Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **459,** 3025–3039 (2016)
Advance Access publication 2016 March 30

doi:10.1093/mnras/stw674

Dark-ages Reionization and Galaxy formation simulation – I. The dynamical lives of high-redshift galaxies

Gregory B. Poole,[1]⋆ Paul W. Angel,[1] Simon J. Mutch,[1] Chris Power,[2] Alan R. Duffy,[1,3] Paul M. Geil,[1] Andrei Mesinger[4] and Stuart B. Wyithe[1]

[1]*School of Physics, University of Melbourne, Parksville, VIC 3010, Australia*
[2]*ICRAR, University of Western Australia, 35 Stirling Highway, Crawley, Western Australia 6009, Australia*
[3]*Centre for Astrophysics and Supercomputing, Swinburne University of Technology, PO Box 218, Hawthorn, VIC 3122, Australia*
[4]*Scuola Normale Superiore, Piazza dei Cavalieri 7, I-56126 Pisa, Italy*

Accepted 2016 March 11. Received 2016 February 23; in original form 2015 August 17

ABSTRACT

We present the Dark-ages Reionization and Galaxy formation Observables from Numerical Simulations (DRAGONS) programme and Tiamat, the collisionless *N*-body simulation programme upon which DRAGONS is built. The primary trait distinguishing Tiamat from other large simulation programme is its density of outputs at high redshift (100 from $z = 35$ to $z = 5$; roughly one every 10 Myr) enabling the construction of very accurate merger trees at an epoch when galaxy formation is rapid and mergers extremely frequent. We find that the friends-of-friends halo mass function agrees well with the prediction of Watson et al. at high masses, but deviates at low masses, perhaps due to our use of a different halo finder or perhaps indicating a break from 'universal' behaviour. We then analyse the dynamical evolution of galaxies during the Epoch of Reionization finding that only a small fraction ($\sim$20 per cent) of galactic haloes are relaxed. We illustrate this using standard relaxation metrics to establish two dynamical recovery time-scales: (*i*) haloes need $\sim$1.5 dynamical times following formation, and (*ii*) $\sim$2 dynamical times following a major (3:1) or minor (10:1) merger to be relaxed. This is remarkably consistent across a wide mass range. Lastly, we use a phase-space halo finder to illustrate that major mergers drive long-lived massive phase-space structures which take many dynamical times to dissipate. This can yield significant differences in the inferred mass build-up of galactic haloes and we suggest that care must be taken to ensure a physically meaningful match between the galaxy formation physics of semi-analytic models and the halo finders supplying their input.

Key words: galaxies: formation – galaxies: high-redshift – cosmology: theory – dark ages, reionization, first stars – early Universe.

1 INTRODUCTION

Following cosmological recombination the baryonic gas filling the Universe became predominantly neutral. The fact that this gas is known to be mostly ionized today (Gunn & Peterson 1965) implies that the intergalactic medium (IGM) underwent a significant *reionization* event at some early point in its history. This fact is responsible for some of the major questions in extragalactic astronomy including: when did this process occur and what were the responsible ionizing sources? Recent observations have begun to provide preliminary answers to this question (e.g. Fan et al. 2006; Ouchi et al. 2010; Planck Collaboration XIII 2015). Soon measurements of highly redshifted 21 cm radio emission (Furlanetto, Oh & Briggs 2006b; Morales & Wyithe 2010) will open an important new observational window for study of the first galaxies, providing the first direct probe of the neutral hydrogen content in the early Universe.

The development of theoretical models that self-consistently include the physics of galaxy formation and intergalactic hydrogen will play a key role in understanding the nature of the first galaxies and in interpreting these observations. This paper is the first in a series describing the Dark-ages Reionization and Galaxy formation Observables from Numerical Simulations (DRAGONS[1]) project which aims to integrate detailed semi-analytic models

---

⋆ E-mail: gpoole@unimelb.edu.au

[1] http://dragons.ph.unimelb.edu.au/

Published by Oxford University Press on behalf of the Royal Astronomical Society

constructed specifically to study galaxy formation at high redshift, with seminumerical models of the galaxy–IGM interaction (Mesinger & Furlanetto 2007; Zahn et al. 2007; Geil & Wyithe 2008). The galaxy formation modelling for DRAGONS is implemented using a set of large *N*-body simulations which we refer to as the Tiamat simulation suite. Tiamat provides a framework within which to implement a semi-analytic model for reionization and to study the formation histories, structure and properties of the dark matter haloes that dictate the formation sites and assembly histories of the first galaxies.

Over the past decade the requirements for simulations aiming to address the structure of reionization and galaxy formation in the Epoch of Reionization (EoR) have been studied extensively. The consensus from previous *N*-body studies (e.g. Iliev et al. 2007; McQuinn et al. 2007; Zahn et al. 2007; Croft & Altay 2008; Lee et al. 2008; Shin, Trac & Cen 2008) and analytic models (e.g. Furlanetto, Zaldarriaga & Hernquist 2004; Wyithe & Morales 2007; Barkana 2009) is that large-scale overdense regions near bright sources ionize first with clustered neighbouring sources contributing to increase the size of ionized regions. Simulations on the scale of 100 Mpc are found to be large enough to correctly capture the structure and duration of reionization, although volumes up to 500^3 Mpc are required to capture all large-scale power due to clustering of H II regions (Iliev et al. 2014).

The challenge is to model galaxy formation in volumes of this size with sufficient resolution. In the cold neutral IGM prior to reionization, molecular cooling may proceed within minihaloes with masses $\sim 10^6\,\mathrm{M}_\odot$. However, the processes principally responsible for regulating galaxy formation are expected to be active in haloes with virial temperatures greater than $T_{\rm min} \sim 10^4$ K, above which atomic hydrogen cooling becomes efficient. On the other hand, the growth of H II regions during reionization is also expected to be influenced by radiative feedback due to suppression of galaxy formation below the cosmological Jeans mass within a heated IGM (e.g. Dijkstra et al. 2004). Together these constraints indicate that sufficient resolution is required to identify halo masses down to $\sim 5 \times 10^7\,\mathrm{M}_\odot$ solar masses within a volume of ∼100 Mpc.

In addition to this dynamic range of scales, for the DRAGONS programme we aim to accurately resolve the relevant time-scales of high-redshift galaxy formation putting an additional constraint on the cadence with which simulation outputs must be generated. At $z \sim 6$ the dynamical time of a galactic disc falls below the lifetime of the least massive Type II supernova progenitor ($\sim 4 \times 10^7$ yr). As a result, snapshots with a cadence of $\sim 10^7$ yr are required to follow galaxy formation correctly during the EoR with a semi-analytic model. This interval is an order of magnitude shorter than needed to describe galaxy formation at redshifts $z \sim 0$.

In this paper, we present the Tiamat suite of collisionless *N*-body simulations which we have run to satisfy these requirements and upon which the DRAGONS programme will be constructed. Given its critical importance as the foundation of the programme, we take this opportunity to present the methodology of constructing this set of simulations and to characterize the populations of galactic haloes obtained. In particular, we shall carefully examine the dynamical evolution of galactic haloes during the reionization era. We seek this understanding because of its potential impact on the structure of galactic haloes, which is of fundamental importance to the physics of galaxy formation at any epoch, including the EoR. In particular, halo concentrations and angular momenta are generally believed to dictate the size and surface density of the disc-like structures in which the majority of star formation occurs. Dynamical disturbances can additionally drive starbursts or affect the stability of these disc-like structures, strongly affecting star formation and forcing morphological transformations which contribute to the assembly of galactic spheroids. This can in turn affect observable galaxy sizes or alter UV fluxes and escape fractions with important effects on the reionization history of the Universe.

At low redshift, it has been shown that a halo's dynamical state can systematically affect the structure and gravitational potential of galactic haloes (e.g. Thomas et al. 1998; Neto et al. 2007; Power, Knebe & Knollmann 2012; Ludlow et al. 2014). These studies have collectively established a set of criteria (which we refer to henceforth as 'standard' relaxation criteria) capable (at low redshifts at least) of separating haloes with disturbed structure from those with relaxed structure. These standard criteria consist of cuts on three metrics for each halo: the separation of its dense centre from its centre of mass ($x_{\rm off}$), its 'pseudo-virial ratio' constructed from its velocity dispersion and gravitational binding energy (φ), and its substructure fraction ($f_{\rm sub}$). When low-redshift haloes are separated into relaxed and unrelaxed samples in this way, substantial effects on halo concentration and (to a lesser extent) spin have been demonstrated. This is of particular importance to studies which aim to understand the processes which establish the 'universal' density profiles of haloes extracted from collisionless *N*-body simulations (Navarro, Frenk & White 1997).

While there is broad agreement at low redshift as to the dependence of halo structure on mass, redshift and dynamical state, recent studies which have attempted to push this understanding to the EoR (e.g. Prada et al. 2012; Dutton & Macciò 2014; Diemer & Kravtsov 2015; Hellwing et al. 2016) have found less consensus. At high redshifts where simulations predict that merger rates are very high, haloes significantly less concentrated, and merger orbital properties quite different, the influence of dynamical state on halo structure may differ from local trends. It is unclear to what degree dynamical disturbance may play a role in the differences in high-redshift halo structure reported in the literature since these studies have not been consistent in their treatment of this issue.

Unfortunately, the details of how the standard relaxation metrics evolve following dynamical disturbances has not been properly explored at any redshift, nor has their efficacy at separating relaxed systems from unrelaxed systems been demonstrated at high redshift. Before presenting a detailed analysis of halo structure at high redshift to understand the discrepancies present in the literature, we aim first to address both of these issues here. Due to the historical focus on low redshifts by galaxy formation models, few large simulation programmes possess sufficient temporal resolution to perform a thorough dynamical analyses at high redshift. Given its fine snapshot temporal resolution, Tiamat represents a unique resource for exploring these issues across the full range of masses most relevant to galaxy formation in the early Universe. We will find that the standard relaxation criteria are effective at identifying systems that are recovering from their formation or from recent significant mergers. With this methodology properly validated at high redshift, we will subsequently perform a detailed analysis of the structure of both relaxed and unrelaxed high-redshift dark matter haloes – including spin parameter and concentration–mass relations – in a companion paper (Angel et al. 2016, hereafter Paper II).

The Tiamat *N*-body simulation hosts a semi-analytic model of galaxy formation named MERAXES, which has been integrated within a seminumerical model for ionization structure. In subsequent papers, we will present this model (Mutch et al. 2015, hereafter Paper III) and use it in a range of studies including high-redshift galaxy luminosity functions (Liu et al. 2015, hereafter Paper IV) and the ionization structure of the IGM (Geil et al. 2015, hereafter Paper V).

Table 1. Box sizes (L), particle counts (N_p), particle mass (m_p), gravitational softening lengths (ϵ) and integration accuracy parameters (η) for the Tiamat simulations as well as the cosmology and halo finding codes used for each.

Simulation	N_p	L (Mpc/h)	m_p ($\mathrm{M}_\odot/h$)	ϵ (kpc/h)	η	Cosmology	Halo finding
Tiamat	2160^3	67.8	2.64×10^6	0.63	0.025	*Planck*-2015	SUBFIND
Medi Tiamat	1080^3	22.6	7.83×10^5	0.42	0.025	*Planck*-2015	SUBFIND
Tiny Tiamat	1080^3	10.0	6.79×10^4	0.19	0.025	*Planck*-2015	SUBFIND
Tiny Tiamat-W07	1024^3	10.0	7.11×10^4	0.20	0.010	*WMAP*-07	SUBFIND and ROCKSTAR

Complementary high-resolution hydrodynamics simulations called *Smaug* (already presented in Duffy et al. 2014) will characterize the basic scaling relationships of early galaxy formation. There will then be a detailed comparison of MERAXES to the results of *Smaug* with suggested constraints of the semi-analytic model based on hydrodynamics (Qin et al. 2016, hereafter Paper VII).

The outline of this paper is as follows. In Section 2, we introduce the construction of the Tiamat simulations including our approaches to halo finding and merger tree construction. In Section 3, we analyse these data products to identify how galactic haloes relax at high redshift, estimate their relaxed fraction and present some preliminary findings on their phase-space structure. Finally, in Section 4, we summarize our study and present our conclusions. Our choice of fiducial cosmology throughout will be a standard spatially flat *Planck* Λ cold dark matter (ΛCDM) cosmology based on 2015 data (Planck Collaboration XIII 2015) (h, Ω_m, Ω_b, Ω_Λ, σ_8, n_s) = (0.678, 0.308, 0.0484, 0.692, 0.815, 0.968) although we will make isolated use of two simulations run with standard spatially flat *WMAP*-5 (Komatsu et al. 2009) (h, Ω_m, Ω_b, Ω_Λ, σ_8, n_s) = (0.727, 0.273, 0.0456, 0.705, 0.812, 0.96) and *WMAP*-7 (Komatsu et al. 2011) (h, Ω_m, Ω_b, Ω_Λ, σ_8, n_s) = (0.702, 0.275, 0.0458, 0.725, 0.816, 0.96) ΛCDM cosmologies.

2 SIMULATIONS

In this section, we present our approach to assembling the Tiamat suite of collisionless N-body simulations. We describe the means by which initial conditions were generated, the simulation code used and its parameters and the means by which bound structures (haloes) were extracted and assembled into merger trees for the analysis presented in subsequent sections. A summary of the most essential metrics of our simulations is presented in Table 1.

2.1 Simulation runs

The simulation data products upon which the DRAGONS programme is being built come primarily from Tiamat; its flagship collisionless N-body simulation. Tiamat has been designed to provide sufficient mass resolution to accurately capture the low-mass galaxy population driving the reionization of the Universe at high redshifts and to do so over a sufficiently large volume to capture the evolving structure of the reionization field, right to the epoch of bubble overlap. It does so with sufficiently high snapshot temporal resolution to fully capture the rapid evolution of both. Specifically, Tiamat consists of a periodic box, 100 Mpc (comoving) on a side sampled with 2160^3 particles with 100 snapshots of particle data recorded at intervals equally spaced in cosmic time from $z = 35$ to $z = 5$ (i.e. every 11.1 Myr). To facilitate several smaller volume but higher resolution studies we have run a series of companion simulations – each consisting of $\sim 10^9$ particles – named *Tiny Tiamat* (10 h^{-1}Mpc box) and *Medi Tiamat* (22.6 h^{-1}Mpc box) using the same snapshot cadence strategy as Tiamat. We have also run a companion 10 h^{-1}Mpc box named *Tiny Tiamat-W07* with a *WMAP*-7 cosmology.

All aspects of the simulations were performed assuming standard ΛCDM cosmologies with the parameters given at the end of Section 1. Initial conditions were generated with the second-order Lagrangian perturbation code 2LPTIC (Crocce, Pueblas & Scoccimarro 2006, with fixes to ensure correct behaviour when particle or displacement field grid cell counts exceed $2^{32} - 1$) at $z = 99$ ($z = 127$ for *Tiny Tiamat-W07*) using a particle load with a regular periodic grid structure and a displacement field computed on regular periodic grids of dimensions 2160^3 for Tiamat and 1080^3 for all other cases. The input power spectrum was generated using CAMB (Lewis, Challinor & Lasenby 2000) with the ΛCDM parameters appropriate to each.

We have run our simulations using GADGET-2 (Springel 2005), a Tree-Particle Mesh (TREEPM) code well suited to large distributed memory systems, with the RAM conserving modifications listed in Poole et al. (2015). Gravitational softening was set to $0.02\bar{d}$ (where $\bar{d} = L/\sqrt[3]{N_p}$ is the mean interparticle spacing of a simulation with N_p particles in a cubic volume of side length L) in all cases (as motivated by Poole et al. 2015) and the integration accuracy parameter was set to $\eta = 0.025$ in all cases except *Tiny Tiamat-W07* where $\eta = 0.01$. For all but *Tiny Tiamat-W07*, this is a slight relaxation of the choice argued for in Poole et al. (2015) but was deemed necessary to reduce the wallclock time for the calculation to required levels.

Lastly, at several places we will seek to compare our high-redshift findings from Tiamat to the dynamical activity of similarly sized systems at low redshift. To facilitate this comparison, we will use the *GiggleZ-HR* simulation (i.e. the highest resolution *GiggleZ* control volume simulation, see Poole et al. 2015 for details). This simulation consists of 1080^3 particles in a 125 h^{-1}Mpc box and was run with a *WMAP*-5 cosmology. Halo finding and tree building was performed in the same way and with the same code versions used for Tiamat. For the purposes of the work we present here, this simulation should provide an adequate comparison between our highest mass high-redshift Tiamat haloes and haloes of similar mass in the low-redshift Universe.

2.2 Halo finding and merger tree construction

We have performed the majority of our halo finding using the widely utilized SUBFIND code of Springel et al. (2001). This code initially identifies collapsed regions of interest using a friends-of-friends (FoF) algorithm (for which we use the standard linking length criterion of $0.2\bar{d}$) and subsequently identifies bound substructures within these FoF groups as locally overdense collections of particles, removing unbound particles through an unbinding procedure. Halo centres are taken to be the position of a halo's most bound particle as identified by SUBFIND unless otherwise stated.

For the *Tiny Tiamat-W07* simulation we have additionally performed the halo finding exercise using the publicly available version of the 6D phase-space code ROCKSTAR (Behroozi, Wechsler & Wu 2013). This code also performs an initial FoF search for structures, this time with a generous linking length of $0.26\bar{d}$, and subsequently searches these objects for distinct structures in phase space. Careful examination of the particle list output from ROCKSTAR revealed numerous cases where particles are assigned to multiple haloes or whole haloes are duplicated (involving approximately 0.1 per cent of particles at $z = 5$ in *Tiny Tiamat-W07*). In the course of rewriting the particle lists to match the standard SUBFIND format required by our analysis codes (i.e., particle lists organized by FoF group, in order of subgroup size and sorted in radial order from the centre of the system), we have eliminated these cases by removing duplicated haloes and allocating particles with multiple halo assignments to the halo with the nearest centre as determined using the shrinking sphere method of Power et al. (2003).

We have constructed merger trees from these halo catalogues following the method to be presented in Poole et al. (2016; in preparation). This approach carefully repairs artefacts introduced by imperfections in the halo finding process and identifies pathologies in the merger trees (e.g. instances when tree branches are broken due to overlinking by the halo finder or as a result of halo fragmentations) through a process of forward and backward matching which scans both ways over multiple snapshots. For our Tiamat trees, we have used 16 snapshots for this process (i.e. $\Delta t_{\rm scan} \sim 160$ Myr) representing more than a full dynamical time even at $z = 5$, which we find sufficient for an accurate calculation.

2.3 Halo mass function

In addition to halo accretion histories and merger trees, the evolving abundance of galactic haloes in Tiamat will be a primary determinant of the galaxy populations and reionization histories we derive from DRAGONS modelling efforts. Parametrizations of these mass functions are of great utility for a wide variety of seminumerical extragalactic calculations and are an excellent way to facilitate comparisons to halo populations from other studies in the literature. For these reasons, and given the excellent combination of volume and mass resolution covered by Tiamat at high redshift, we present here the Tiamat halo mass functions and a parametrized fit to them.

A great deal of effort has been invested in the literature on methods of accurately and robustly estimating the halo mass function across a wide range of masses, cosmologies and redshifts and we refer the interested reader to the studies of Press & Schechter (1974), Jenkins et al. (2001), Lukić et al. (2007), Reed et al. (2007), Tinker et al. (2008), Knebe et al. (2013), Murray, Power & Robotham (2013) and references therein for a detailed account of the subject. Our emphasis here is merely to present the halo mass function in the most recent *Planck* cosmology across ranges in mass and redshift relevant to galaxy formation during the EoR. Given the recent interest in the 'universal' FoF halo mass function parametrization presented by Watson et al. (2013, see Appendix A for details), we focus here on a parametrization of that form. We have used a Monte Carlo Markov Chain (MCMC) approach to perform this fit and present the details of our approach, the resulting best-fitting parameters and their covariance in Appendix A.

In Fig. 1, we present the FoF halo mass functions derived from the SUBFIND catalogues extracted from the *Tiny Tiamat*, *Medi Tiamat* and Tiamat simulations and compare these to the four-parameter 'universal' FoF mass function of Watson et al. (2013) as fit originally in that work (dotted lines) and as refitted in this work (solid lines). Some small but significant differences are found between the original Watson et al. (2013) parametrization and our re-parametrization, particularly at the lowest masses and highest redshifts. There are many possible sources for this including numerical reasons, such as those arising from systematic differences between our halo finder and that employed in the Watson et al. (2013) study, as well as more physical reasons such as systematic differences in FoF linking during the EoR when the most massive haloes lie in highly filamentary regions where structures are more prone to overlinking by this algorithm. All of our data for the Tiamat fits is at high redshift while the Watson et al. (2013) fits are to data spanning a much wider range of redshifts, so the difference may also simply reflect the fact that the FoF halo mass function deviates from a universal form, at least at high redshifts. Given the important contributions low-mass galaxies make to the ionizing photon budget during the EoR, such differences are important to note.

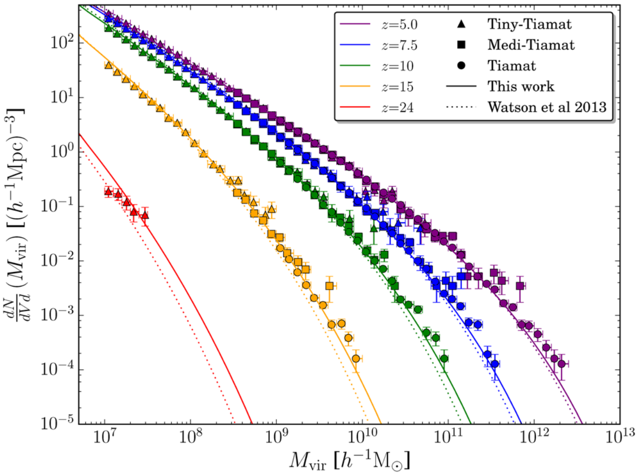

Figure 1. Tiamat mass functions at redshifts $z = 5,7.5,10,15$ and 24 compared to the fitting formula of Watson et al. (2013). Dotted lines present results for the original fit presented in Watson et al. (2013) and solid lines present results of the refitting performed in this study.

3 ANALYSIS

In this section, we examine the time-scales by which galactic haloes at high redshift relax following formation and mergers. We relate these time-scales to their dynamical age and to intervals between merger events and will find that only towards the end of the EoR do significant numbers of haloes exist in relaxed states. Lastly, we will also present some preliminary results about their phase-space structure that may be of consequence for the application of semi-analytic models at high redshift.

3.1 Dynamical recovery time-scales

It has long since been shown (and perhaps not surprising) that the structure of haloes extracted from collisionless *N*-body simulations has a significant dependence on the dynamical state of the system (Thomas et al. 1998; Neto et al. 2007; Power et al. 2012; Ludlow et al. 2014). At low redshifts at least, cuts on three metrics for quantifying the dynamical state of haloes have found success at separating systems with disturbed structure from those with relaxed structure: the offset parameter ($x_{\rm off}$), given by the displacement of the densest centre of a halo from its centre of mass; the virial ratio (φ), given by $2K/|U|$, where K is the kinetic energy and U is the halo's gravitational binding energy (see section 5.1 of Poole et al. 2006, and references therein, for a detailed description of virialization); and the substructure fraction ($f_{\rm sub}$), which we take here to be the ratio of the particle count of all but the most massive of a FoF halo's substructures to its total particle count. Each have simple physical interpretations as measures of dynamical state. Elevated values of $f_{\rm sub}$ naturally arise during the earliest stages of a merger when a halo is naturally split between multiple similarly sized substructures. Elevated values of φ are found prior to the dissipation of orbital energy following a merger. Lastly, elevated values of $x_{\rm off}$ are a natural result of the movement of a halo's dense core as it orbits the centre of mass of its system following even minor disturbances. The standard values for relaxed systems which we adopt are those proposed originally by Neto et al. (2007) and recently confirmed to be successful in the study of halo density profiles (Ludlow et al. 2014); specifically, $x_{\rm off} < 0.07$, $\varphi < 1.35$ and $f_{\rm sub} < 0.1$. To date, a careful examination of how these metrics evolve following dynamical disturbances has not been performed however, leaving the physical nature of these cuts unclear. Additionally, it is unclear how appropriate they are for high-redshift studies.

In what follows we shall study the evolution of these relaxation metrics following three sorts of mass accretion event capable of driving dynamical disturbances: halo formation (defined as the point at which a halo last reached 50 per cent of its present mass; a standard choice in the field, with other similar choices resulting in no qualitative change to our results), mergers between a primary halo and a secondary halo at least one third its mass (so-called major, or 3:1 mergers) or mergers between a primary halo and a secondary halo at least one tenth its mass (so-called minor, or 10:1 mergers). Throughout our analysis we will measure intervals of time for a halo at redshift z in units of its dynamical time, which we take to be 10 per cent of the Hubble time at that redshift. Times in this dimensionless system of units will be denoted by τ. At all redshifts, the Hubble time is $\tau = 10$ in this system. The three times since a halo last experienced each of these events will be referred to as *dynamical ages* and denoted by $\tau_{\rm form}$, $\tau_{3:1}$ and $\tau_{10:1}$, respectively. The dynamical ages required for our relaxation criteria ($x_{\rm off}$, φ and $f_{\rm sub}$) to return to and maintain standard values for relaxed haloes following these events are used to motivate two *recovery times*: a formation recovery time and a merger recovery time.

In Fig. 2, we show the evolution of the distribution of $x_{\rm off}$, φ and $f_{\rm sub}$ for two mass-selected halo populations[2] ($M_{\rm vir} = 10^8$–$10^{8.5}\,h^{-1}\,M_\odot$ taken from *Tiny Tiamat* and $M_{\rm vir} = 10^{10.5}$–$10^{11}\,h^{-1}\,{\rm M}_\odot$ taken from Tiamat) between $z = 5$ and $z = 7.5$ as a function of their three dynamical ages (we have looked at various redshift and mass ranges, finding no evidence of changes to any reported trends). General trends are quickly apparent for each metric as a function of all three dynamical ages. Since different environments, merger orbital properties and the oscillatory nature of $x_{\rm off}$ and φ (see Poole et al. 2006, for an analysis) all lead to scatter in each metric as a function of τ, we focus our discussion here on the path traced by the distribution peak of each metric, unless stated otherwise.

We find that $x_{\rm off}$ starts with high values of $\sim$0.2 at the time of halo formation, declining to our relaxed level of 0.07 at $\tau_{\rm form} \sim 1.5$ and then to baseline levels of $x_{\rm off} \sim 0.04$ afterwards. Following 3:1 and 10:1 mergers, peak levels occur roughly one dynamical time after a merger begins with relaxed levels obtained at $\tau_{3:1} \sim \tau_{10:1} {\sim} 2$. Peak values of 0.2 and roughly 0.07 are reached following 3:1 and 10:1 mergers, suggesting that mergers are progressively less likely to excite the system above our $x_{\rm off} \sim 0.07$ relaxation criterion as mass ratios drop below 10 per cent.

Interestingly, the virial ratio shows significantly less evolution following both formation and merger events. In all cases, the distribution peak sits at levels similar to our $\varphi \sim 1.35$ relaxation criteria at times when $x_{\rm off}$ lies above its relaxation criteria. Once $x_{\rm off}$ is found to drop below relaxation levels (or shortly before) φ can be seen to decline somewhat from values of $\sim$1.35 to $\sim$1. Interestingly, the lower mass haloes from the *Tiny Tiamat* simulation behave similarly to the higher mass haloes in Tiamat, although with a significantly higher tail in the 95 per cent confidence interval. Generally however, φ exhibits much less sensitivity to dynamical disturbances and relaxes to baseline levels quicker than $x_{\rm off}$, suggesting that it is a much less robust discriminator of dynamical state.

Lastly, $f_{\rm sub}$ shows a very simple and well-defined behaviour following dynamical events. At formation, a wide range of values are seen about a distribution peak of $\sim$0.2. A slow decline to baseline values follows. After merger events, $f_{\rm sub}$ increases by expected amounts: 30 per cent for 3:1 mergers and 10 per cent for 10:1 mergers. The subsequent decline in $f_{\rm sub}$ is more rapid than what is seen following formation, with levels dropping at a rate of approximately 20 per cent per dynamical time. We also find that substructure fractions return to standard relaxed values 30–50 per cent faster than core offsets following dynamical disturbances. Despite this, because $f_{\rm sub}$ is most sensitive to dynamical disturbance in the earliest stages of mergers, it is an effective compliment to the $x_{\rm off}$ statistic which exhibits a slight delay in reacting during merger events.

We conclude then that, of the three metrics we study here, the $x_{\rm off}$ statistic is the most effective single measure of dynamical state. It is sensitive to disturbances from mergers greater than approximately 10:1 and retains this sensitivity for approximately 2 dynamical times afterwards. Of course, this quantity is expected to oscillate following dynamical disturbances and shows a delay in responding to merger events, making the joint application of a complimentary and (ideally) independent metric necessary. The $f_{\rm sub}$ metric is effective in this regard but loses sensitivity at late times when $x_{\rm off}$ continues to maintain elevated levels. The virial ratio is significantly less discriminating than these statistics but evolves in ways consistent with the relaxation of $x_{\rm off}$ and $f_{\rm sub}$. Additional details (including specific numbers) regarding the relative influence of each statistic on setting relaxed halo population sizes in Tiamat can be found in Paper II.

From Fig. 2 we also note the remarkable similarity between the relaxation evolution of these two halo populations for all three metrics despite spanning a range of 1000 in mass. In Fig. 3, we present the distribution of all three metrics at three fixed times spanning the important range of their relaxation following each of the three dynamical events presented in Fig. 2. With the exception of $f_{\rm sub}$, all metrics are essentially independent of mass throughout the period of relaxation following halo formation or mergers larger than 10:1. The differing trends of $f_{\rm sub}$ with mass for each simulation is a numerical effect arising from their differing resolutions as a function of mass and is an expected result. Despite this one numerical effect, this figure clearly illustrates that high-redshift haloes recover

---

[2] The mass ranges selected for Fig. 2 were chosen because they yield sufficient statistics (575890 haloes contribute to the Tiamat panel) while consisting of haloes large enough to resolve both the last 3:1 and 10:1 mergers in their progenitor lines. A halo mass of $\log{(M_{\rm vir}/{\rm M}_\odot)} = 10.25$ consists of 6840 particles in Tiamat. The secondary halo of a 10:1 merger has only 684 particles in this case. Given that these haloes are doubling their mass every 1 to 2 dynamical times and we want to follow remnants for 3–4 dynamical times following merger events, these secondary haloes could have been as small as 150 particles when they merged with their descendant.

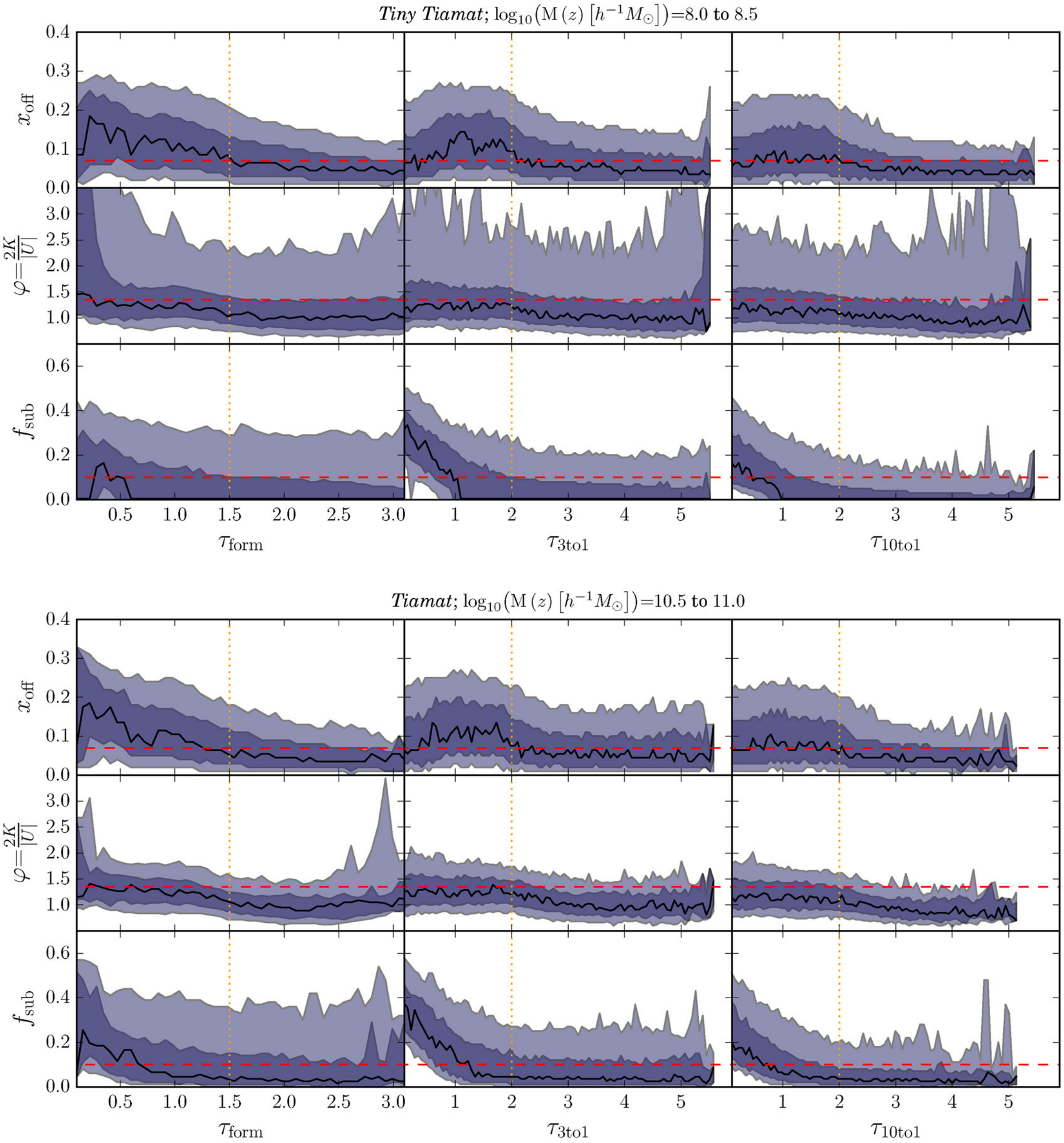

Figure 2. Dependence of the offset parameter (x_{off}), virial ratio (φ) and the substructure fraction (f_{sub}) with the time since a FoF halo's progenitor line: [left] achieved 50 per cent of its current mass (τ_{form}), [mid] last experienced a 3:1 (or larger) merger ($\tau_{3:\,1}$) or [right] last experienced a 10:1 (or larger) merger ($\tau_{10:\,1}$). For a halo at redshift z, times are measured in units of its dynamical time, taken to be 10 per cent of the Hubble time at z (in this system, the Hubble time is always $\tau = 10$). Black lines trace the peak of the distribution while dark and light shaded regions represent 68 per cent and 95 per cent confidence intervals about this peak, respectively. Standard relaxation criteria ($x_{\mathrm{off}} = 0.07$, $\varphi = 1.35$ and $f_{\mathrm{sub}} = 0.1$) are labelled with dashed red lines and our fiducial recovery times (see text) are labelled with dotted orange lines. All results are accumulated for haloes over the redshift range $5 \le z \le 7.5$ for two masses (top, $M_{\mathrm{vir}} = 10^{8}$–$10^{8.5}\,h^{-1}\,\mathrm{M}_\odot$ from *Tiny Tiamat* and bottom, M_{vir}=$10^{10.5}$–$10^{11}\,h^{-1}\,M_\odot$ from Tiamat) spanning a factor of 1000 in mass.

from formation and merger events within a time which is highly insensitive to their mass.

These results suggest that following formation or mergers greater than 10:1, a small and fixed number of pericentric passages of the material disturbed at large radius in the merger remnant are required for relaxation (many more passages may be involved at small radii where densities are higher and dynamical times shorter). If this is the case, the mass independence of relaxation could be seen as a product of the fact that halo crossing times depend only on their mean density, which is defined in terms of a fixed overdensity, and independent of mass. Secondary factors which could influence halo relaxation include halo concentrations, shapes and merger orbital properties. While we have not yet been able to explore the mass dependence of halo shape and merger orbital properties and any possible influence they may exert on halo relaxation, we show in the next paper in this series (Paper II) that halo concentrations are nearly mass-invariant at $z > 5$. This may be partially responsible for the mass invariance of halo relaxation at high redshift in a way that may break down at low redshifts where halo concentrations do in fact have a significant mass dependence.

From these results, we define two mass-independent recovery times separating relaxed and unrelaxed systems at high redshift: $\tau_{\mathrm{form}} = 1.5$ and $\tau_{\mathrm{merge}} = 2$. Our expectation is that high-redshift haloes which have doubled their mass within one and a half

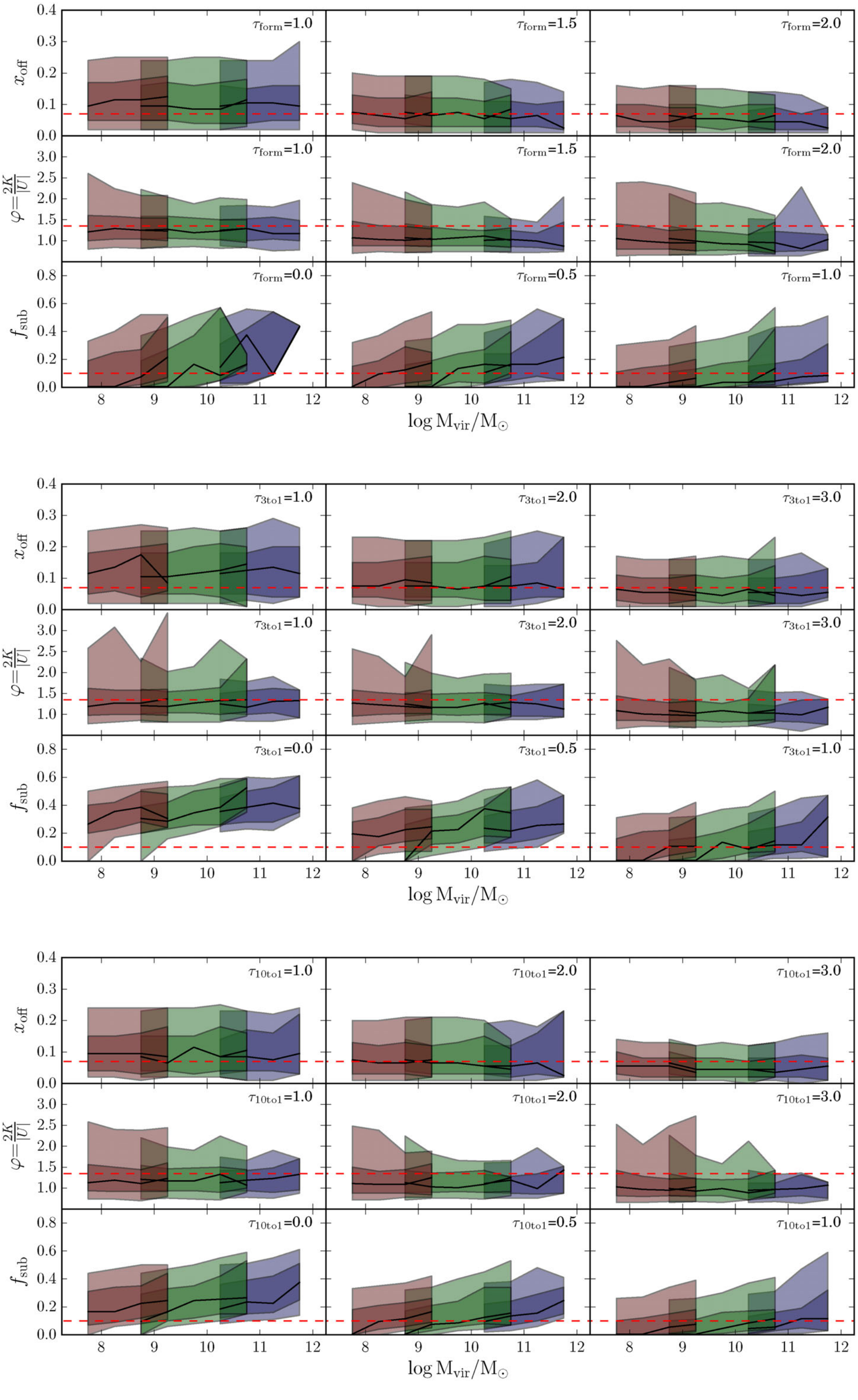

Figure 3. Dependence of the offset parameter ($x_{\rm off}$), virial ratio (φ) and the substructure fraction ($f_{\rm sub}$) on halo mass at a series of fiducial times measured since a FoF halo's progenitor line: [top] achieved 50 per cent of its current mass ($\tau_{\rm form}$), [mid] last experienced a 3:1 (or larger) merger ($\tau_{3:1}$) or [bottom] last experienced a 10:1 (or larger) merger ($\tau_{10:1}$). For a halo at redshift z, times are measured in units of its dynamical time, taken to be 10 per cent of the Hubble time at z (in this system, the Hubble time is always $\tau = 10$). Black lines trace the peak of the distribution while dark and light shaded regions represent 68 per cent and 95 per cent confidence intervals about this peak, respectively. Red, green, and blue represents results taken from the *Tiny Tiamat*, *Medi Tiamat*, and Tiamat simulations, respectively. Standard relaxation criteria ($x_{\rm off} = 0.07$, $\varphi = 1.35$ and $f_{\rm sub} = 0.1$) are labelled with dashed red lines. All results are accumulated for haloes over the redshift range $5 \le z \le 7.5$. The fact that each frame is essentially independent of mass (with the exception of $f_{\rm sub}$, which is tilted across the range of each simulation, as expected, due to resolution effects) reflects the mass independence of how each metric relaxes following halo formation or mergers more significant than 10:1.

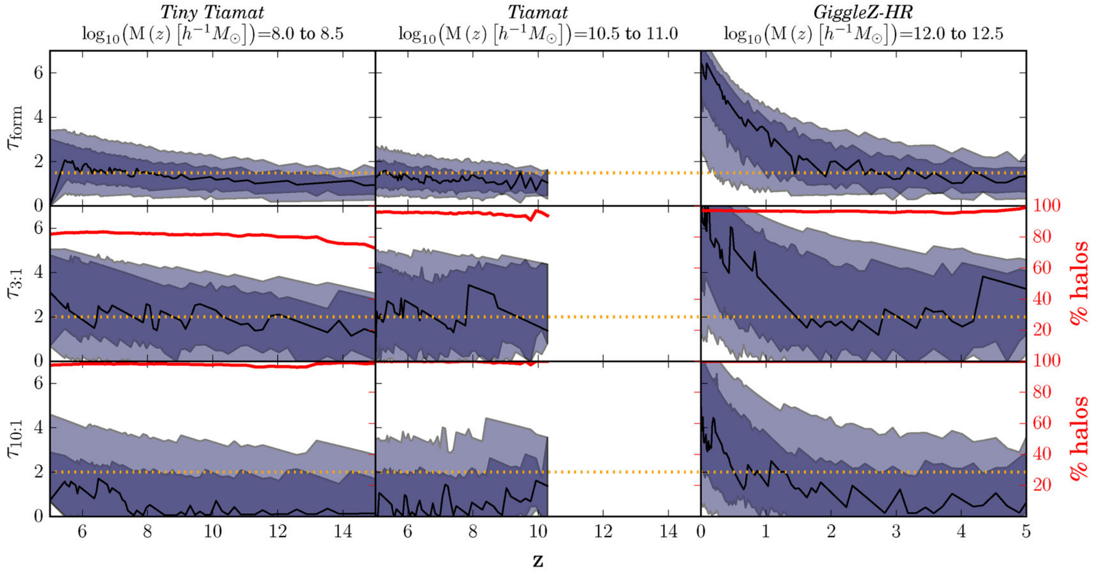

Figure 4. Redshift dependence of the dynamical age (black lines show the distribution peak while dark and light shaded regions represent 68 per cent and 95 per cent confidence intervals about this peak, respectively) of FoF haloes for two cases at $z > 5$ spanning a range of 1000 in mass ([left] $M_{\mathrm{vir}} = 10^8$–$10^{8.5}\ h^{-1}\ M_\odot$ from *Tiny Tiamat* and [middle] $M_{\mathrm{vir}} = 10^{10.5}$–$10^{11}\ h^{-1}\ M_\odot$ from Tiamat) contrasted with [right] the $z < 5$ evolution of $M_{\mathrm{vir}} = 10^{12}$–$10^{12.5}$ FoF haloes from the *GiggleZ-HR* simulation. Our fiducial recovery times ($\tau_{\mathrm{form}} = 1.5$ and $\tau_{\mathrm{merge}} = 2$; haloes above these lines are capable of being relaxed) are labelled with dotted orange lines and the fraction of haloes that have experienced 3:1 or 10:1 mergers are shown in red.

dynamical times ($\tau_{\mathrm{form}} < 1.5$) or which have experienced mergers larger than 10:1 within two dynamical times ($\tau_{\mathrm{merge}} < 2$) are likely to be disturbed.

3.2 Relaxed fraction evolution

How then do the fractions of haloes meeting these recovery criteria evolve with redshift? In Fig. 4, we plot the evolution of the distributions of τ_{form},$\tau_{3:\,1}$and $\tau_{10:\,1}$ as a function of redshift for the two mass-selected populations presented in Fig. 2. The distributions of dynamical ages are shown in blue (black lines show the distribution peak while dark and light shaded regions represent 68 per cent and 95 per cent confidence intervals about this peak, respectively) and the fraction of the population which have had 3:1 or 10:1 mergers are plotted in red (nearly 100 per cent in all cases). The τ_{form} and τ_{merge} recovery times obtained in Section 3.1 are indicated with dotted orange lines. Our expectation is that haloes existing above the indicated recovery times for all three cases should be relaxed.

We can see from this figure that the distribution of all three dynamical ages evolves very little from $z = 15$ to $z = 5$. Over this redshift range, the distribution of formation ages is very narrow and peaked very close to our formation recovery time-scale of $\tau_{\mathrm{form}} = 1.5$. The near constant value of τ_{form} during this epoch is consistent with early mass accretion histories which are exponential, as found previously by several other authors (e.g. Wechsler et al. 2002; McBride, Fakhouri & Ma 2009; Correa et al. 2015). The distribution of times since 3:1 mergers is much broader and is also peaked near our merger recovery time-scale $\tau_{\mathrm{merge}} = 2$. This tells us that typical haloes at high redshifts across all galactic masses are doubling their mass on timeframes that only barely permit relaxation while simultaneously, major mergers are occurring at rates which only barely permit recovery between events. The situation is importantly different for minor mergers. In this case we find that haloes experience minor mergers at rates which are much too rapid (on average) to permit dynamical relaxation between events.

In Fig. 4, we also plot the low-redshift evolution of our three dynamical ages as found in the *GiggleZ-HR* simulation. We note that the distributions compare well between Tiamat and *GiggleZ-HR* at $z = 5$ despite slightly different cosmologies and the somewhat higher masses depicted by *GiggleZ-HR*, validating its use for qualitative comparisons. Here we find that the narrow distribution of formation ages, broad distribution of merger ages and short times between 10:1 mergers persists almost unchanged until approximately $z \sim 2$. At this time, we find that $M_{\mathrm{vir}} = 10^{12}$–$10^{12.5}\ h^{-1}\ \mathrm{M}_\odot$ haloes begin to become progressively older, as typical formation ages and times since major mergers increase and times since minor mergers creep above merger recovery times of $\tau_{\mathrm{merge}} = 2$ by $z = 0$. The increase of τ_{form} at lower redshifts corresponds to a transition in these haloes' mass accretion histories from an exponential form to a linear form, as discussed already in the literature (e.g. McBride et al. 2009; Correa et al. 2015).

The fractions of haloes which meet these recovery criteria as a function of redshift is presented explicitly in Fig. 5. Here we see more clearly (in blue) the trends we identified from Fig. 4: the disappearance of haloes with formation times less than $\tau_{\mathrm{form}} = 1.5$ and the sustained low levels of haloes having had sufficient time to recover from their most recent mergers. We have added to these plots (in red) the fraction of haloes that simultaneously satisfy our standard x_{off}, φ and f_{sub} relaxation criteria. Remarkably, the fraction of relaxed haloes and the fraction having had sufficient time to recover from their last 10:1 (or larger) merger are very similar across a wide range of masses and redshifts. We conclude from this that the standardized relaxation criteria of Neto et al. (2007) are effectively identifying systems that have been disturbed by 10:1 (or larger) mergers. It should be noted however that our recovery criteria of $\tau_{\mathrm{form}} = 1.5$ and $\tau_{\mathrm{merge}} = 2$ have been calibrated at high redshift and may need adjustment at low redshift, where haloes are substantially more concentrated and the orbital properties of merging systems are significantly different, with more circular orbits requiring longer to relax. This is likely the reason why our estimates of the recovered

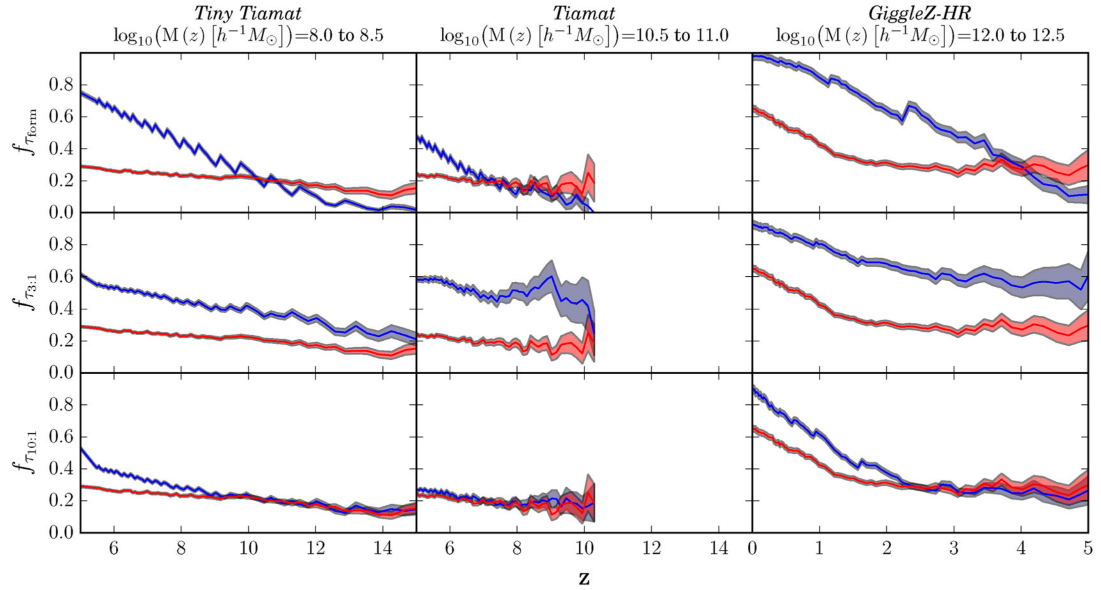

Figure 5. Redshift dependence of the fraction of FoF haloes meeting our recovery (in blue) and relaxation (in red) criteria (black lines show the fraction of the population while shaded regions depict Poisson uncertainties) for two cases at $z > 5$ spanning a range of 1000 in mass ([left] $M_{\mathrm{vir}} = 10^{8}$–$10^{8.5}\ h^{-1}\ \mathrm{M}_\odot$ from *Tiny Tiamat* and [middle] $M_{\mathrm{vir}} = 10^{10.5}$–$10^{11}\ h^{-1}\ M_\odot$ from Tiamat) contrasted with [right] the $z < 5$ evolution of $M_{\mathrm{vir}} = 10^{12}$–$10^{12.5}$ FoF haloes from the *GiggleZ-HR* simulation. In all cases, the recovered fraction roughly follows the fraction of haloes that have been able to recover from their last 10:1 (or larger) merger, suggesting that minor mergers are principally responsible for regulating the relaxed fraction of a halo population. Deviations between the two at $z < 2$, when growth stops being exponential (indicated by τ_{form}∼constant), may reflect important changes at low redshift to the processes of dynamical recovery.

fraction (measured using criteria calibrated at $z > 5$) exceeds the relaxed population (as measured directly from x_{off}, φ and f_{sub}) at $z < 2$.

Taken together, we see that at high redshifts ($z > 5$), the fraction of relaxed haloes drops to levels of ∼20 per cent at all galactic masses. Combined with the rapid decline in the number density of haloes with redshift at this time, we conclude that the abundance of relaxed galactic haloes prior to the EoR drops to very low levels. This should make it very challenging to assemble large populations of relaxed haloes at $z > 10$, which is of particular concern for studies seeking to understand the processes acting to establish universal density profiles for collisionless systems at high redshift.

3.3 Large and long-lived phase-space structures

As an exercise during the development phase of the Tiamat simulations, we analysed one of our simulations (*Tiny Tiamat-W07*) with the ROCKSTAR halo finder, allowing us to study the effect of halo finding on our semi-analytic modelling campaign. Doing so has yielded an interesting new insight into the dynamical lives of high-redshift galactic haloes.

In Fig. 6, we depict a relatively massive but otherwise typical halo at $z = 5$ extracted from the *Tiny Tiamat-W07* simulation. In this figure, the whole FoF system (as reported by ROCKSTAR) is depicted in the top panel while subsequent panels depict the configuration-space (left) and velocity-space structure (right) of the most massive substructure (in yellow) and the second most massive substructure (in cyan) as determined by SUBFIND (middle) and ROCKSTAR (bottom). There is a stark difference between the results from these two halo finders. We can see clearly – in a manner we find to be shared by all massive haloes at high redshift – that while the halo appears relatively undisturbed with unremarkable substructure, it in fact consists primarily of two very massive subhaloes which are distinct in phase space.

Phase-space halo finders such as ROCKSTAR are of course designed to separate halo substructures in this way, but it is not entirely clear that this is a desired result for applications in galaxy formation modelling. While approaches differ in detail, the central premise of all semi-analytic galaxy formation models is that the total matter assembly provided by their merger tree inputs can be reliably mapped to a faithful description of the baryonic assembly of galactic haloes. Problems may arise if the collisional fluids (particularly the hot haloes) associated with multiple collisionless systems oscillating through each other for >3 dynamical times cannot follow the collisionless material of their initial hosts. Substantial amounts of this gas will be stripped or rapidly coalesce into one hot halo, loosing its association with its original collisionless component while that material continues to orbit. This is the case with the Bullet Cluster for instance, albeit at a different mass scale and redshift. It is also the situation modelled by McCarthy et al. (2008) who find that the stripping of a galaxy's hot halo (due to tides, ram pressure stripping and hydrodynamic instabilities) is extremely efficient up to and during its first pericentric passage (i.e. ∼1 dynamical time following accretion). The amount of material removed varies with halo mass, concentration and orbit, but is substantial and typically in the range of 60–80 per cent for the broad range of cases they examine.

If such structures were short lived, the impact on our galaxy formation model would likely be insignificant. However, they are in fact long lived in dynamical terms. Following the format of Fig. 2, Fig. 7 presents a comparison of the evolving substructure fractions of FoF haloes extracted from *Tiny Tiamat-W07* using SUBFIND to those obtained from ROCKSTAR as functions of the dynamical ages τ_{form},$\tau_{3:1}$ and $\tau_{10:1}$. While we see the familiar decline of f_{sub} following formation and mergers presented in Fig. 2 in the SUBFIND trees, the ROCKSTAR trees exhibit a much slower decline, reaching constant

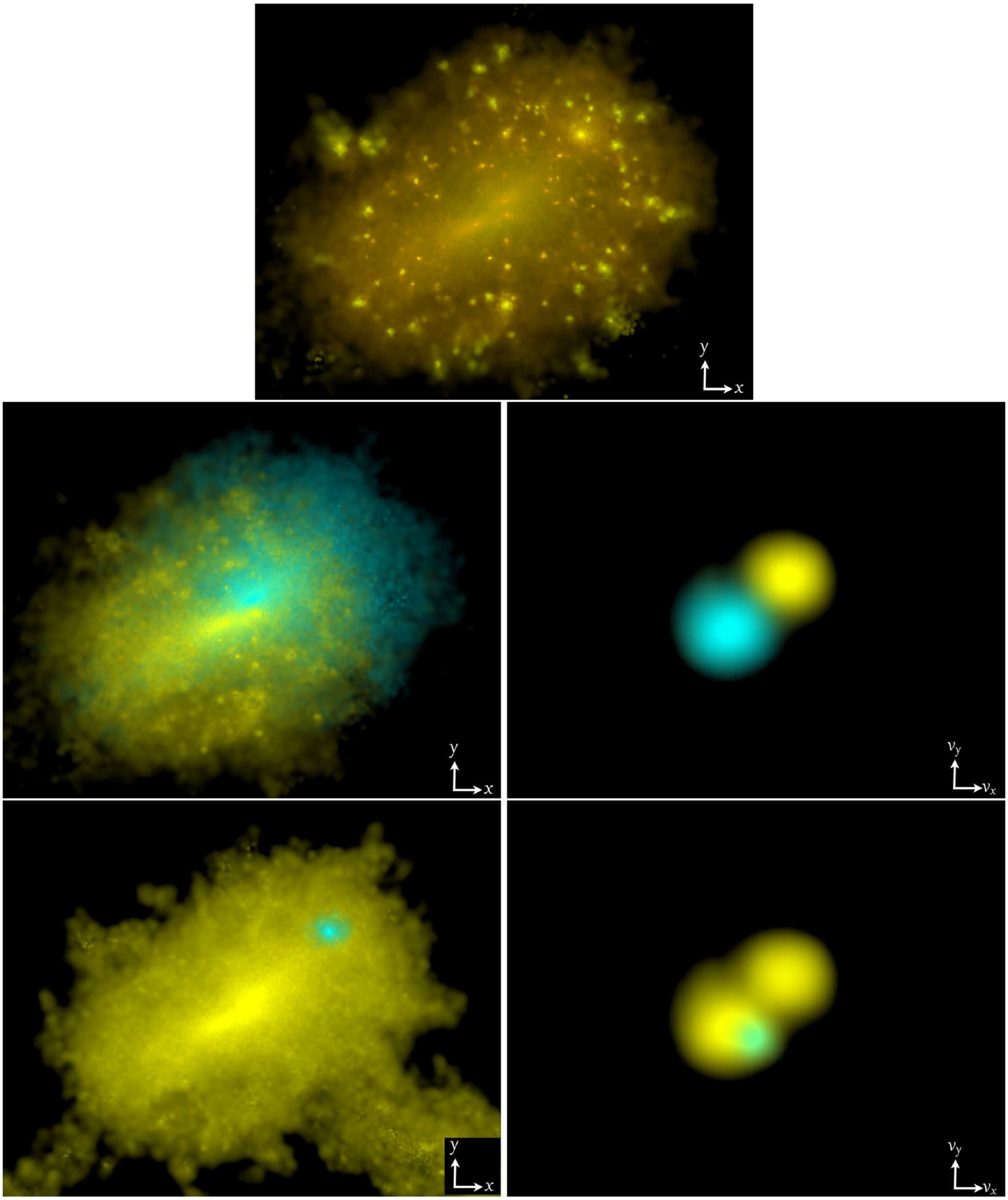

Figure 6. A $M_{\rm vir} = 4.4 \times 10^{10}\,h^{-1}\,{\rm M}_{\odot}$ halo at $z = 5$ in the *Tiny Tiamat-W07* simulation. The top panel is a configuration-space rendering of the FoF halo as identified by ROCKSTAR with colours set by the column weighted particle velocity dispersion. The bottom four panels depict the [left] real-space and [right] velocity-space structure of the two most massive substructures (yellow for the most massive, cyan for the second most massive) identified in this FoF halo by [middle] ROCKSTAR and [bottom] SUBFIND. Luminance is set by the logarithm of the integrated column through the system in all cases. In the case of the cyan structure in the SUBFIND case, luminance has been arbitrarily increased by a factor of 10 to increase its contrast. Fields of view for all configuration and velocity space images are 200 h^{-1}kpc (comoving) and 1000 km s^{-1}, respectively. Large distinct substructures such as the ones identified here by ROCKSTAR are common and long lived at high redshifts.

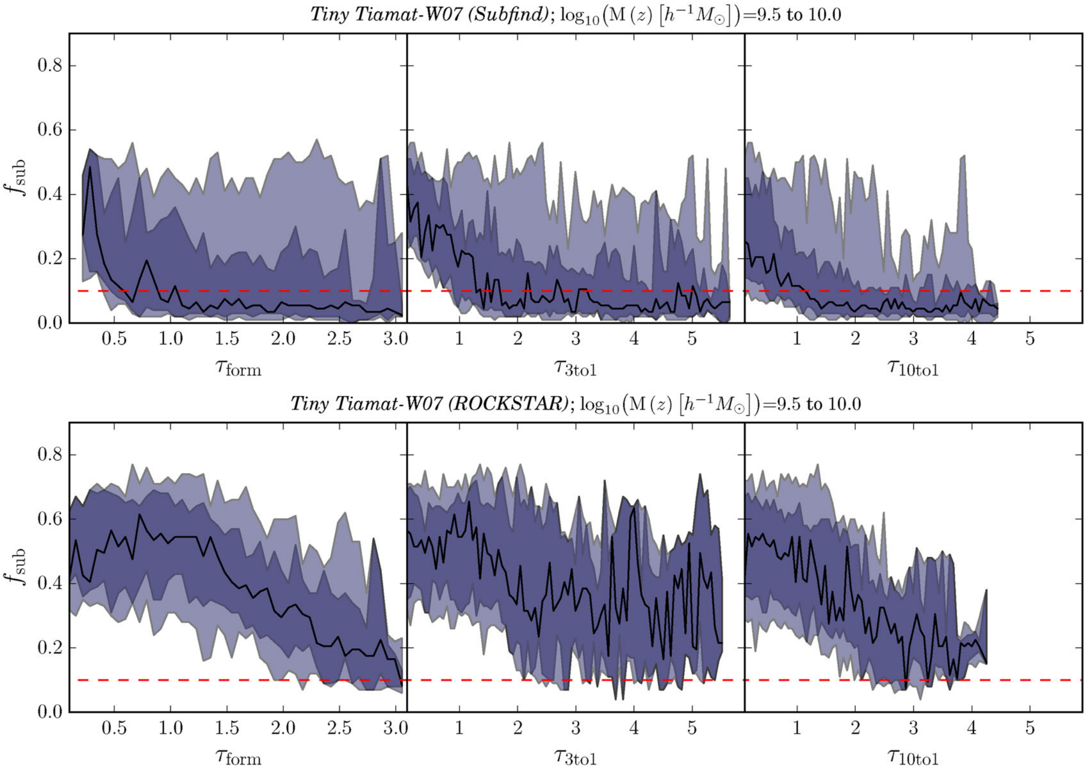

Figure 7. Comparison of the evolution of substructure fractions (f_{sub}) derived from [top] SUBFIND and [bottom] ROCKSTAR, in a well resolved and populated mass range in the *Tiny Tiamat-W07* simulation as a function of time since a FoF halo's progenitor line: [left] achieved 50 per cent of its current mass (τ_{form}), [mid] last experienced a 3:1 (or larger) merger ($\tau_{3:\,1}$) or [right] last experienced a 10:1 (or larger) merger ($\tau_{10:\,1}$). Black lines show the distribution peak while dark and light shaded regions represent 68 per cent and 95 per cent confidence intervals about this peak, respectively. For a halo at redshift z, times are measured in units of its dynamical time, taken to be 10 per cent of the Hubble time at z (in this system, the Hubble time is always $\tau = 10$). ROCKSTAR haloes have much higher substructure fractions due to the longer recovery times required for major mergers to lose their identity in phase space as opposed to the shorter times required to lose their identity in configuration space.

levels only after $\tau \gtrsim 5$, sustaining levels well above our standard relaxation criteria even after that.

On the other hand, if these substructures were rare, their impact on galaxy formation modelling would again be minimal. They are in fact very common. To illustrate their prevalence and to quantify the magnitude of this effect, we present the substructure fractions of our SUBFIND and ROCKSTAR $z = 5$ *Tiny Tiamat-W07* halo catalogues in Fig. 8. While the FoF halo mass functions for the two catalogues are virtually identical (except at the highest masses where the larger linking length used by ROCKSTAR unsurprisingly yields more systems, presumably due to overlinking), the substructure fractions at the highest (and most resolved) masses of the two catalogues are very different. Substructure fractions are 50–60 per cent at the highest masses in ROCKSTAR indicating that only around half of the mass in these systems is assigned to the most massive component of the system. This is a consequence of a very different splitting of the top level of the FoF group's substructure hierarchy.

Suggestions of this effect can be seen in the recent work of Behroozi et al. (2015). While these authors find that substructure properties like position and velocity generally agree between configuration and phase-space halo finders, they find that substantial differences in masses can occur. They also find strong disagreements in the frequency and duration of major mergers, particularly at redshifts $z > 1$. These differences are likely related to the situation presented in Fig. 6.

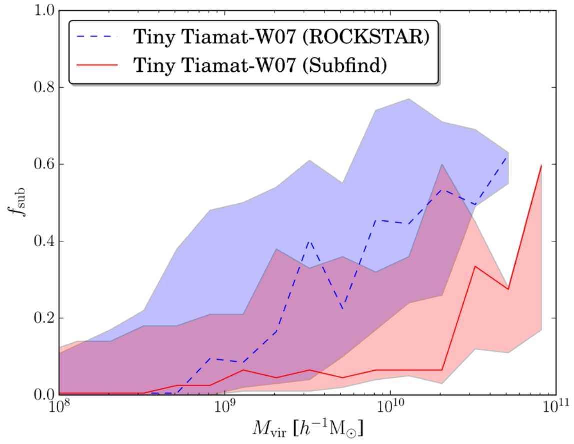

Figure 8. Substructure fractions in the *Tiny Tiamat-W07* simulation as measured by ROCKSTAR (in blue) and SUBFIND (in red) at $z = 5$. Lines (solid for SUBFIND, dashed for ROCKSTAR) depict the peak of the distribution while shaded regions depict the 68 per cent confidence interval about this peak. The higher substructure fractions in ROCKSTAR are a product of long-lived phase-space structures resulting from major mergers which dissolve quickly (in a couple dynamical times) in configuration space but retain their identity in velocity space.

We emphasize that we make no attempt here to advocate for one halo finding approach over another. Rather, we seek to make the point that care should be taken – particularly at high redshift where major mergers are frequent and the sorts of large, diffuse phase-space structures we illustrate in Fig. 6 are likely most prevalent – to ensure that each semi-analytic model is matched, in a physically meaningful way, to the nature of the substructure hierarchy supplied by the halo finder contributing to its input. Such differences may lead (once tuneable parameters are adjusted to yield accurate fits to observations) to significant systematics with mass in the evolution of merger trees which could masquerade as physical processes as diverse as mass dependances in dust properties, photon escape fractions, feedback and cooling. A detailed account of how the cooling and feedback modelling of DRAGONS (using MERAXES with trees derived from SUBFIND halo finding) compares to the *Smaug* hydrodynamic simulations of Duffy et al. (2014) will be presented in Paper VII, where a direct halo-by-halo comparison of the two methodologies will be presented.

We take this opportunity to point out one other possible important astrophysical consequence of large bulk phase-space structures such as this. Recent studies have begun to investigate the possibility that heating from dark matter annihilation may be observable in the redshifted 21 cm background from $z > 30$ (e.g. Furlanetto, Oh & Pierpaoli 2006a; Evoli, Mesinger & Ferrara 2014; Mack 2014; Schon et al. 2015). If phase-space structures such as these prove to be common at this epoch, important changes to inferred annihilation cross-sections may result.

4 SUMMARY AND CONCLUSIONS

We have introduced the DRAGONS programme and presented the Tiamat collisionless *N*-body simulation suite upon which it is constructed. The abundance of FoF structures populating Tiamat is a good match to the 'universal' model proposed by Watson et al. (2013) at high masses, but we find a suppression of low-mass systems, possibly due to differences in our halo finding procedure or perhaps indicating a deviation from 'universal' behaviour, at least at large redshifts.

Using Tiamat we have also illustrated the dynamically violent conditions experienced by galactic haloes at large redshift. We find that across a wide range of galactic mass ($M_{\rm vir} = 10^8$–$10^{11}\,h^{-1}\,{\rm M}_{\odot}$) above $z = 5$, haloes relax from their formation and from mergers in essentially the same way and in the same amount of time: within one and a half dynamical times in the case of their formation (i.e. $\tau_{\rm form} = 1.5$) and within two dynamical times following mergers involving a primary and a secondary larger than 10 per cent of its mass (i.e. $\tau_{\rm merge} = 2$).

The distribution of formation times and times since major mergers maintain approximately these time-scales across all redshifts above $z = 5$ while the time between minor mergers is typically significantly less. Relaxed fractions maintain levels of less than 20 per cent at $z > 5$ as a result. Using the *GiggleZ-HR* simulation (which extends to $z = 0$, albeit with poorer resolution) we find that this remains true for $M_{\rm vir} = 10^{12}\,h^{-1}\,{\rm M}_{\odot}$ haloes until $z \sim 2$. It appears that the rate of minor mergers principally regulate a halo population's relaxed fraction, as measured by standard metrics. Combined with the rapid decline of the halo mass function at redshifts $z > 10$, the abundance of relaxed haloes prior to the EoR must be extremely low.

Using the phase-space halo finder ROCKSTAR, we also demonstrate that high-redshift haloes host large and long-lived substructures that go undetected to halo finders such as SUBFIND which utilize configuration-space information only. This results in substructure fractions that are much higher for ROCKSTAR than for SUBFIND, with probable implications for semi-analytic models of galaxy formation at high redshift.

Taken together, these results illustrate the dynamically violent circumstances under which galaxy formation proceeds in the early Universe. The consequences are many and significant, including implications for photon escape fractions, efficiencies of feedback from winds (both stellar and active galactic nucleus) and the efficiency of spheroid assembly. These in turn can have important consequences for the reionization history of the Universe during the EoR and observed galaxy sizes.

ACKNOWLEDGEMENTS

This research was supported by the Victorian Life Sciences Computation Initiative (VLSCI), grant ref. UOM0005, on its Peak Computing Facility hosted at the University of Melbourne, an initiative of the Victorian Government, Australia. Part of this work was performed on the gSTAR national facility at Swinburne University of Technology. gSTAR is funded by Swinburne and the Australian Governments Education Investment Fund. This research programme is funded by the Australian Research Council through the ARC Laureate Fellowship FL110100072 awarded to JSBW. CP acknowledges support of ARC DP130100117 and DP140100198, and ARC Future Fellowship FT130100041. AM acknowledges support from the European Research Council (ERC) under the European Unions Horizon 2020 research and innovation programme (grant agreement no. 638809 AIDA). We thank Volker Springel for making his GADGET-2 and SUBFIND codes available to us and Peter Behroozi for making his ROCKSTAR code available. We also thank N. Gnedin for several very useful comments on our manuscript.

REFERENCES

Angel P. W., Poole G. B., Ludlow A. D., Duffy A. R., Geil P. M., Mutch S. J., Mesinger A., Wyithe J. S. B., 2016, MNRAS, 459, 2106 (Paper II)
Barkana R., 2009, MNRAS, 397, 1454
Behroozi P. S., Wechsler R. H., Wu H.-Y., 2013, ApJ, 762, 109
Behroozi P. et al., 2015, MNRAS, 454, 3020
Bhattacharya S., Heitmann K., White M., Lukić Z., Wagner C., Habib S., 2011, ApJ, 732, 122
Correa C. A., Wyithe J. S. B., Schaye J., Duffy A. R., 2015, MNRAS, 450, 1521
Crocce M., Pueblas S., Scoccimarro R., 2006, MNRAS, 373, 369
Croft R. A. C., Altay G., 2008, MNRAS, 388, 1501
Diemer B., Kravtsov A. V., 2015, ApJ, 799, 108
Dijkstra M., Haiman Z., Rees M. J., Weinberg D. H., 2004, ApJ, 601, 666
Duffy A. R., Wyithe J. S. B., Mutch S. J., Poole G. B., 2014, MNRAS, 443, 3435
Dutton A. A., Macciò A. V., 2014, MNRAS, 441, 3359
Evoli C., Mesinger A., Ferrara A., 2014, J. Cosmol. Astropart. Phys., 11, 024
Fan X. et al., 2006, AJ, 132, 117
Furlanetto S. R., Zaldarriaga M., Hernquist L., 2004, ApJ, 613, 1
Furlanetto S. R., Oh S. P., Pierpaoli E., 2006a, Phys. Rev. D, 74, 103502
Furlanetto S. R., Oh S. P., Briggs F. H., 2006b, Phys. Rep., 433, 181
Geil P. M., Wyithe J. S. B., 2008, MNRAS, 386, 1683
Geil P. M., Mutch S. J., Poole G. B., Angel P. W., Duffy A. R., Mesinger A., Wyithe J. S. B., 2015, preprint (arXiv:1512.00564) (Paper V)
Gunn J. E., Peterson B. A., 1965, ApJ, 142, 1633
Hellwing W. A., Frenk C. S., Cautun M., Bose S., Helly J., Jenkins A., Sawala T., Cytowski M., 2016, MNRAS, 457, 3492

Iliev I. T., Mellema G., Shapiro P. R., Pen U.-L., 2007, MNRAS, 376, 534

Iliev I. T., Mellema G., Ahn K., Shapiro P. R., Mao Y., Pen U.-L., 2014, MNRAS, 439, 725

Jenkins A., Frenk C. S., White S. D. M., Colberg J. M., Cole S., Evrard A. E., Couchman H. M. P., Yoshida N., 2001, MNRAS, 321, 372

Knebe A. et al., 2013, MNRAS, 435, 1618

Komatsu E. et al., 2009, ApJS, 180, 330

Komatsu E. et al., 2011, ApJS, 192, 18

Lee K.-G., Cen R., Gott J. R., Trac H., 2008, ApJ, 675, 8

Lewis A., Challinor A., Lasenby A., 2000, ApJ, 538, 473

Liu C., Mutch S. J., Angel P. W., Duffy A. R., Geil P. M., Poole G. B., Mesinger A., Wyithe J. S. B., 2015, preprint (arXiv:1512.00563) (Paper IV)

Ludlow A. D., Navarro J. F., Angulo R. E., Boylan-Kolchin M., Springel V., Frenk C., White S. D. M., 2014, MNRAS, 441, 378

Lukić Z., Heitmann K., Habib S., Bashinsky S., Ricker P. M., 2007, ApJ, 671, 1160

McBride J., Fakhouri O., Ma C.-P., 2009, MNRAS, 398, 1858

McCarthy I. G., Frenk C. S., Font A. S., Lacey C. G., Bower R. G., Mitchell N. L., Balogh M. L., Theuns T., 2008, MNRAS, 383, 593

Mack K. J., 2014, MNRAS, 439, 2728

McQuinn M., Lidz A., Zahn O., Dutta S., Hernquist L., Zaldarriaga M., 2007, MNRAS, 377, 1043

Mesinger A., Furlanetto S., 2007, ApJ, 669, 663

Morales M. F., Wyithe J. S. B., 2010, ARA&A, 48, 127

Murray S. G., Power C., Robotham A. S. G., 2013, MNRAS, 434, L61

Mutch S. J., Geil P. M., Poole G. B., Angel P. W., Duffy A. R., Mesinger A., Wyithe J. S. B., 2015, preprint (arXiv:1512.00562) (Paper III)

Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493

Neto A. F. et al., 2007, MNRAS, 381, 1450

Ouchi M. et al., 2010, ApJ, 723, 869

Planck Collaboration XIII, 2015, preprint (arXiv:1502.1589)

Poole G. B., Fardal M. A., Babul A., McCarthy I. G., Quinn T., Wadsley J., 2006, MNRAS, 373, 881

Poole G. B. et al., 2013, MNRAS, 429, 1902

Poole G. B. et al., 2015, MNRAS, 449, 1454

Power C., Navarro J. F., Jenkins A., Frenk C. S., White S. D. M., Springel V., Stadel J., Quinn T., 2003, MNRAS, 338, 14

Power C., Knebe A., Knollmann S. R., 2012, MNRAS, 419, 1576

Prada F., Klypin A. A., Cuesta A. J., Betancort-Rijo J. E., Primack J., 2012, MNRAS, 423, 3018

Press W. H., Schechter P., 1974, ApJ, 187, 425

Qin D. et al., 2016, MNRAS, in press (Paper VII)

Reed D. S., Bower R., Frenk C. S., Jenkins A., Theuns T., 2007, MNRAS, 374, 2

Schon S., Mack K. J., Avram C. A., Wyithe J. S. B., Barberio E., 2015, MNRAS, 451, 2840

Shin M.-S., Trac H., Cen R., 2008, ApJ, 681, 756

Springel V., 2005, MNRAS, 364, 1105

Springel V., White S. D. M., Tormen G., Kauffmann G., 2001, MNRAS, 328, 726

Thomas P. A. et al., 1998, MNRAS, 296, 1061

Tinker J., Kravtsov A. V., Klypin A., Abazajian K., Warren M., Yepes G., Gottlöber S., Holz D. E., 2008, ApJ, 688, 709

Watson W. A., Iliev I. T., D'Aloisio A., Knebe A., Shapiro P. R., Yepes G., 2013, MNRAS, 433, 1230

Wechsler R. H., Bullock J. S., Primack J. R., Kravtsov A. V., Dekel A., 2002, ApJ, 568, 52

Wyithe J. S. B., Morales M. F., 2007, MNRAS, 379, 1647

Zahn O., Lidz A., McQuinn M., Dutta S., Hernquist L., Zaldarriaga M., Furlanetto S. R., 2007, ApJ, 654, 12

APPENDIX A: MASS FUNCTION FITTING

In this section, we present additional details regarding the results of the FoF halo mass function fitting that we present and discuss in Section 2.3. We seek to test the parametrized universal FoF halo mass function presented by Watson et al. (2013) with the Tiamat simulation suite. This mass function follows the convention introduced by Jenkins et al. (2001, see Lukić et al. 2007 for a good review) whereby $n(M, z)$, the number density of FoF haloes with mass M at redshift z, is separated into a 'scaled-mass function' component $f(\sigma, z)$ expected to be independent with redshift (as predicted by the analytic theory of Press & Schechter 1974, and its extensions) and terms which encapsulate the linear growth of the matter density field

$$\frac{\mathrm{d}n}{\mathrm{d}\log M} = \frac{\rho_b}{M} f(\sigma, z) \frac{\mathrm{d}\ln\sigma^{-1}}{\mathrm{d}\log M}, \tag{A1}$$

where $\rho_b(z)$ is the mean background matter density as a function of redshift and σ is the variance of the linearly evolved matter density field smoothed with a spherical tophat on a scale R encompassing (on average) the halo mass M (i.e. $R{=}3M/[4\pi\rho_b(z)]^{1/3}$).

As simulations have become increasingly more precise and halo finding approaches increasingly diverse, a large number of parametrizations for $f(\sigma, z)$ have been proposed with varying degrees of complexity and redshift evolution. A 'universal' (i.e. redshift-independent) form for the mass function of structures extracted with the FoF algorithm was presented by Watson et al. (2013) with the form

$$f(\sigma) = A\left(\frac{\beta}{\sigma}+1\right)^{\alpha} e^{-\gamma/\sigma^2}, \tag{A2}$$

where A and β are effectively two 'normalization parameters' and α and γ are effectively two 'shape parameters'. We have fit this functional form to the mass functions presented in Fig. 1 extracted from Tiamat,*Tiny Tiamat* and *Medi Tiamat* at redshifts $z = 5$, 7.5, 10, 15 and 24. To do so we have used the implementation of the Metropolis–Hastings MCMC algorithm first presented in Poole et al. (2013) and utilized in several studies since. We have sought to minimize systematic differences in our comparison to the Watson et al. (2013) fitting results. We have thus followed their approach and restricted our fit to haloes with more than 1000 particles. We have also corrected for finite box size effects using the method employed by Lukić et al. (2007) and Bhattacharya et al. (2011, whereby fluctuations on scales larger than the box size are excluded from the tophat filtering calculation which maps halo mass to the variance of the matter density field), used the expression for 1σ Poisson uncertainties introduced by Watson et al. (2013)

$$\sigma_{\pm} = \sqrt{N+\frac{1}{4}}\pm\frac{1}{2} \tag{A3}$$

for computing the χ^2 likelihoods that we use for the fitting and applied a slight resolution correction to our FoF masses of the form

$$M_{\mathrm{FoF}} = m_p N_p\left(1 - N_p^{-0.6}\right), \tag{A4}$$

where N_p is the number of particles in the FoF halo and m_p is the particle mass.

We present the posterior distribution function (PDF) which emerges from this fit in Fig. A1 and compare the parametrization of equation (A2), as obtained across all redshifts by Watson et al. (2013), to that obtained here for $z \geq 5$. A table of fitted values and their covariance are presented in Table A1.

At the low-mass (high-σ) end of the fit shown in Fig. A2, we find some significant differences between the Tiamat results and the predictions of Watson et al. (2013) but good agreement otherwise. Why then are our best-fitting parameters so different? We can see from Fig. A1 that very strong degeneracies exist between the

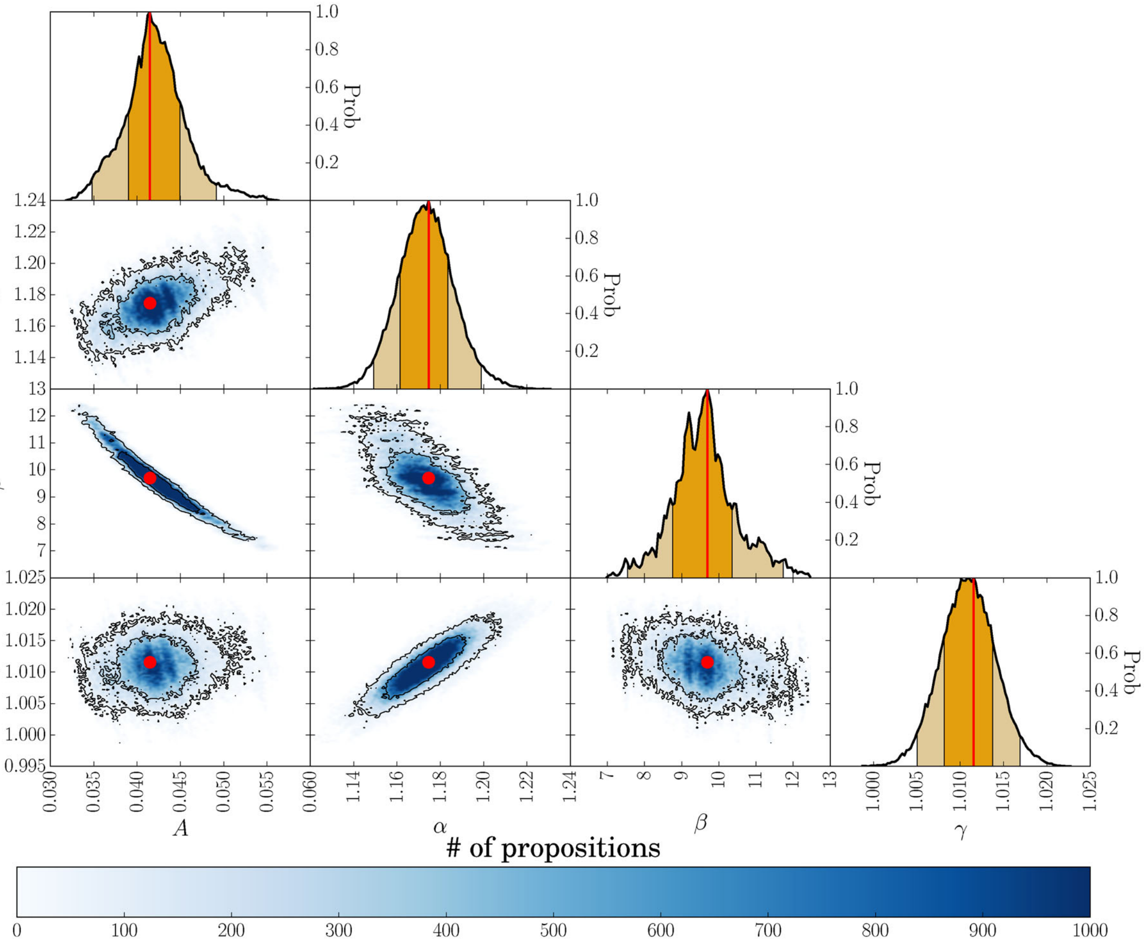

Figure A1. One and two-dimensional projections of the posterior distribution function (PDF) of our MCMC fit of the parametrized universal FoF halo mass function presented by Watson et al. (2013, see equation A2) to the FoF halo mass functions extracted from the Tiamat simulation suite. Blue-scale images show the PDF in terms of the number of propositions used to sample it, with black contours showing 68 per cent and 95 per cent confidence regions. Dark and light orange shaded regions show the 68 per cent and 95 per cent confidence intervals of the one-dimensional projections, respectively. Red points and lines mark the most probable location in this parameter space, which we quote as the best-fitting parameters of our fit.

Table A1. Best-fitting values for and covariance between the Watson et al. (2013) mass function parameters as fit to SUBFIND haloes extracted from the Tiamat suite of *Planck*-2015 EoR simulations.

Parameter	Watson fit	Tiamat fit	covariance with A	covariance with α	covariance with β	covariance with γ
A	0.282	0.0333	6.89×10^{-5}	1.65×10^{-4}	-1.93×10^{-2}	1.55×10^{-5}
α	2.163	1.153	1.65×10^{-4}	5.10×10^{-4}	-4.77×10^{-2}	6.82×10^{-5}
β	1.406	12.33	-1.93×10^{-2}	-4.77×10^{-2}	5.80	-4.81×10^{-3}
γ	1.210	1.01	1.55×10^{-5}	6.82×10^{-5}	-4.81×10^{-3}	1.29×10^{-5}

normalization parameters (A and β) and shape parameters (α and γ) of equation (A2). Taking the shape parameters first, the suppression of low-mass systems requires a reduction of α to flatten the slope and an adjustment to A to alter the low-mass normalization. In the case of the normalization parameters, their product directly sets the normalization of $f(\sigma)$ at the low-σ end. The product of these two parameters is strongly constrained by the data, as illustrated by the $A\beta$ ∼constant form of the PDF in Fig. A1. Given the adjustments discussed above which are needed to fit the high-σ end of the function, our fit becomes pushed to a different part of this degeneracy. While our best-fitting parameters for A and β are very different from the Watson et al. (2013) values, the products are very similar with our fit yielding $A\beta = 0.407$ and the Watson et al. (2013) values yielding $A\beta = 0.396$. This reflects very similar results between the two studies at the high-mass (or low-σ) end, despite these very different fit values.

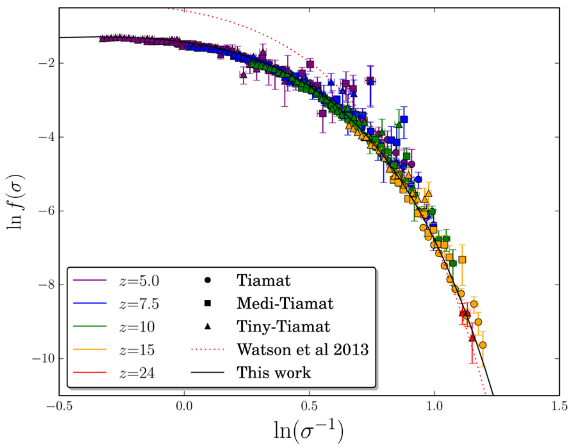

Figure A2. Comparison of the scaled mass function of Watson et al. (2013) to data from Tiamat (shown by the coloured data points, following the same format as Fig. 1) and to the parametrization derived from Tiamat and presented in this work.

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