID list differs for each.

Group Name	Units	Description
/Snapshot_135/SubfindID_cam0,1,2,3	-	The SUBFIND IDs these values correspond to (different for each camera view, but the same for all bands and fields). 10654,10618,10639,10620 entries.
/Snapshot_135/band_name/Gini_cam0,1,2,3	-	The G_ini coefficient, which measures the relative distribution of the galaxy pixel flux values.
/Snapshot_135/band_name/M20_cam0,1,2,3	-	M_{20}, the second-order moment of the brightest 20% of the galaxy's flux.
/Snapshot_135/band_name/C_cam0,1,2,3	-	The concentration parameter C.
/Snapshot_135/band_name/RP_cam0,1,2,3	*kpc*	The elliptical Petrosian radius r_P.
/Snapshot_135/band_name/RE_cam0,1,2,3	*kpc*	The elliptical half-light radius r_E.

5.3.6 Stellar Circularities, Angular Momenta, Axis Ratios

A catalog for the circularities, angular momenta and axis ratios of the stellar component, for Illustris-1 galaxies at z=0. Data is available for all subhalos with stellar mass (inside twice the stellar half mass radius) bigger than $10^9 M_\odot$. For complete definitions on the calculation of each value, see Table 5.23 and Genel et al. (2015), where they were presented and used. The first four quantities are calculated after alignment with the angular momentum vector, and measure the quantities inside 10 times the stellar half-mass radius. The "Circ*" fields are based on the distribution of the circularity parameter ϵ of the individual stars, as defined in Equation (2) of Marinacci et al. (2014a).

Table 5.23: Details of the supplementary data catalog: Stellar Circularities, Angular Momenta, and Axis Ratios.

Group Name	Units	Description
/Snapshot_135/ SubfindID	-	The SUBFIND IDs these values correspond to (27345 entries).
/Snapshot_135/ SpecificAngMom	km/s × kpc	The specific angular momentum of the stars.
/Snapshot_135/ CircAbove07Frac	-	The fraction of stars with $\epsilon > 0.7$. This is a common definition of the disk stars - those with significant (positive) rotational support.
/Snapshot_135/ CircAbove07 MinusBelowNeg07Frac	-	The fraction of stars with $\epsilon > 0.7$ minus the fraction of stars with $\epsilon < -0.7$. This removes the contribution of the bulge to the disk, assuming the bulge is symmetric around $\epsilon = 0$.
/Snapshot_135/ CircTwiceBelow0Frac	-	The fraction of stars with $\epsilon < 0$, multiplied by two. This is another common way in the literature to define the bulge.
/Snapshot_135/ MassTensorEigenVals	kpc	Three numbers for each galaxy, which are the eigenvalues of the mass tensor of the stellar mass inside the stellar $2R_{1/2}$. This means that in a coordinate system that is aligned with the eigenvectors (principal axes), the component i equals $M_i \equiv \sqrt{\sum_j m_j r_{j,i}^2} / \sqrt{\sum_j m_j}$, where j enumerates over stellar particles inside that radius, $r_{j,i}$ is the distance of stellar particle j in the i axis from the most bound particle of the galaxy, and m_j is its mass, and $i \in (1, 2, 3)$. They are sorted such that $M_1 < M_2 < M_3$. Example use: $M_1/\sqrt{M_2 M_3}$ can represent the flatness of the galaxy.
/Snapshot_135/ ReducedMass TensorEigenVals	-	Similar to the above, except less weight is given to further away particles. The orientation of the system is the same, but the quantity measured for each axis is instead $M_i \equiv \sqrt{\sum_j m_j r_{j,i}^2 / R_j^2} / \sqrt{\sum_j m_j}$, where $R_j \equiv \sum_i r_{j,i}^2$ is the distance of star j from the centre of the galaxy.

5.4 Data Access

There are two complementary ways to access the Illustris data products.

1. Raw files can be directly downloaded, and example scripts are provided as a starting point for local analysis.

2. A web-based API can be used, either through a web browser or programmatically in an analysis script, to perform common search and extraction tasks.

These two approaches can be combined. For example, you may be forced to download the full redshift zero group catalog in order to perform a complex search not supported by the API. After locally determining a sample of interesting galaxies, you could then extract their individual merger trees (and/or raw particle data) without needing to download the full simulation merger tree (or a full snapshot).

Both approaches are documented below, while "getting started" tutorials for several languages (currently: Python, IDL, and Matlab) can be found online.

5.4.1 Direct File Download and Example Scripts

All of the primary data products for Illustris are released in HDF5 format. This is a portable, self-describing, binary specification suitable for large numerical datasets, for which file access routines are available in all common computing languages. We use only the basic features of the format: groups, attributes, and datasets, with one and two dimensional numeric arrays.

In order to maintain reasonable filesizes, most outputs are split across multiple file "pieces" (or "chunks"). For example, each snapshot of Illustris-1 is split into 512 sequentially numbered files. Individual links to each file chunk are available through the web-based API, and a snapshot can be downloaded in its entirety with a single wget command. Direct download links for other snapshots, simulations, and file types (such as group catalogs or merger trees) can be found at the appropriate URLs, as described below. Pre-computed sha256 checksums are provided for all files so that their integrity can be verified.

The provided example scripts (in IDL, Python, and Matlab) provide such basic I/O functionality as: (i) reading a given particle type and/or data field from the snapshot files, (ii) reading only the particle subset from the snapshot corresponding to a halo or subhalo, (iii) extracting the full subtree or main progenitor branch from either SubLink or LHaloTree for a given subhalo, (iv) walking a tree to count the number of mergers, (v) reading the entire group catalog at one snapshot, (vi) reading specific fields from the group catalog, or the entries for a single halo or subhalo. We expect they will provide a useful starting point for writing any analysis task, and intend them as a 'minimal working examples' which are short and simple enough that they can be quickly understood and extended.

5.4.2 Web-based API

We have implemented a web-based interface (API) which can respond to a variety of user requests and queries. It is a well-defined interface between the user and the Illustris data products, which is expressed in terms of the required input(s) and expected output(s)

for each type of request. The provided functionality is independent, as much as possible, from the underlying data structure, heterogeneity, format, and access methods. The API can be used in addition to, or in place of, the download and local analysis of large data files. At a high level, the API allows a user to **search**, **extract**, **visualize**, and **analyze**. In each case, the goal is to reduce the data response size, either by extracting an unmodified subset, or by calculating a derivative quantity.

By specific example, the following types of requests can be handled through the current API, for any simulation at any snapshot:

- List the available simulations, their snapshots, and all associated metadata.
- List all objects in the SUBFIND group catalog and their properties.
- Search with numeric range(s) over any field(s) present in the SUBFIND group catalogs.
- Return all fields from the group catalog for a specific halo or subhalo.
- Return a full snapshot cutout of the particle/cell data for a given halo or subhalo.
- Return a subset of this 'group cutout' containing only specified particle/cell type(s), and/or specific field(s) for each type.
- Return the complete merger history, or just the main progenitor branch, for a given subhalo, for any of the merger trees.
- Download all raw snapshot, group catalog, merger tree, and supplementary data catalog files which exist.
- Download subsets of raw snapshot files, containing only specified particle/cell type(s), and/or specific field(s) for each type.

- Crossmatch subhalos between full physics runs and their dark matter only analogues.
- Traverse relationships between halos and subhalos, for instance from a satellite subhalo to its parent FoF group to the primary (central) subhalo of that group.
- Traverse descendant and primary progenitor links across adjacent snapshots, as available in the SubLink merger trees.
- View or render visualizations of the different components (e.g. dark matter, gas, stars) of halos and subhalos, when available.
- Retrieve or calculate additional properties, beyond what is available in the group catalogs, for halos and subhalos, when available.

The Illustris data access API is available at the following permanent URL:

- `http://www.illustris-project.org/api/`

Simple Python examples for working with the API are provided in Appendix B.

We provide a list of endpoints, their descriptions, and return types. All accept only GET requests. To provide long-term consistency, we anticipate that the API structure described herein will never change. As additional data products, simulations, tools, and analysis tasks are developed and released, new endpoints will be added. In order to take advantage of new features as they are introduced, we recommend a user consult the up to date API reference available on the website. Tables 5.24, 5.25 and 5.26 provide descriptions of each currently available endpoint.

Table 5.24: API Endpoint Descriptions and Reference (I): simulation and snapshot meta-data, subhalos and halos.

Endpoint	Description	Return Type
/api/	list all simulations currently accessible to the user	json,api (?format=)
/api/{sim_name}/	list metadata (including list of all snapshots+redshifts) for {sim_name}	json,api (?format=)
/api/{sim_name}/snapshots/	list all snapshots which exist for this simulation	json,api (?format=)
/api/{sim_name}/snapshots/{num}/	list metadata for snapshot {num} of simulation {sim_name}	json,api (?format=)
/api/{sim_name}/ snapshots/z={redshift}/	redirect to the snapshot which exists closest to {redshift} (with a maximum allowed error of 0.1 in redshift)	json,api (?format=)
define [base] = /api/{sim_name}/snapshots/{num} or [base] = /api/{sim_name}/snapshots/z={redshift}		
Subfind Subhalos		
[base]/subhalos/	paginated list of all subhalos for this snapshot of this run	json,api (?format=)
[base]/subhalos/?{search_query}	execute {search_query} over all subhalos, return those satisfying the search with basic fields and links to /subhalos/{id}	json,api (?format=)
[base]/subhalos/{id}	list available data fields and links to all queries for subhalo {id}	json,api (?format=)
[base]/subhalos/{id}/info.json	extract all group catalog fields for subhalo {id}	json (.ext)
[base]/subhalos/{id}/cutout.hdf5	return snapshot cutout of subhalo {id}, all particle types and fields	HDF5 (.ext)
[base]/subhalos/{id}/cutout.hdf5 ?{cutout_query}	return snapshot cutout of subhalo {id} corresponding to the {cutout_query}	HDF5 (.ext)
FoF Halos		
[base]/halos/{halo_id}/	list what we know about this FoF halo, in particular the 'child_subhalos'	json,api (?format=)
[base]/halos/{halo_id}/info.json	extract all group catalog fields for halo {halo_id}	json (.ext)
[base]/halos/{halo_id}/cutout.hdf5	return snapshot cutout of halo {halo_id}, all particle types and fields	HDF5 (.ext)
[base]/halos/{halo_id}/ cutout.hdf5?{cutout_query}	return snapshot cutout of halo {halo_id} corresponding to the {cutout_query}	HDF5 (.ext)

Table 5.25: API Endpoint Descriptions and Reference (II): merger trees, supplementary data catalogs, file downloads.

Endpoint	Description	Return Type
	Merger Trees	
[base]/subhalos/{id}/ lhalotree/full.hdf5	retrieve full tree (flat HDF5 format or hierchical/nested JSON format)	HDF5,json (.ext)
[base]/subhalos/{id}/ lhalotree/mpb.hdf5	retrieve only main progenitor branch (towards higher redshift for this subhalo)	HDF5,json (.ext)
[base]/subhalos/{id}/ sublink/full.hdf5	same as above for 'lhalotree' but for sublink	HDF5,json (.ext)
[base]/subhalos/{id}/ sublink/mpb.hdf5	same as above for 'lhalotree' but for sublink	HDF5,json (.ext)
	supplementary data: stellar mocks	
[base]/subhalos/{id}/ stellar_mocks/broadband.fits	download raw broadband fits file for subhalo {id}	FITS (.ext)
[base]/subhalos/{id}/stellar_mocks/ broadband.hdf5?view={view}	download subset of broadband fits file for subhalo {id}: all 36 bands for view number {view}	HDF5 (.ext)
[base]/subhalos/{id}/stellar_mocks/ broadband.hdf5?band=band	download subset of broadband fits file for subhalo {id}: all 4 views for band {band} (1-indexed number, or name)	HDF5 (.ext)
[base]/subhalos/{id}/ stellar_mocks/image.png	download stellar mock png 2D image (subhalo particles only)	PNG (.ext)
[base]/subhalos/{id}/ stellar_mocks/image_fof.png	download stellar mock png 2D image (all group particles)	PNG (.ext)
[base]/subhalos/{id}/ stellar_mocks/image_gz.png	download stellar mock png 2D image ('galaxy zoo' image w/ realistic noise and background)	PNG (.ext)
[base]/subhalos/{id}/ stellar_mocks/sed.txt	download stellar mock integrated 1D SED for subhalo {id}	txt,json (.ext)

Table 5.26: API Endpoint Descriptions and Reference (III): file downloads.

Endpoint	Description	Return Type
direct file downloads		
define [base] = /api/sim_name/files		
[base]/	list of each 'files' type available for this simulation (excluding those attached to specific snapshots)	json,api (?format=)
[base]/snapshot-{num}/	list of all the actual file chunks to download snapshot {num}	json,api (?format=)
[base]/snapshot-{num}.{chunknum}.hdf5	download chunk {chunknum} of snapshot {num}	HDF5 (.ext)
[base]/snapshot-{num}.{chunknum}.hdf5?{cutout_query}	download only {cutout_query} of chunk {chunknum} of snapshot {num}	HDF5 (.ext)
[base]/groupcat-{num}/	list of all the actual file chunks to download group catalog (fof/subfind) for snapshot {num}	json,api (?format=)
[base]/groupcat-{num}.{chunknum}.hdf5	download chunk {chunknum} of group catalog for snapshot {num}	HDF5 (.ext)
[base]/lhalotree/	list of all the actual file chunks to download LHaloTree merger tree for this simulation	json,api (?format=)
[base]/lhalotree.{chunknum}.hdf5	download chunk {chunknum} of LHaloTree merger tree for this simulation	HDF5 (.ext)
[base]/sublink/	list of all the actual file chunks to download SubLink merger tree for this simulation	json,api (?format=)
[base]/sublink.{chunknum}.hdf5	download chunk {chunknum} of SubLink merger tree for this simulation	HDF5 (.ext)

API Access Details

Each API endpoint can return a response in one or more data types. When multiple options exist, a specific return format can be requested through one of the following methods.

- "(?format=)" indicates that the return type is chosen by supplying such a querystring, appended to the URL.
- "(.ext)" indicates that the return type is chosen by supplying the desired file extension in the URL.

Search and Cutout Requests. Several API functions accept additional, optional parameters, which are described here.

{search_query} is an AND combination of restrictions over any of the supported fields, where the relations supported are 'greater than' (gt), 'greater or equal to' (gte), 'less than' (lt), 'less than or equal to' (lte), 'equal to'. The first four work by appending e.g. '__gt=val' to the field name (using a double underscore). For example:

- mass_dm__gt=90.0
- mass__gt=10.0&mass__lte=20.0
- vmax__lt=100.0&len__gas=0&vmaxrad__gt=20.0

{cutout_query} is a concatenated list of particle fields, separated by particle type. The allowed particle types are 'dm','gas','stars','bhs'. The field names are exactly as in

the snapshots ("all" is allowed). Omitting all particle types will return the full cutout: all types, all fields. For example:

- gas=Masses,Coordinates,Velocities
- dm=Coordinates&stars=all

Authentication. All API requests require authentication, and therefore also user registration. Each request must provide, along with the details of the request itself, the unique "API Key" of the user making the request. A user can send their API key in the querystring, by appending it to the URL as:

- ?api_key=d22d1f16b894a0b894ec31

A user can alternatively send their API key in HTTP header. This is particularly useful for wget commands or within scripts (see the API tutorial). Note that if a user is logged in to the website, then requests *from the browser* are automatically authenticated. Navigating the Browsable API works in this way.

5.4.3 Further Online Tools

Subhalo Search Form

We provide a simple search form through which users can query the subhalo database. The search capabilities that exist in the API are exposed in a more human-friendly interface, to enable exploration without the need to write code or write URLs by hand. For example, objects can be selected based on total mass, stellar mass, star formation

rate, gas metallicity, or size. The output is a familiar spreadsheet type format, which lists properties from the group catalogs. In addition, each subhalo row provides links to a common set of web-based tools for introspection. These include a full listing of all catalog fields, a form for selecting particle types and initiating an extraction of particles from the snapshot, merger tree visualization, and links to pre-rendered images, when available.

Explorer

The Illustris Explorer[3] is an experiment in the visualization, exploration, and dissemination of large data sets – in particular, those generated by large, astrophysical simulations such as Illustris. It uses the approach of thin-client interaction with derived data products, in this case, pre-computed imagery layered under group catalog information. In Figure 5.5 a full box slice of the simulation is shown in projection, with a depth of 15 Mpc/h, revealing a fifth of the total volume of Illustris at $z = 0$. All the imagery is rendered and saved as hierarchical image pyramids (see also Bertin et al. (2015)), while rapid search over group properties spatially overlays the results within this volume. All mass components of the simulation are present: the continuous gas and dark matter fields, stellar light from individual stars, and black holes. We have found the interface particularly useful in exploring the spatial relationships between these four components and the discrete halos and subhalos identified with substructure finding algorithms.

[3] `www.illustris-project.org/explorer/`

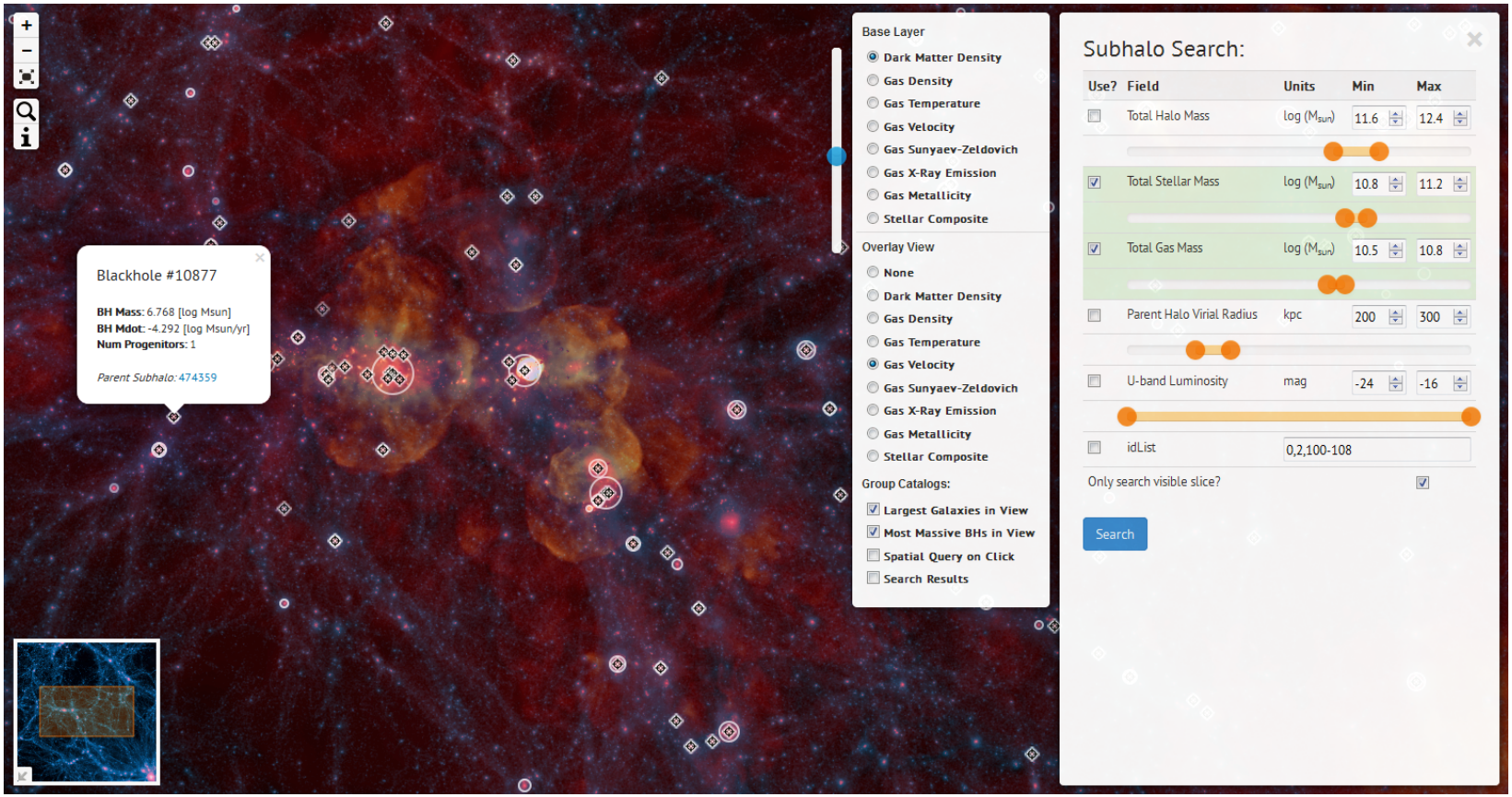

Figure 5.5 The current Illustris Explorer interface. The main view shows a gas velocity projection overlaid on the dark matter density field. The most massive galaxies currently visible are shown with circles, while black holes are represented with crosshairs. The overview in the lower left corner provides orientation on larger scales. Clicking at any location will launch a spatial search for the nearest subhalos, while clicking on a BH particle will query its details, including a link to its parent subhalo. The central panel controls image layer selection. The right panel presents a simple search interface over subhalo properties.

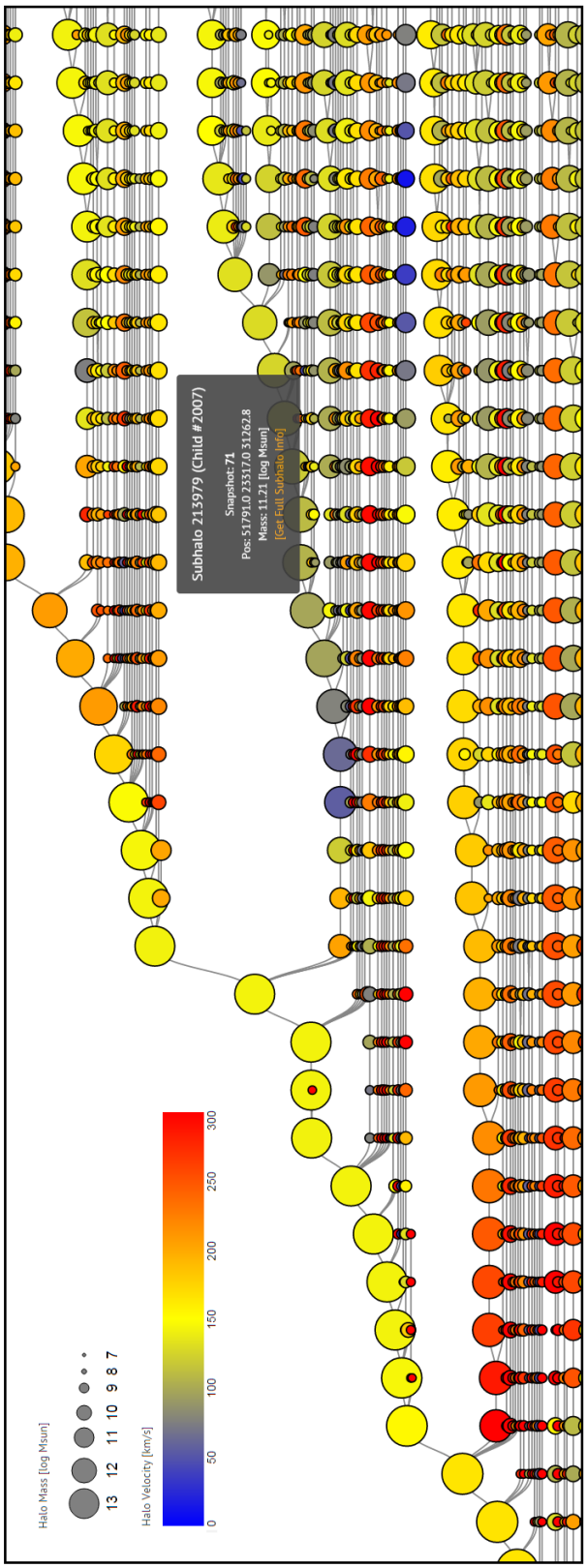

Figure 5.6 Example of interactive merger tree exploration. We show a zoomed-in portion of the SubLink tree for the 500th most massive central subhalo of Illustris-1 at $z = 0$ (ID 395444). Vector based, client-side rendering means that each node can be interacted with individually. One is shown displaying an informational popup, which includes a link back into the API for inspecting that particular progenitor subhalo. Here we show tree node size scaled with total halo mass in $\log \mathrm{M}_\odot$, and color mapped to subhalo velocity magnitude in km/s.

Merger Tree

As a demonstration of the potential of rich client applications built on top of the Illustris API, we show in Figure 5.6 the currently available interface for interactively exploring the merger trees.[4] A zoomed-in portion of the SubLink tree for the 500th most massive central subhalo of Illustris-1 at $z = 0$ is shown. For any run, snapshot, and subhalo combination, the browser requests a parseable representation of the merger tree from the API (in JSON format), and renders it using the scalable vector graphics (SVG) backend of the d3 javascript visualization library. Because the tree is vector based, and client side, each node can be interacted with individually. Here the informational popup provides a link, back into the API, where the details of the selected progenitor subhalo can be interrogated.

5.5 Scientific Remarks and Cautions

The Illustris Simulations (particularly Illustris-1) have been shown to resolve many details of the small-scale properties of galaxies, as well as the evolution of stars and gas within the cosmic web. Illustris-1 reproduces many observational facts on the demographics and properties of the galaxy populations at various epochs, and on the distribution of gas on large scales. As described in Section 5.2, this has been achieved with a comprehensive galaxy formation model which is intended to account for all the primary processes that are believed to be important for the formation and evolution of

[4]If logged in, you can launch this viewer from inside the Explorer, by selecting a subhalo ID or subhalo circle marker after a search, or through the general subhalo search form.

galaxies.

However, the enormous dynamical range and the variety and complexity of physics phenomena involved in these numerical endeavours necessarily involve some modeling uncertainties. We have identified below the known problems and points of caution in the Illustris simulated output that any user of the public data must be aware of before embarking on the analysis of the released products. These points should be carefully taken into account before advancing scientific conclusions or making comparisons to observational results.

5.5.1 Caveats with the Illustris Galaxy Formation Model

Limitations in the Illustris implementations of the stellar and AGN feedback, and possibly of the adopted star-formation recipe, determine a series of issues in the simulated galaxy populations and gas content of halos in comparison to observational constraints. These all point to an inefficient quenching of the star formation in galaxies at different masses and regimes, and in some cases also to qualitatively not-realistic behaviors of the feedback models. In particular, the following issues applicable to the highest-resolution realization (Illustris-1) must be noted.

- The cosmic star formation rate density is too high at $z \lesssim 1$, possibly because of an inefficient quenching of galaxies residing in halos of $10^{11-12}\mathrm{M}_{\odot}$ (see Figs. 8 and 2 in Vogelsberger et al., 2014a; Genel et al., 2014, respectively).
- The stellar mass function at $z \lesssim 1$ is too high both at the high and the low ends of the sampled stellar mass range, $M_{\star} \lesssim 10^{10}\mathrm{M}_{\odot}$ and $M_{\star} \gtrsim 10^{11.5}\mathrm{M}_{\odot}$, see Fig.11,

Vogelsberger et al. (2014a) and Fig.3, Genel et al. (2014).

- The physical extent of galaxies can be a factor of a few larger than observed for $M_\star \lesssim 10^{10.7}\mathrm{M}_\odot$ (see Fig. 9 in Snyder et al., 2015).
- The galaxy color distribution deviates from observations observations in that it does not exhibit a clear bimodality between red and blue galaxies, and the green-valley and the blue cloud appear over populated with respect to to the red sequence (especially for $M_\star \gtrsim 10^{10}\mathrm{M}_\odot$ (see Fig.14 in Vogelsberger et al., 2014a).
- About 10 percent of disk galaxies in the mass range $M_\star \sim 10^{10.5-11}\mathrm{M}_\odot$ at $z = 0$ exhibit strong stellar and gaseous ring-like features, and appear as an additional sub-population in the $G_{\mathrm{ini}} - M_{20}$ plane (see Fig. 5 in Snyder et al., 2015); such features appear to be even more frequent at higher redshifts. Via fragmentation, stellar rings may give rise to spurious stellar clumps that the SUBFIND algorithm identifies as subhalos but whose origin and existence is not necessarily physically well motivated (see also below).
- The total gas within $R_{500\mathrm{c}}$ is underestimated at late times by a factor 3-10 in halos with $M_{500\mathrm{c}} \sim 10^{13-14}\mathrm{M}_\odot$, because of the too violent operation mode of the Illustris radio-mode feedback (see Fig. 10 in Genel et al., 2014).
- For similar reasons, the bolometric X-ray luminosity in the hot coronae of elliptical galaxies is by many factors lower than in spiral galaxies, contradicting observational constraints (see Section 5.2 of Bogdan et al., 2015); and the predictions for the Sunyaev-Zel'dovich signals from Illustris clusters are not reliable (Popa et al. 2015, in prep).

For some items of this list we have intentionally omitted more specific quantifications of the tensions with observations for two reasons: on the one side, not all observational results are in agreement among each other, making quantitative statements necessarily partial; on the other side, excruciating care is necessary to properly map simulated variables into observationally-derived quantities. For example, we notice that the adopted low star-formation density threshold value and the low thermal energy content of galactic winds may be the cause for spurious star-formation in the circumgalactic medium around Milky Way-like galaxies, at large distances from the natural, dense sites of star formation activity (i.e. disks, see Marinacci et al. (2014a)). However, no observational data are available to properly quantify such phenomenon. Similarly, the impact of the AGN feedback on the dark-matter distribution within Illustris halos might be overestimated, but direct observational constraints are lacking.

Furthermore, while a first analysis of the stellar ages of Illustris galaxies seemed to reveal an overestimation of the predicted stellar ages for $M_\star \lesssim 10^{10.5}\mathrm{M}_\odot$ galaxies (see Fig. 25, Vogelsberger et al. (2014a)), we have now recognized that such a comparison to observations is rather inconclusive, as the shape of the age-mass relation of galaxies strongly depends, in the first place, on whether stellar ages are measured by mass- or light- weighting.

To better inform which features of the simulations should be trusted when making science conclusions, note also the following points more directly related to numerical choices:

- In both the snapshots and halo catalogs, metallicity values should be used and interpreted with care. These depend on the underlying choices for stellar evolution

and metal enrichment, with tabulated yields being uncertain and continuously updated. Furthermore, no metallicity floor has been imposed to the output data, so that metallicities of a small fraction of gas and star elements adopt minuscule, unrealistic values. Any user should feel free to adopt the most convenient and appropriate metallicity floor.

- In the SUBFIND catalogs, relatively-low mass, stellar- or gas-dominated objects at small galactocentric distances from their host halos may be artifacts and should be considered with care. These may be the results of the fragmentation of aforementioned stellar rings in disk galaxies, and may appear as outliers in halos/galaxies scaling relations involving sizes, masses, metallicities and mass-to-light ratios.

- Low-mass BHs in relatively low-mass subhalos should also be considered with care, particularly those hosted in satellite subhalos of more massive galaxies or at low redshifts. Because spurious motions of BH particles are prevented by repositioning the BH on halo potential minimum, in some cases, low-mass BHs in satellite galaxies are repositioned on the central halo on artificially short timescales. These "empty" satellites may then be repopulated with new BH seeds, regardless of redshift. The vast majority of these late-forming, satellite-hosted seeds do not grow significantly before merging with the central BH, so the effects are largely confined to BHs with mass $< 10^6 M_\odot$.

5.6 Community Considerations

5.6.1 Citation

To support proper attribution, recognize the effort of individuals involved, and monitor ongoing usage and impact, we request the following. Any publication making use of data from the Illustris simulations should cite this release paper (Nelson et al. 2015c) as well as the original paper introducing the project (Vogelsberger et al., 2014b). Furthermore, extensive use of the data, or studies of galaxy properties and populations, should cite if appropriate Vogelsberger et al. (2014a) as well as Genel et al. (2014). Any investigation of the black hole population should cite if appropriate Sijacki et al. (2014).

Finally, use of any of the supplementary data products should include the relevant citation. A full and up to date list will be maintained on the Illustris website. At the time of publication, this includes use of the SubLink merger trees (Rodriguez-Gomez et al., 2015), the redshift zero synthetic stellar images (Torrey et al., 2015), the subsequently derived morphological parameters (Snyder et al., 2015), and the stellar angular momentum, circularity measurements, and axis ratios (Genel et al., 2015).

5.6.2 Collaboration and Contributions

The full snapshots of Illustris-1 are sufficiently large that it will be prohibitive for most users to acquire or store a large number. We note that transferring 1.5 TB (per snapshot) at 10 MB/s will take roughly 42 hours. As a result, projects which require access to the entire snapshot set may benefit from closer interaction with members of the Illustris

collaboration. In particular, many team members are open to more direct collaboration, which can include guest access to compute resources which are local to full copies of the data. We welcome ideas for joint projects, so long as they intersect with the interests of collaboration members and do not overlap with existing efforts. We suggest, practically, to contact the author(s) who have already published work using Illustris data in related scientific topics.[5]

We also welcome contributions to the data release. These can take the form of either analysis code, or computed data products.

- If you would like to develop an (expensive) analysis routine, we can run it against one or all simulations or snapshots. The resulting data can be made immediately public through the Illustris API. Alternatively, the resulting data can be made privately available until an initial publication is released, and then released publicly.
- If you would like to develop an (inexpensive, fast) analysis routine, we can integrate it into the Illustris API, such that it can be requested on demand for any object. In this case, analysis should be restricted to subhalo or halo particles, and take at most a few seconds.
- Finally, if you produce a data set derived from the Illustris simulations, and would like to make it publicly available, we can host and distribute it alongside the other supplementary data catalogs.

If you are interested in any of these options, or have other ideas, please get in touch.

[5]See `http://www.illustris-project.org/results/` for a list.

5.6.3 Future Data Releases

We anticipate release of additional data in the near future, for which further documentation will be provided online.

Rockstar and Consistent-Trees

We plan to release ROCKSTAR group catalogs and the CONSISTENT-TREES merger trees built upon them for the six Illustris boxes in the near future, and will provide further documentation at that time. These group catalogs can include a different subhalo population than identified with the SUBFIND algorithm, particularly during mergers. The algorithm used to construct the C-Trees also has fundamental differences to both LHALOTREE and SUBLINK, inserting 'ghost' nodes or modifying properties of existing nodes such that objects in the tree may not map 1-to-1 to the group catalogs from which they were constructed. The output format and structure also differ substantially from either of the two other trees.

These additional catalogs can provide a powerful comparison and consistency check for any scientific analysis. We also anticipate that some users will simply be more familiar with these outputs, or need them as inputs to other tools.

Additional Supplementary Data Catalogs

The $z = 0$ "stellar mocks" multi-band images are being generated for twelve additional snapshots of Illustris-1 at $0.5 < z < 9$. These will include two sets of mock images in 47 common filters, one observing galaxies redshifted to the appropriate epoch and

the other observing galaxies in their rest frame. In addition, we expect to add maps of mass, metallicity, gas and stellar velocity, and gas and stellar velocity dispersion in the same projections as these synthetic images. Subsequently, we will also release the non-parametric morphology catalogs for the high redshift galaxy populations.

We expect to release a mock strong lensing catalog, which includes properties of galaxies that most resemble the observed lenses in term of mass/velocity dispersion.The following properties will be available: the Einstein radius R_E, the projected and 3d radial profile slopes, dark matter fraction within R_E, central stellar velocity dispersion, anisotropic parameters, effective radius, Sersic index, light ellipticity and orientation. This data will be available at several redshifts from $z = 0$ to $z = 1$, assuming fiducial source redshifts.

Additional details on the black holes will be provided: high time resolution outputs of black hole properties, and enumeration of all black hole merger events. This data is new and independent from the snapshots.

The catalog related to stellar circularities will be extended to all snapshots of Illustris-1.

Stellar assembly and merger history catalogs will be released, including details such as in-situ/ex-situ fractions, stellar mass formed pre/post infall, number of major and minor mergers in different time intervals and time since recent merger events. This data will be available for all subhalos at all snapshots of Illustris-123.

Dark-matter halo catalogs at selected snapshots will be released including dark-matter density profiles fit parameters, fit-independent concentration estimates, halo formation times, and halo shapes.

Mock images and property catalogs of Illustris-1 stellar halos will be released, at a selection of snapshots between z=0 to z=2.

We plan to publish lightcone images, whereby we transform raw simulation data from all snapshots into self-consistent mock-observed survey fields, in HST and JWST filters.

Additional Simulations

Several smaller simulations related to Illustris have been discussed in previous papers, including a series of 25Mpc/h boxes with variations on the input feedback parameters. These can be released in the future if there is community interest. Ongoing and future projects, including higher resolution "zooms" of individual systems, as well as larger volumes, will also be released through this platform in the future.

API Functionality Expansion

There is significant room for the development of additional features in the web-based API. In particular, for (i) on-demand visualization tasks, (ii) on-demand analysis tasks, and (iii) client-side, browser based tools for data exploration and visualization. For example, (i) requesting an image of projected gas density for a given halo, (ii) requesting a power-law radial slope measurement of a stellar halo or best-fit NFW parameters, and (iii) an interactive 3D representation of the subhalos within a given halo. We welcome community input and direct contributions in any of these directions.

5.7 Summary and Conclusions

We have made publicly available all the simulated data associated with the Illustris project at the permanent URL:

- `http://www.illustris-project.org/data/`

The Illustris project includes a series of large-scale, cosmological simulations ideal for studying the formation and evolution of galaxies. The simulation suite consists of three runs at increasing resolution levels of the same $(106.5\ \mathrm{Mpc})^3$ cosmological volume, with and without baryonic physics included. The high-resolution simulations (Illustris-1 and Illustris-1-Dark) include several million gravitationally bound structures, and the $z = 0$ Illustris-1 volume contains $\sim$7000 well-resolved galaxies with stellar mass exceeding $10^{10}\mathrm{M}_{\odot}$. The galaxies sampled in this volume span a range of environments and formation histories, allowing for a wide range of science topics to be addressed using the simulation data. For all six realizations, we are releasing the following data products:

- the raw snapshots at all 136 available redshifts down to $z = 0$;
- the friends-of-friends and SUBFIND halo/galaxy catalogs at the same 136 available redshifts down to $z = 0$;
- the SUBLINK and LHALOTREE merger trees;
- the raw snapshots of four sub regions of the full volume, for each full physics run, output with significantly higher time frequency;
- supplementary data catalogs currently focused on properties of the Illustris-1 $z = 0$ galaxy population.

We anticipate release of additional data post-processed products in the near future, for which further documentation will be provided online. Although the total data volume associated with the Illustris project which is presently released is sizeable, ∼265 TB, we have made a significant effort to make this data accessible to the broader community. Specifically, the simulation data is available either via direct download of the raw files or via web-based API queries for common search, extraction, and analysis tasks. Extensive documentation on the format and contents of all released datasets is included both in this paper as well as online, where it will be progressively extended. Additionally, we have made basic I/O scripts and starting examples in IDL, Python, and Matlab available to enable users to analyze and work with the raw data. The resulting data products have widespread applications and provide a powerful tool for the interpretation of extragalactic observations. By making this data publicly available, we hope to maximize the scientific return from the considerable computational resources invested into running the Illustris simulation suite.

Appendix A: Technical Details

The vast majority of past simulation data releases have made use of relational database systems (i.e. MySQL, PostgreSQL, or commercial options) as the primary mechanism for user interaction as well as data distribution. Following the impressive success of the SDSS Skyserver, and starting notably for theory with the Millenium simulation database, users were invited to write and submit raw SQL queries to these databases. Most non-trivial tasks require complex queries which can join multiple tables together across foreign key relations, as well as an awareness of the indexing systems and their use. The power of the query language is offset for most non-experts by the unusual approach, which requires abandoning common methods for the local analysis of astronomical data sets, most notably by the writing of small code snippets, which can have loops and if-else type decision branches.

In the present effort we have made use of a relational database in the usual way, to hold the full outputs of the group finding algorithms (and not the raw particle data). We exported all group catalogs into the database, with one InnoDB table per run. Each table is partitioned on snapshot number, and has only a single composite B-Tree index on (snapshot,subhalo_id). The goal was to enable rapid search over arbitrary parameter combinations, primarily at a single snapshot. Therefore we did not adopt a merger tree centric ordering (as in Lemson & Virgo Consortium, 2006). In fact, by releasing multiple merger trees we also wished to emphasize the fact that there is no ground truth for the merger history of any object, where by definition such an ordering is useful for only one tree. Our snapshot ordering scheme suffers the same limitation – it is specifically

reflective of the SUBFIND group finder employed on-the-fly.

Instead of allowing users to submit SQL queries, we hide the existence of the database behind an API facade. This implies that each piece of functionality must be exposed through an API endpoint. In addition to the subhalo catalogs, we export all relevant metadata for simulation runs and snapshots into the databases, which enables the overall API structure.

The Illustris API is based on a representational state transfer architecture (REST, see Fielding, 2000). Requests and responses are transferred over HTTP, and GET is the only supported request verb (meaning that the system is read-only from the user perspective). Individual resources, or "endpoints", are identified by their unique URL. The system is stateless, meaning that each request is independent of any previous requests, and must include sufficient information to handle it. The default response type is JSON, a human-readable text format which can be parsed by all modern languages and clients. Because the primary purpose of the API is to serve scientific data sets, HDF5 is chosen as the default response type for binary data. For many resources, the response can be requested in any number of supported formats, which currently include CSV, JSON, HDF5, FITS, PNG, and plain text. The ability of a client to navigate the API and discover available resources is crucial. We generally adopted the principle of Hypermedia as the Engine of Application State (HATEOAS), meaning that users can discover and request resources in the API without needing to know its structure in advance. This is achieved by stating all relationships between objects in terms of the absolute URL at which each object can be found. For example, the final code listing in Appendix B below uses the hyperlinked relationship from a given subhalo to its descendant at a different redshift to walk through a merger tree.

In general, we aim to support only relatively light queries, which the user should anticipate will complete in a few seconds at most. There is no queued or batch query system, where long running queries can be submitted and their progress periodically polled. There is no per-user remote storage (e.g. "MyDB"). Together, this greatly simplifies the design of the system, with the implied thought that the typical user workflow will be to download and process specific datasets on their local machine. We have elected to design a system with a split between a front end, which is exposed to the user, and back end resources. This allows for the two to be in different locations, and for multiple back ends to be supported. In particular, our division is such that the front end handles (i) the Illustris website itself, including (ii) all user details: registration, management, authentication. (iii) All statistics and record keeping. (iv) The full API structure, and responding to API requests at all endpoints. (v) The database, holding both simulation metadata, and the group catalogs. Currently only one back end is in use, and consists of a public-facing machine on the same local network as the data. It handles:

- Serving raw data files. In this case, several distributed filesystems are locally mounted. Requests are translated into the appropriate system path, and given back to Apache to serve directly via XSendFile.
- Extract subsets of data files are also served. In this case, the pre-calculated offsets are used in order to only read the requested data from disk. This data is either read into a memory structure in the format requested by the client, or subsequently converted to the requested format. In particular, binary extractions from HDF5 containers are read into an in-memory HDF5 "image". The raw bytestream of this

image is then transferred to the client from memory, such that no temporary copy of the data subset need be saved.

The back end is stateless, has no database or persistent local storage of any kind, and no knowledge of the user making each request. In order to provide authentication, which forms the basis of usage monitoring, permission levels, bandwidth throttling and rate limits, the following steps are taken:

1. The user makes a request to the API on the front end, including their API-Key.
2. The front end authenticates (verifies their identity) and authorizes (checks sufficient permissions) the user.
3. The front end verifies the validity of the request, including the existence of the requested data.
4. If the request can be satisfied from data available in the front end database (e.g. simulation metadata, subhalo fields), the response is returned directly.
5. If the request requires data from the back end, the appropriate path (URL) is constructed.
6. The front end generates a hash-based message authentication code (HMAC) by concatenating a time-based one-time password (TOTP, see RFC 6238) with a pre-shared secret key and the request URL itself.
7. This token is appended to the back end request URL, which is then sent to the client with a REDIRECT request.

8. The client makes the request to the back end.

9. The back end verifies the request by computing the current TOTP and constructing the same hash using the pre-shared secret key.

The use of the time-varying key means that each request to the back end is attached to a specific request from a specific user. The advantage of this approach is that the front end can redirect clients to data at any back end resource while avoiding the bandwidth burden of making the request itself and forwarding the data on to the client.

At the software level, the Illustris data release makes use of a large number of software projects. It is realized on a common open source software stack: CentOS, Apache, and MySQL. On the front end, Python is used to handle all dynamic web content through the Django web framework with several packages including the Django REST framework. The website uses the Bootstrap framework, the jQuery javascript library, MathJax and pygments rendering. The Explorer interface uses the Leaflet tile map engine, as well as the two-dimensional R-Tree indexing capabilities in MySQL to locate subhalos and black holes inside in the visible bounding box. Currently there is no support for spatial indexing in higher dimensions, so using the database for 3D (periodic) distance queries would require a custom solution (Lemson et al., 2011). Client-side visualizations, currently for the merger trees, use the d3 javascript data visualization library, and three.js for WebGL. On the back end, the HDF5 library with the h5py, numpy, and fitsio Python packages provide the bulk of the data interaction layer.

This back end is currently only focused on storage and data delivery, and we do not yet have any system in place to allow temporary, guest access to compute resources which are local to the data itself. However, we envision that this could change

in the future. The data delivery portal has access to the compute resources of the cluster, and instead of defining specific, pre-written analysis functions, we would like to provide a familiar environment for the execution of arbitrary user programs. There has been significant recent development related to remote, multi-user, rich interfaces to computational kernels. In particular, the Jupyter notebook environment (previously called IPython, Pérez & Granger (2007)) can be spawned, on demand, inside sand-boxed Docker instances, through a web-based portal with authentication provided by the existing user registration system. This means that users could develop analysis routines in any language (Jupyter support includes Python, IDL, Matlab, Julia, and many others) and execute them, in the same interface, on the remote cluster.

Finally, the read-only, highly structured nature of simulation output motivates more efficient approaches. In particular, bitmap indexing over HDF5 as in FastQuery (Chou et al., 2011; Byna et al., 2012) together with a SQL-like query layer (Wang et al., 2013). When these technologies are slightly more mature, the need to place a copy of raw simulation data into a relational database can be foregone. Instead, the DB can be used only to handle meta-data, and fast indexed search and queries can be made directly against structured binary data on disk. We anticipate that such an approach might be relevant for future data release efforts.

Appendix B: API Code Examples

To be explicit by way of example, the following are absolute URLs for the Illustris API covering some of its functionality, where the type of the request should be clear from the preceding documentation.

- http://www.illustris-project.org/api/Illustris-2/
- http://www.illustris-project.org/api/Illustris-2/snapshots/68/
- http://www.illustris-project.org/api/Illustris-1/snapshots/135/subhalos/73664/
- http://www.illustris-project.org/api/Illustris-1/snapshots/135/subhalos/73664/stellar_mocks/broadband.fits
- http://www.illustris-project.org/api/Illustris-1/snapshots/135/subhalos/73664/stellar_mocks/sed.txt
- http://www.illustris-project.org/api/Illustris-1/snapshots/80/halos/523312/cutout.hdf5?dm=Coordinates&gas=all
- http://www.illustris-project.org/api/Illustris-3/snapshots/135/subhalos?mass__gt=10.0&mass__lt=20.0
- http://www.illustris-project.org/api/Illustris-2/snapshots/68/subhalos/50000/sublink/full.hdf5
- http://www.illustris-project.org/api/Illustris-2/snapshots/68/subhalos/50000/sublink/mpb.json
- http://www.illustris-project.org/api/Illustris-1/files/groupcat-135.5.hdf5
- http://www.illustris-project.org/api/Illustris-2/files/snapshot-135.10.hdf5
- http://www.illustris-project.org/api/Illustris-2/files/snapshot-135.3.hdf5?dm=all
- http://www.illustris-project.org/api/Illustris-3/files/sublink.2.hdf5

In the online documentation we provide a complete getting started guide for the web-based API, as well as a cookbook of common tasks, in Python, IDL, and Matlab. Here we include just four examples taken from that documentation, and only in Python, to give a flavor of the approach. The task numbers are taken from the online version.

Task 0: First, we define a helper function, to make the HTTP response, and check for errors. If the response is JSON, automatically parse it. If the response is binary data, automatically save it to a file.

```
>>> def get(path, params=None):
>>>   # make HTTP GET request to path
>>>   headers = {"api-key":"INSERT_API_KEY_HERE"}
>>>   r = requests.get(path, params=params, headers=headers)
>>>
>>>   # raise exception if response code is not HTTP SUCCESS (200)
>>>   r.raise_for_status()
>>>
>>>   if r.headers['content-type'] == 'application/json':
>>>     return r.json() # parse json responses automatically
>>>
>>>   if 'content-disposition' in r.headers:
>>>     filename = r.headers['content-disposition'].split("filename=")[1]
>>>     with open(filename, 'wb') as f:
>>>       f.write(r.content)
>>>     return filename # return the filename string
```

Task 1: For Illustris-1 at $z = 0$, get all the fields available for the subhalo with id=0 and print its total mass and stellar half mass radius.

```
>>> url = "http://www.illustris-project.org/api/"
>>> url += "Illustris-1/snapshots/135/subhalos/0/"
>>> r = get(url)
>>> r['mass']
22174.8

>>> r['halfmassrad_stars']
12.395
```

Task 2: For Illustris-1 at $z = 2$, search for all subhalos with total mass $10^{11.9}\mathrm{M}_\odot < M < 10^{12.1}\mathrm{M}_\odot$, print the number returned, and the SUBFIND IDs of the first five results.

```
>>> # first convert log solar masses into group catalog units
>>> mass_min = 10**11.9 / 1e10 * 0.704
>>> mass_max = 10**12.1 / 1e10 * 0.704
>>>
>>> params = {'mass__gt':mass_min, 'mass__lt':mass_max}
>>>
>>> # make the request
>>> url = "http://www.illustris-project.org/api/"
>>> url += "Illustris-1/snapshots/z=2/subhalos/"
>>> subhalos = get(url, params)
>>> subhalos['count']
550

>>> ids = [ subhalos['results'][i]['id'] for i in range(5) ]
>>> ids
[1, 1352, 5525, 6574, 12718]
```

Task 11: Download the entire Illustris-1 $z = 0$ snapshot including *only the positions, masses, and metallicities of stars* (in the form of 512 HDF5 files). In this example, since we only need these three fields for stars only, we can reduce the download and storage size from ∼1.5 TB to ∼17 GB.

```
>>> base_url = "http://www.illustris-project.org/api/Illustris-1/"
>>> sim_metadata = get(base_url)
>>> params = {'stars':'Coordinates,Masses,GFM_Metallicity'}
>>>
>>> for i in range(sim_metadata['num_files_snapshot']):
>>>     file_url = base_url + "files/snapshot-135." + str(i) + ".hdf5"
>>>     saved_filename = get(file_url, params)
>>>     print saved_filename
```

Task 8: For Illustris-1 at $z = 2$, for five specific SUBFIND IDs (from above: 1, 1352, 5525, 6574, 12718), locate the $z = 0$ descendant of each by using the API to walk down the SUBLINK descendant links.

```
>>> ids = [1, 1352, 5525, 6574, 12718]
>>> z0_desc_ids = [-1]*len(ids)
>>>
>>> for i,id in enumerate(ids):
>>>     start_url = "http://www.illustris-project.org/api/"
>>>     start_url += "Illustris-1/snapshots/z=2/subhalos/"
>>>     start_url += str(id)
>>>     sub = get(start_url)
>>>
>>>     while sub['desc_sfid'] != -1:
>>>         # request the full subhalo details through the sublink URL
>>>         sub = get(sub['related']['sublink_descendant'])
>>>         if sub['snap'] == 135:
>>>             z0_desc_ids[i] = sub['id']
>>>
>>>     # note: possible that descendant branch did not reach z=0
>>>     if z0_desc_ids[i] >= 0:
>>>         print 'Descendant of '+str(id)+' is '+str(z0_desc_ids[i])

Descendant of 1 at z=0 is 30465
Descendant of 1352 at z=0 is 41396
Descendant of 5525 at z=0 is 99148
Descendant of 6574 at z=0 is 51811
Descendant of 12718 at z=0 is 194303
```

Chapter 6

Conclusions and Future Directions

Our understanding of the processes which drive the formation and evolution of galaxies has been propelled forward through the development of accurate and comprehensive numerical models. At the same time, observations not only of the galaxies themselves, but also of the gaseous reservoirs surrounding them, have increasingly offered new insights and, perhaps more importantly, unsolved puzzles. It is only through the continuous interplay between simulations and observations that we can hope to arrive at a comprehensive and *correct* theory for the complex cycle of baryons in to, and out of, galaxies across cosmic time.

The work presented herein provides the first application of the moving-mesh technique to the problem of cosmological gas accretion. In Chapter 2 we demonstrated that conclusions as to the importance of hot versus cold mode accretion are qualitatively changed using this more robust numerical method. The Monte Carlo tracer particle technique was developed and applied to quantify the thermal history of accreting gas. This tracer technique was subsequently used in Chapter 3 to address the impact of

feedback. We found that the energy input from star formation and AGN feedback channels suppresses smooth accretion direct from the intergalactic medium, setting up a strong galactic-scale fountain which provides a substantial late time accretion rate of recycled gas. Chapter 4 transitioned from cosmological volumes to a series of zoom simulations of individual galaxy halos, thereby affording significantly higher spatial resolution. Studying the halo gas structure in detail we demonstrated the significant angular variability of features such as the virial shock, as well as the increasing complexity of the multi-phase interaction between filamentary inflows and the quasi-static hot halo with better resolution. Finally, Chapter 5 presented our effort to publicly release the entirety of the simulation data produced as part of the Illustris project. We anticipate that this will facilitate broad and informative comparisons between state of the art theoretical models and the wealth of available observational data.

Numerical simulations studying the formation and evolution of galaxies remain challenging. The shear diversity of physical processes at play, together with their inherent complexities and non-linear interactions, implies that many significant approximations and assumptions are required, and always will be. To date, most cosmological simulations, including those we present in this work, ignore several areas of fundamental physics. These include the impact of magnetic fields, any explicit treatment of radiative transfer or radiation fields, non-thermal sources of pressure including cosmic rays, non-equilibrium chemistry, molecules and their effects, and many others. In addition, finite computational resources imply a finite spatial resolution. Therefore, many of the most important processes occur below the resolution limit of our simulations, and always will. Together, these two inescapable truths of physical complexity and finite resolution combine to delineate a daunting task, and yet improvements and refinements in the next

generation of simulations will undeniably continue to make progress in both areas.

Moving forward, we note several promising directions to pursue. First, to explore the impact of local radiation fields on the state of cosmic gas. Modern cosmological simulations almost uniformly neglect the production and transport of spatially inhomogeneous radiation fields (from stars, black holes, gas, radiative shocks, and unresolved backgrounds) and their impact on the thermodynamic state and cooling properties of gas. We would like to develop a robust and physically accurate model for local radiation sources, and the impact of non-equilibrium ionization and cooling effects in circumgalactic gas. This would necessarily leverage methods for radiative transport through the Voronoi-based computational mesh. Our motivation is two-fold. First, as opposed to most existing models for feedback, which expel gas from the galaxy after it has already accreted, these radiation-based processes can potentially act as a "preventive" source of feedback, stopping halo gas from cooling in the first place. An accounting of such preventive feedback may be fundamentally required to simultaneously reproduce the observed amount of stellar mass in dark matter halos as well as secondary properties of those halos, such as their x-ray luminosities. Second, a more accurate treatment of radiation is required to determine the chemical and kinematical signatures of gas inflows and outflows in the circumgalactic regime, in order to make direct comparison with observational data. Recent observations of the gas content of the circumgalactic medium have revealed the ubiquitous existence of significant amounts of mass in cold, metal-enriched gas clouds, which appear to survive significantly longer than expected in the halos of massive galaxies. This remains a point of tension for all simulations to date, which do not produce such large covering fractions of low ionization state gas.

A second clear direction is the improvement of the current feedback models. Existing

models for energy input from AGN feedback as well as galactic-scale outflows driven by feedback from star formation are crude. Ongoing work on the shortcomings identified in the Illustris physics models motivates refinements and improvements in the physics. Perhaps more than any other aspect of the simulations, significant improvements to the physical models and numerical implementations of feedback are not only warranted, but likely to have the largest impact on the simulated circumgalactic gas and its observational signatures. Details of the interaction of an outflowing wind with the halo gas, how far wind material can travel, how much it can enrich the IGM, and its recycling properties – the balance between inflow and outflow and the time scale of reincorporation – may all be quite sensitive to the models. In reality, outflows are observed to be multi-phase, with gas distributed at different densities, temperatures, metallicities, and velocities. None of these complications are considered at present. We can improve the sophistication of feedback modeling either by using constraints provided by observations, or by conducting smaller scale simulations where these processes can be directly resolved. Both are promising avenues, and the high-resolution zooms of individual halos, together with their comparison to observations, provide an ideal test bed for the implementation of new, more realistic treatments of galactic winds.

The third clear direction is the need to solve the issue of limited resolution in the halo. If the observed absorption in cold metal ions is indeed caused by structures significantly below the resolution limit of current simulations, any direct comparison between the two may be premature. The constraints on the physical size of such absorbers range from sub-parsec to hundreds of parsecs. Obtaining $\sim 100\,\mathrm{pc}$ hydrodynamic resolution at the virial radius by simply running higher resolution simulations, even of isolated halos, will be prohibitively expensive. However, new numerical techniques may enable novel ways

to focus resolution within the halo regime. If we can avoid clustering the computational effort in the densest regions of space – within galaxies themselves – much higher spatial resolutions can be achieved in lower density media. In theory, the refinement criterion in grid based simulations can be based on any localized properties in the simulation. In the moving mesh technique, for instance, we can enforce a target gas cell mass criterion, through adaptive refinement and de-refinement of Voronoi cells, as a function of distance from the center of a halo. Alternatively, sub-resolution models for multiphase gas outside of galaxies may need to be developed, in analogy to those used to represent the unresolved phases of dense interstellar gas.

On one hand, the speed of computers, the size of simulations, the quality and diversity of observational data are all continually increasing, and to claim that *now* somehow represents a unique point in time is somewhat narrow in outlook. And yet, *now* does seem exactly the point where, for the first time, numerical realizations of the currently accepted standard model of cosmology have succeeded in reproducing many fundamental features of observed galaxy populations, from the early universe to the present day. The next generation of cosmological simulations promises to address many of the outstanding questions and further our understanding of galaxies, and of the universe as a whole. Although it remains to be seen what discoveries may yet surprise, the history of our field suggests that some surprises are, rather likely, waiting to be discovered.

References

Aarseth, S. J. 1963, MNRAS, 126, 223

Abadi, M. G., Navarro, J. F., Steinmetz, M., & Eke, V. R. 2003, ApJ, 591, 499

Agertz, O., & Kravtsov, A. V. 2014, ArXiv e-prints

Agertz, O., Kravtsov, A. V., Leitner, S. N., & Gnedin, N. Y. 2013, ApJ, 770, 25

Agertz, O., Teyssier, R., & Moore, B. 2009, MNRAS, 397, L64

Agertz, O., et al. 2007, MNRAS, 380, 963

Aragon-Calvo, M. A., Neyrinck, M. C., & Silk, J. 2014, ArXiv e-prints

Aumer, M., White, S. D. M., Naab, T., & Scannapieco, C. 2013, MNRAS, 434, 3142

Bachmann, L. K., Dolag, K., Hirschmann, M., Almudena Prieto, M., & Remus, R.-S. 2014, ArXiv e-prints

Balsara, D. S. 1995, Journal of Computational Physics, 121, 357

Barkana, R., & Loeb, A. 2001, Phys. Rep., 349, 125

Barnes, J., & Efstathiou, G. 1987, ApJ, 319, 575

Barnes, J., & Hut, P. 1986, Nature, 324, 446

Barro, G., et al. 2013, ApJ, 765, 104

Bauer, A., & Springel, V. 2012, MNRAS, 3102

Bauer, A., Springel, V., Vogelsberger, M., Genel, S., Torrey, P., Sijacki, D., Nelson, D., & Hernquist, L. 2015, ArXiv e-prints

Behroozi, P. S., Wechsler, R. H., Wu, H.-Y., Busha, M. T., Klypin, A. A., & Primack, J. R. 2013, ApJ, 763, 18

Bellovary, J., Brooks, A., Volonteri, M., Governato, F., Quinn, T., & Wadsley, J. 2013, ApJ, 779, 136

Bernyk, M., et al. 2014, ArXiv e-prints

Bertin, E., Pillay, R., & Marmo, C. 2015, Astronomy and Computing, 10, 43

Besla, G., Kallivayalil, N., Hernquist, L., van der Marel, R. P., Cox, T. J., & Kereš, D. 2010, ApJ, 721, L97

—. 2012, MNRAS, 421, 2109

Bird, S., Haehnelt, M., Neeleman, M., Genel, S., Vogelsberger, M., & Hernquist, L. 2015, MNRAS, 447, 1834

Bird, S., Vogelsberger, M., Haehnelt, M., Sijacki, D., Genel, S., Torrey, P., Springel, V., & Hernquist, L. 2014, MNRAS, 445, 2313

Bird, S., Vogelsberger, M., Sijacki, D., Zaldarriaga, M., Springel, V., & Hernquist, L. 2013, MNRAS, 429, 3341

Birnboim, Y., & Dekel, A. 2003a, MNRAS, 345, 349

—. 2003b, MNRAS, 345, 349

Birnboim, Y., Dekel, A., & Neistein, E. 2007, MNRAS, 380, 339

Bogdan, A., et al. 2015, ArXiv e-prints

Booth, C. M., Agertz, O., Kravtsov, A. V., & Gnedin, N. Y. 2013, ApJ, 777, L16

Bower, R. G., Benson, A. J., Malbon, R., Helly, J. C., Frenk, C. S., Baugh, C. M., Cole, S., & Lacey, C. G. 2006, MNRAS, 370, 645

Boylan-Kolchin, M., Springel, V., White, S. D. M., Jenkins, A., & Lemson, G. 2009, MNRAS, 398, 1150

Brammer, G. B., et al. 2012, ApJS, 200, 13

Brook, C. B., Stinson, G., Gibson, B. K., Roškar, R., Wadsley, J., & Quinn, T. 2012, MNRAS, 419, 771

Brook, C. B., Stinson, G., Gibson, B. K., Shen, S., Macciò, A. V., Obreja, A., Wadsley, J., & Quinn, T. 2014, MNRAS, 443, 3809

Brooks, A. M., Governato, F., Quinn, T., Brook, C. B., & Wadsley, J. 2009, ApJ, 694, 396

Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000

Bryan, G. L., & Norman, M. L. 1998, ApJ, 495, 80

Bullock, J. S., Dekel, A., Kolatt, T. S., Kravtsov, A. V., Klypin, A. A., Porciani, C., & Primack, J. R. 2001, ApJ, 555, 240

Buser, R. 1978, A&A, 62, 411

Byna, S., et al. 2012, in Proceedings of the International Conference on High Performance Computing, Networking, Storage and Analysis, SC '12 (Los Alamitos, CA, USA: IEEE Computer Society Press), 59:1–59:12

Cen, R. 1992, ApJS, 78, 341

—. 2014, ApJ, 789, L21

Cen, R., Miralda-Escudé, J., Ostriker, J. P., & Rauch, M. 1994, ApJ, 437, L9

Ceverino, D., Klypin, A., Klimek, E. S., Trujillo-Gomez, S., Churchill, C. W., Primack, J., & Dekel, A. 2014, MNRAS, 442, 1545

Chabrier, G. 2003, PASP, 115, 763

Chandrasekhar, S. 1943, ApJ, 97, 255

Chou, J., Wu, K., & Prabhat. 2011, in Cluster Computing (CLUSTER), 2011 IEEE International Conference on, 455–464

Churchill, C. W., Vander Vliet, J. R., Trujillo-Gomez, S., Kacprzak, G. G., & Klypin, A. 2014, ArXiv e-prints

Ciotti, L., & Ostriker, J. P. 2007, ApJ, 665, 1038

Colella, P., & Woodward, P. R. 1984, Journal of Computational Physics, 54, 174

Cooley, J. W., & Tukey, J. W. 1965, Math. Comput., 19, 297

Costa, T., Sijacki, D., & Haehnelt, M. G. 2014, MNRAS, 444, 2355

Creasey, P., Theuns, T., Bower, R. G., & Lacey, C. G. 2011, MNRAS, 415, 3706

Crighton, N. H. M., Hennawi, J. F., Simcoe, R. A., Cooksey, K. L., Murphy, M. T., Fumagalli, M., Prochaska, J. X., & Shanks, T. 2015, MNRAS, 446, 18

Crocce, M., Fosalba, P., Castander, F. J., & Gaztañaga, E. 2010, MNRAS, 403, 1353

Croton, D. J., et al. 2006, MNRAS, 365, 11

Dahlen, T., et al. 2004, ApJ, 613, 189

Danovich, M., Dekel, A., Hahn, O., Ceverino, D., & Primack, J. 2014, preprint, (arXiv:1407.7129)

Danovich, M., Dekel, A., Hahn, O., & Teyssier, R. 2012, MNRAS, 422, 1732

Davé, R., Finlator, K., & Oppenheimer, B. D. 2012, MNRAS, 421, 98

Davis, M., Efstathiou, G., Frenk, C. S., & White, S. D. M. 1985, ApJ, 292, 371

Davis, M., et al. 2003, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 4834, Discoveries and Research Prospects from 6- to 10-Meter-Class Telescopes II, ed. P. Guhathakurta, 161–172

De Lucia, G., & Blaizot, J. 2007, MNRAS, 375, 2

Dehnen, W. 2000, ApJ, 536, L39

Dekel, A., & Birnboim, Y. 2006, MNRAS, 368, 2

Dekel, A., Zolotov, A., Tweed, D., Cacciato, M., Ceverino, D., & Primack, J. R. 2013, MNRAS, 435, 999

Dekel, A., et al. 2009, Nature, 457, 451

Di Matteo, T., Colberg, J., Springel, V., Hernquist, L., & Sijacki, D. 2008, ApJ, 676, 33

Di Matteo, T., Springel, V., & Hernquist, L. 2005, Nature, 433, 604

Dolag, K., Borgani, S., Murante, G., & Springel, V. 2009, MNRAS, 399, 497

Dubois, Y., Pichon, C., Devriendt, J., Silk, J., Haehnelt, M., Kimm, T., & Slyz, A. 2013, MNRAS, 428, 2885

Dubois, Y., Pichon, C., Haehnelt, M., Kimm, T., Slyz, A., Devriendt, J., & Pogosyan, D. 2012, MNRAS, 423, 3616

Dubois, Y., et al. 2014, MNRAS, 444, 1453

Evrard, A. E., et al. 2002, ApJ, 573, 7

Faucher-Giguere, C.-A., Hopkins, P. F., Keres, D., Muratov, A. L., Quataert, E., & Murray, N. 2014, ArXiv e-prints

Faucher-Giguère, C.-A., Kereš, D., & Ma, C.-P. 2011, MNRAS, 417, 2982

Faucher-Giguère, C.-A., Lidz, A., Zaldarriaga, M., & Hernquist, L. 2009, ApJ, 703, 1416

Feldmann, R., & Mayer, L. 2015, MNRAS, 446, 1939

Feng, Y., Di Matteo, T., Croft, R., & Khandai, N. 2014, MNRAS, 440, 1865

Ferland, G. J., Korista, K. T., Verner, D. A., Ferguson, J. W., Kingdon, J. B., & Verner, E. M. 1998, PASP, 110, 761

Fielding, R. T. 2000, PhD thesis, aAI9980887

Ford, A. B., Davé, R., Oppenheimer, B. D., Katz, N., Kollmeier, J. A., Thompson, R., & Weinberg, D. H. 2014, MNRAS, 444, 1260

Ford, A. B., Oppenheimer, B. D., Davé, R., Katz, N., Kollmeier, J. A., & Weinberg, D. H. 2013, MNRAS, 432, 89

Freeth, T., et al. 2006, Nature, 444, 587

Frenk, C. S., White, S. D. M., & Davis, M. 1983, ApJ, 271, 417

Frenk, C. S., et al. 1999, ApJ, 525, 554

Fumagalli, M., Hennawi, J. F., Prochaska, J. X., Kasen, D., Dekel, A., Ceverino, D., & Primack, J. 2014, ApJ, 780, 74

Gabor, J. M., & Bournaud, F. 2014, MNRAS, 437, L56

Gabor, J. M., & Davé, R. 2012, MNRAS, 427, 1816

Genel, S., Dekel, A., & Cacciato, M. 2012, MNRAS, 425, 788

Genel, S., Fall, S. M., Hernquist, L., Vogelsberger, M., Snyder, G. F., Rodriguez-Gomez, V., Sijacki, D., & Springel, V. 2015, ArXiv e-prints

Genel, S., Genzel, R., Bouché, N., Naab, T., & Sternberg, A. 2009, ApJ, 701, 2002

Genel, S., Vogelsberger, M., Nelson, D., Sijacki, D., Springel, V., & Hernquist, L. 2013, MNRAS, 435, 1426

Genel, S., et al. 2014, MNRAS, 445, 175

George, M. R., Fabian, A. C., Sanders, J. S., Young, A. J., & Russell, H. R. 2009, MNRAS, 395, 657

Gingold, R. A., & Monaghan, J. J. 1977, MNRAS, 181, 375

Godunov, S. K. 1959, Matematicheskii Sbornik, 89, 271

Górski, K. M., Hivon, E., Banday, A. J., Wandelt, B. D., Hansen, F. K., Reinecke, M., & Bartelmann, M. 2005, ApJ, 622, 759

Gottlöber, S., & Yepes, G. 2007, ApJ, 664, 117

Gray, J., Szalay, A. S., Thakar, A. R., Kunszt, P. Z., Stoughton, C., Slutz, D., & vandenBerg, J. 2002, eprint arXiv:cs/0202014

Greengard, L., & Rokhlin, V. 1987, Journal of Computational Physics, 73, 325

Greggio, L. 2005, A&A, 441, 1055

Grogin, N. A., et al. 2011, ApJS, 197, 35

Guo, Q., et al. 2011, MNRAS, 413, 101

Hahn, O., & Abel, T. 2011, MNRAS, 415, 2101

Hahn, O., & Angulo, R. 2015, preprint, (arXiv:1501.01959)

Hahn, O., Teyssier, R., & Carollo, C. M. 2010, MNRAS, 405, 274

Hanasz, M., Lesch, H., Naab, T., Gawryszczak, A., Kowalik, K., & Wóltański, D. 2013, ApJ, 777, L38

Hennawi, J. F., et al. 2006, ApJ, 651, 61

Henriques, B., White, S., Thomas, P., Angulo, R., Guo, Q., Lemson, G., Springel, V., & Overzier, R. 2014, ArXiv e-prints

Hernquist, L. 1987, ApJS, 64, 715

Hernquist, L., & Katz, N. 1989a, ApJS, 70, 419

—. 1989b, ApJS, 70, 419

Hobbs, A., Read, J., Power, C., & Cole, D. 2012, preprint, (arXiv:1207.3814)

Hockney, R. W., & Eastwood, J. W. 1981, Computer Simulation Using Particles

Holmberg, E. 1941, ApJ, 94, 385

Hopkins, P. F., Kereš, D., Oñorbe, J., Faucher-Giguère, C.-A., Quataert, E., Murray, N., & Bullock, J. S. 2014, MNRAS, 445, 581

Hummels, C. B., Bryan, G. L., Smith, B. D., & Turk, M. J. 2013, MNRAS, 430, 1548

Hunt, J. C. R. 1998, Annual Review of Fluid Mechanics, 30, D13

Iapichino, L., Adamek, J., Schmidt, W., & Niemeyer, J. C. 2008, MNRAS, 388, 1079

Joung, M. R., Putman, M. E., Bryan, G. L., Fernández, X., & Peek, J. E. G. 2012, ApJ, 759, 137

Kang, H., Ostriker, J. P., Cen, R., Ryu, D., Hernquist, L., Evrard, A. E., Bryan, G. L., & Norman, M. L. 1994, ApJ, 430, 83

Karakas, A. I. 2010, MNRAS, 403, 1413

Katz, N., & Gunn, J. E. 1991, ApJ, 377, 365

Katz, N., Hernquist, L., & Weinberg, D. H. 1992, ApJ, 399, L109

Katz, N., Keres, D., Dave, R., & Weinberg, D. H. 2003, in Astrophysics and Space Science Library, Vol. 281, The IGM/Galaxy Connection. The Distribution of Baryons at z=0, ed. J. L. Rosenberg & M. E. Putman, 185

Katz, N., Weinberg, D. H., & Hernquist, L. 1996, ApJS, 105, 19

Kereš, D., & Hernquist, L. 2009, ApJ, 700, L1

Kereš, D., Katz, N., Fardal, M., Davé, R., & Weinberg, D. H. 2009, MNRAS, 395, 160

Kereš, D., Katz, N., Weinberg, D. H., & Davé, R. 2005, MNRAS, 363, 2

Kereš, D., Vogelsberger, M., Sijacki, D., Springel, V., & Hernquist, L. 2012, MNRAS, 425, 2027

Keshet, U., Waxman, E., Loeb, A., Springel, V., & Hernquist, L. 2003, ApJ, 585, 128

Khandai, N., Di Matteo, T., Croft, R., Wilkins, S. M., Feng, Y., Tucker, E., DeGraf, C., & Liu, M.-S. 2014, preprint, (arXiv:1402.0888)

Kim, J., Park, C., Rossi, G., Lee, S. M., & Gott, III, J. R. 2011, Journal of Korean Astronomical Society, 44, 217

Klypin, A. A., & Shandarin, S. F. 1983, MNRAS, 204, 891

Klypin, A. A., Trujillo-Gomez, S., & Primack, J. 2011, ApJ, 740, 102

Lemson, G., Budavári, T., & Szalay, A. 2011, in Proceedings of the 23rd International Conference on Scientific and Statistical Database Management, SSDBM'11 (Berlin, Heidelberg: Springer-Verlag), 509–526

Lemson, G., & Virgo Consortium, t. 2006, ArXiv Astrophysics e-prints

Lemson, G., & Zuther, J. 2009, Mem. Soc. Astron. Italiana, 80, 342

Lemson, G., et al. 2014, ArXiv e-prints

LeVeque, R. J. 2002, Finite volume methods for hyperbolic problems, Vol. 31 (Cambridge university press)

Lewis, A., Challinor, A., & Lasenby, A. 2000, ApJ, 538, 473

Li, Y., & Bryan, G. L. 2014, ApJ, 789, 54

Lotz, J. M., Primack, J., & Madau, P. 2004, AJ, 128, 163

LSST Science Collaboration et al. 2009, ArXiv e-prints

Lu, Y., Kereš, D., Katz, N., Mo, H. J., Fardal, M., & Weinberg, M. D. 2011, MNRAS, 416, 660

Lucy, L. B. 1977, AJ, 82, 1013

Makino, N., Sasaki, S., & Suto, Y. 1998, ApJ, 497, 555

Maoz, D., Mannucci, F., & Brandt, T. D. 2012, MNRAS, 426, 3282

Marinacci, F., Pakmor, R., & Springel, V. 2014a, MNRAS, 437, 1750

Marinacci, F., Pakmor, R., Springel, V., & Simpson, C. M. 2014b, MNRAS, 442, 3745

Matteucci, F., Panagia, N., Pipino, A., Mannucci, F., Recchi, S., & Della Valle, M. 2006, MNRAS, 372, 265

McKee, C. F., & Ostriker, J. P. 1977, ApJ, 218, 148

Miniati, F. 2014, ApJ, 782, 21

Monaco, P., Benson, A. J., De Lucia, G., Fontanot, F., Borgani, S., & Boylan-Kolchin, M. 2014, MNRAS, 441, 2058

Monaghan, J. J. 1992, ARA&A, 30, 543

—. 2005, Reports on Progress in Physics, 68, 1703

Muñoz, D. J., Springel, V., Marcus, R., Vogelsberger, M., & Hernquist, L. 2012, MNRAS, 63

Murante, G., Calabrese, M., De Lucia, G., Monaco, P., Borgani, S., & Dolag, K. 2012, ApJ, 749, L34

Muratov, A. L., Keres, D., Faucher-Giguere, C.-A., Hopkins, P. F., Quataert, E., & Murray, N. 2015, ArXiv e-prints

Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, ApJ, 490, 493

Navarro, J. F., & White, S. D. M. 1993, MNRAS, 265, 271

Nelson, D., Genel, S., Vogelsberger, M., Springel, V., Sijacki, D., Torrey, P., & Hernquist, L. 2015, MNRAS, 448, 59

Nelson, D., Vogelsberger, M., Genel, S., Sijacki, D., Kereš, D., Springel, V., & Hernquist, L. 2013, MNRAS, 429, 3353

Norman, M. L., Paschos, P., & Abel, T. 1998, Mem. Soc. Astron. Italiana, 69, 455

Oñorbe, J., Garrison-Kimmel, S., Maller, A. H., Bullock, J. S., Rocha, M., & Hahn, O. 2014, MNRAS, 437, 1894

Ocvirk, P., Pichon, C., & Teyssier, R. 2008, MNRAS, 390, 1326

Okamoto, T., Frenk, C. S., Jenkins, A., & Theuns, T. 2010, MNRAS, 406, 208

Oppenheimer, B. D., & Davé, R. 2006, MNRAS, 373, 1265

—. 2008, MNRAS, 387, 577

Oppenheimer, B. D., Davé, R., Kereš, D., Fardal, M., Katz, N., Kollmeier, J. A., & Weinberg, D. H. 2010, MNRAS, 406, 2325

O'Shea, B. W., Nagamine, K., Springel, V., Hernquist, L., & Norman, M. L. 2005, ApJS, 160, 1

Overzier, R., Lemson, G., Angulo, R. E., Bertin, E., Blaizot, J., Henriques, B. M. B., Marleau, G.-D., & White, S. D. M. 2013, MNRAS, 428, 778

Pakmor, R., Bauer, A., & Springel, V. 2011, MNRAS, 418, 1392

Pakmor, R., Springel, V., Bauer, A., Mocz, P., Munoz, D. J., Ohlmann, S. T., Schaal, K., & Zhu, C. 2015, ArXiv e-prints

Park, C. 1990, MNRAS, 242, 59P

Pérez, F., & Granger, B. E. 2007, Computing in Science and Engineering, 9, 21

Petitjean, P., Aracil, B., Srianand, R., & Ibata, R. 2000, A&A, 359, 457

Pieri, M. M., et al. 2014, MNRAS, 441, 1718

Pillepich, A., et al. 2014, MNRAS, 444, 237

Planck Collaboration et al. 2014, A&A, 571, A1

Portinari, L., Chiosi, C., & Bressan, A. 1998, A&A, 334, 505

Press, W. H., & Schechter, P. 1974, ApJ, 187, 425

Prochaska, J. X., Hennawi, J. F., & Simcoe, R. A. 2013, ApJ, 762, L19

Prochaska, J. X., Lau, M. W., & Hennawi, J. F. 2014, ApJ, 796, 140

Puchwein, E., & Springel, V. 2013, MNRAS, 428, 2966

Putman, M. E., Peek, J. E. G., & Joung, M. R. 2012, ARA&A, 50, 491

Rahmati, A., Pawlik, A. H., Raičevic, M., & Schaye, J. 2013, MNRAS, 430, 2427

Rahmati, A., Schaye, J., Bower, R. G., Crain, R. A., Furlong, M., Schaller, M., & Theuns, T. 2015, ArXiv e-prints

Rasera, Y., Alimi, J.-M., Courtin, J., Roy, F., Corasaniti, P.-S., Füzfa, A., & Boucher, V. 2010, in American Institute of Physics Conference Series, Vol. 1241, American Institute of Physics Conference Series, ed. J.-M. Alimi & A. Fuözfa, 1134–1139

Rauch, M., Sargent, W. L. W., Barlow, T. A., & Carswell, R. F. 2001, ApJ, 562, 76

Rees, M. J., & Ostriker, J. P. 1977, MNRAS, 179, 541

Riebe, K., et al. 2013, Astronomische Nachrichten, 334, 691

Rodriguez-Gomez, V., et al. 2015, ArXiv e-prints

Rosdahl, J., & Blaizot, J. 2012, MNRAS, 423, 344

Rubin, K. H. R., Hennawi, J. F., Prochaska, J. X., Simcoe, R. A., Myers, A., & Wingyee Lau, M. 2014, ArXiv e-prints

Rudie, G. C., Steidel, C. C., Shapley, A. E., & Pettini, M. 2013, ApJ, 769, 146

Rudie, G. C., et al. 2012, ApJ, 750, 67

Salem, M., & Bryan, G. L. 2014, MNRAS, 437, 3312

Sales, L. V., Navarro, J. F., Theuns, T., Schaye, J., White, S. D. M., Frenk, C. S., Crain, R. A., & Dalla Vecchia, C. 2012, MNRAS, 423, 1544

Sales, L. V., et al. 2015, MNRAS, 447, L6

Sánchez Almeida, J., Elmegreen, B. G., Muñoz-Tuñón, C., & Elmegreen, D. M. 2014, A&A Rev., 22, 71

Sarkar, K. C., Nath, B. B., Sharma, P., & Shchekinov, Y. 2014, ArXiv e-prints

Sazonov, S. Y., Ostriker, J. P., Ciotti, L., & Sunyaev, R. A. 2005, MNRAS, 358, 168

Scannapieco, C., et al. 2012, MNRAS, 423, 1726

Schaal, K., & Springel, V. 2015, MNRAS, 446, 3992

Schaller, M., et al. 2014, ArXiv e-prints

Schaye, J., Carswell, R. F., & Kim, T.-S. 2007, MNRAS, 379, 1169

Schaye, J., et al. 2015, MNRAS, 446, 521

Seitenzahl, I. R., Röpke, F. K., Fink, M., & Pakmor, R. 2010, MNRAS, 407, 2297

Shen, S., Madau, P., Guedes, J., Mayer, L., Prochaska, J. X., & Wadsley, J. 2013, ApJ, 765, 89

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

Sijacki, D., Springel, V., & Haehnelt, M. G. 2009, MNRAS, 400, 100

Sijacki, D., Vogelsberger, M., Genel, S., Springel, V., Torrey, P., Snyder, G., Nelson, D., & Hernquist, L. 2014, ArXiv e-prints

Sijacki, D., Vogelsberger, M., Kereš, D., Springel, V., & Hernquist, L. 2012, MNRAS, 424, 2999

Silk, J. 1977, ApJ, 211, 638

Simcoe, R. A., Sargent, W. L. W., Rauch, M., & Becker, G. 2006, ApJ, 637, 648

Skillman, S. W., Warren, M. S., Turk, M. J., Wechsler, R. H., Holz, D. E., & Sutter, P. M. 2014, ArXiv e-prints

Smith, B., Sigurdsson, S., & Abel, T. 2008, MNRAS, 385, 1443

Snyder, G. F., et al. 2015, ArXiv e-prints

Sparre, M., et al. 2015, MNRAS, 447, 3548

Spergel, D. N., et al. 2003, ApJS, 148, 175

Springel, V. 2005, MNRAS, 364, 1105

—. 2010a, MNRAS, 401, 791

—. 2010b, ARA&A, 48, 391

Springel, V., Di Matteo, T., & Hernquist, L. 2005a, MNRAS, 361, 776

Springel, V., & Hernquist, L. 2002, MNRAS, 333, 649

—. 2003, MNRAS, 339, 289

Springel, V., White, S. D. M., Tormen, G., & Kauffmann, G. 2001a, MNRAS, 328, 726

Springel, V., Yoshida, N., & White, S. D. M. 2001b, New Astronomy, 6, 79

Springel, V., et al. 2005b, Nature, 435, 629

—. 2008, MNRAS, 391, 1685

Srisawat, C., et al. 2013, MNRAS, 436, 150

Steidel, C. C., Erb, D. K., Shapley, A. E., Pettini, M., Reddy, N., Bogosavljević, M., Rudie, G. C., & Rakic, O. 2010, ApJ, 717, 289

Stewart, K. R., Brooks, A. M., Bullock, J. S., Maller, A. H., Diemand, J., Wadsley, J., & Moustakas, L. A. 2013, ApJ, 769, 74

Stewart, K. R., Kaufmann, T., Bullock, J. S., Barton, E. J., Maller, A. H., Diemand, J., & Wadsley, J. 2011, ApJ, 738, 39

Stinson, G. S., Brook, C., Macciò, A. V., Wadsley, J., Quinn, T. R., & Couchman, H. M. P. 2013, MNRAS, 428, 129

Stoughton, C., et al. 2002, AJ, 123, 485

Suresh, J., Bird, S., Vogelsberger, M., Genel, S., Torrey, P., Sijacki, D., Springel, V., & Hernquist, L. 2015, preprint, (arXiv:1501.02267)

Sutherland, R. S., & Dopita, M. A. 1993, ApJS, 88, 253

Suto, Y., Sasaki, S., & Makino, N. 1998, ApJ, 509, 544

Szalay, A. S., Gray, J., Thakar, A. R., Kunszt, P. Z., Malik, T., Raddick, J., Stoughton, C., & vandenBerg, J. 2002a, eprint arXiv:cs/0202013

Szalay, A. S., Gray, J., & VandenBerg, J. 2002b, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 4836, Survey and Other Telescope Technologies and Discoveries, ed. J. A. Tyson & S. Wolff, 333–338

Szalay, A. S., Kunszt, P. Z., Thakar, A. R., Gray, J., & Slutz, D. 2000, in Astronomical Society of the Pacific Conference Series, Vol. 216, Astronomical Data Analysis Software and Systems IX, ed. N. Manset, C. Veillet, & D. Crabtree, 405

Tasker, E. J., Brunino, R., Mitchell, N. L., Michielsen, D., Hopton, S., Pearce, F. R., Bryan, G. L., & Theuns, T. 2008, MNRAS, 390, 1267

Teyssier, R. 2002, A&A, 385, 337

The Dark Energy Survey Collaboration. 2005, ArXiv Astrophysics e-prints

Thielemann, F.-K., Nomoto, K., & Yokoi, K. 1986, A&A, 158, 17

Thielemann, F.-K., et al. 2003, Nuclear Physics A, 718, 139

Toro, E. 1999, Riemann Solvers and Numerical Methods for Fluid Dynamics: A Practical Introduction, Applied mechanics: Researchers and students (Springer)

Torrey, P., Vogelsberger, M., Genel, S., Sijacki, D., Springel, V., & Hernquist, L. 2014, MNRAS, 438, 1985

Torrey, P., Vogelsberger, M., Sijacki, D., Springel, V., & Hernquist, L. 2012, MNRAS, 427, 2224

Torrey, P., et al. 2015, MNRAS, 447, 2753

Travaglio, C., Hillebrandt, W., Reinecke, M., & Thielemann, F.-K. 2004, A&A, 425, 1029

Turner, M. L., Schaye, J., Steidel, C. C., Rudie, G. C., & Strom, A. L. 2014a, ArXiv e-prints

—. 2014b, MNRAS, 445, 794

Übler, H., Naab, T., Oser, L., Aumer, M., Sales, L. V., & White, S. D. M. 2014, MNRAS, 443, 2092

van de Voort, F., & Schaye, J. 2012, MNRAS, 2882

van de Voort, F., Schaye, J., Booth, C. M., & Dalla Vecchia, C. 2011a, MNRAS, 415, 2782

van de Voort, F., Schaye, J., Booth, C. M., Haas, M. R., & Dalla Vecchia, C. 2011b, MNRAS, 414, 2458

van Leer, B. 1977, J. Comput. Phys., 23, 276

—. 1979, Journal of Computational Physics, 32, 101

Vazza, F., Gheller, C., & Brunetti, G. 2010, A&A, 513, A32

Viola, M., Monaco, P., Borgani, S., Murante, G., & Tornatore, L. 2008, MNRAS, 383, 777

Vogelsberger, M., Genel, S., Sijacki, D., Torrey, P., Springel, V., & Hernquist, L. 2013, MNRAS, 436, 3031

Vogelsberger, M., Sijacki, D., Kereš, D., Springel, V., & Hernquist, L. 2012, MNRAS, 425, 3024

Vogelsberger, M., et al. 2014a, MNRAS, 444, 1518

—. 2014b, Nature, 509, 177

von Hoerner, S. 1960, ZAp, 50, 184

VonNeumann, J., & Richtmyer, R. D. 1950, Journal of Applied Physics, 21, 232

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

Wang, Y., Su, Y., & Agrawal, G. 2013, in , 335–342

Wellons, S., et al. 2015, MNRAS, 449, 361

Wetzel, A. R., & Nagai, D. 2014, ArXiv e-prints

White, S. D. M., & Frenk, C. S. 1991, ApJ, 379, 52

White, S. D. M., Frenk, C. S., & Davis, M. 1983, ApJ, 274, L1

White, S. D. M., & Rees, M. J. 1978, MNRAS, 183, 341

Wiersma, R. P. C., Schaye, J., & Smith, B. D. 2009a, MNRAS, 393, 99

Wiersma, R. P. C., Schaye, J., Theuns, T., Dalla Vecchia, C., & Tornatore, L. 2009b, MNRAS, 399, 574

Woods, R. M., Wadsley, J., Couchman, H. M. P., Stinson, G., & Shen, S. 2014, MNRAS, 442, 732

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

Xu, G. 1995, ApJS, 98, 355

Yepes, G., Kates, R., Khokhlov, A., & Klypin, A. 1997, MNRAS, 284, 235

York, D. G., et al. 2000, AJ, 120, 1579

Yoshida, N., Stoehr, F., Springel, V., & White, S. D. M. 2002, MNRAS, 335, 762

Yoshikawa, K., Yoshida, N., & Umemura, M. 2013, ApJ, 762, 116

Yu, Q., & Tremaine, S. 2002, MNRAS, 335, 965

Zavala, J., Balogh, M. L., Afshordi, N., & Ro, S. 2012, MNRAS, 426, 3464

Zel'dovich, Y. B. 1970, A&A, 5, 84

Zhu, Q., Hernquist, L., & Li, Y. 2014, ArXiv e-prints