EDGE: a new approach to suppressing numerical diffusion in adaptive mesh simulations of galaxy formation

Andrew Pontzen,[1]★ Martin P. Rey [ID],[1,2] Corentin Cadiou [ID],[1] Oscar Agertz [ID],[2] Romain Teyssier [ID],[3] Justin Read [ID][4] and Matthew D. A. Orkney[4]

[1]*Department of Physics and Astronomy, University College London, London WC1E 6BT, UK*
[2]*Lund Observatory, Department of Astronomy and Theoretical Physics, Lund University, Box 43, SE-221 00 Lund, Sweden*
[3]*Institute for Computational Science, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland*
[4]*Department of Physics, University of Surrey, Guildford GU2 7XH, UK*

ABSTRACT

We introduce a new method to mitigate numerical diffusion in adaptive mesh refinement (AMR) simulations of cosmological galaxy formation, and study its impact on a simulated dwarf galaxy as part of the 'EDGE' project. The target galaxy has a maximum circular velocity of $21\,\mathrm{km\,s^{-1}}$ but evolves in a region that is moving at up to $90\,\mathrm{km\,s^{-1}}$ relative to the hydrodynamic grid. In the absence of any mitigation, diffusion softens the filaments feeding our galaxy. As a result, gas is unphysically held in the circumgalactic medium around the galaxy for $320\,\mathrm{Myr}$, delaying the onset of star formation until cooling and collapse eventually triggers an initial starburst at $z = 9$. Using genetic modification, we produce 'velocity-zeroed' initial conditions in which the grid-relative streaming is strongly suppressed; by design, the change does not significantly modify the large-scale structure or dark matter accretion history. The resulting simulation recovers a more physical, gradual onset of star formation starting at $z = 17$. While the final stellar masses are nearly consistent ($4.8 \times 10^6\,\mathrm{M_\odot}$ and $4.4 \times 10^6\,\mathrm{M_\odot}$ for unmodified and velocity-zeroed, respectively), the dynamical and morphological structure of the $z = 0$ dwarf galaxies are markedly different due to the contrasting histories. Our approach to diffusion suppression is suitable for any AMR zoom cosmological galaxy formation simulations, and is especially recommended for those of small galaxies at high redshift.

Key words: methods: numerical – galaxies: dwarf – cosmology: miscellaneous.

1 INTRODUCTION

Numerical hydrodynamics plays a central role in understanding the formation and evolution of galaxies, and their impact on the Universe (e.g. Somerville & Davé 2015). However, numerical methods for solving the hydrodynamical equations are required to trade accuracy in different regimes. The two most established methods, smoothed particle hydrodynamics (SPH) and adaptive mesh refinement (AMR) have well-documented, complementary strengths and weaknesses (Agertz et al. 2007; Springel 2010a; Teyssier 2015). While the 'subgrid' physics (i.e. differing approaches to incorporating unresolved astrophysical processes) dominates the uncertainty budget of galaxy formation simulations (Scannapieco et al. 2012; Kim et al. 2016), errors in hydrodynamics can also cause systematic shifts in galactic properties, so should be studied and suppressed to the greatest possible degree.

One approach is to combine the best of SPH and AMR worlds using a moving mesh (e.g. Springel 2010b; Weinberger, Springel & Pakmor 2020) or other hybrid solver (Hopkins 2015). On the other hand, the comparative immaturity of the resulting numerical integrators means that their properties are still relatively unexplored. Movement in the mesh is known to exacerbate numerical noise

associated with the grid itself, which can then act as a source of instabilities; the importance of this to physical solutions remains unclear (McNally, Lyra & Passy 2012; Muñoz et al. 2014). There are also strong practical and sociological reasons why groups choose to work with codes into which they have invested significant effort and where strengths and weaknesses are familiar.

In this paper, we will focus on the best-known weakness of AMR: its results are dependent on the rate at which the fluid moves relative to the grid, breaking Galilean invariance (e.g. Robertson et al. 2010). The effect of grid-relative motions is to cause numerical diffusion; such diffusion is in fact required to produce stable hydrodynamical solutions, so is in equal part a strength and weakness of successful schemes (Teyssier 2015). In some circumstances, the diffusion is negligible but if the entire region being simulated moves rapidly relative to the grid, diffusion may degrade the effective resolution to be many times worse than the grid spacing would suggest.

In this work, we will study diffusion in AMR simulations of dwarf galaxies, and discuss a new route to mitigating its effects that we initially designed for the EDGE[1] project (Agertz et al. 2020). The EDGE regime is particularly challenging: we target small galaxies (virial velocities $\simeq 20\,\mathrm{km\,s^{-1}}$), where the impact of numerical diffusion is most significant compared with the gravitational potential. While the

---

★ E-mail: a.pontzen@ucl.ac.uk

[1]Engineering dwarfs at galaxy formation's EDGE.

mean velocity in the cosmological box is zero by construction, the galaxies typically move at $\simeq 80\,\mathrm{km\,s^{-1}}$ relative to the grid, an order of magnitude higher than the internal velocity dispersions.

The rate of flow is only one factor in determining the extent of numerical diffusion; it is also sensitive to the order of the hydrodynamic solver, the cell size, and the time-step (e.g. Robertson et al. 2010). Since AMR demonstrably converges in the limit of infinitely small cells, the simplest route to minimising diffusion is to use an aggressive refinement scheme (e.g. Few et al. 2016); however, as we will see in this work, such an approach is computationally expensive in general. While most refinement schemes are quasi-Lagrangian, i.e. aiming to achieve roughly equal mass per grid cell, using a super-Lagrangian scheme (e.g. Chabanier et al. 2020) is one route to suppressing numerical diffusion in a more targeted (and therefore computationally cheaper) region. A separate option is to use higher-order slope estimation schemes; however, according to Godunov's theorem, only first-order linear schemes are monotonicity preserving, meaning higher-order solutions are oscillatory and may generate negative densities. Slope limiters, which smoothly reduce the order of the solver in regions with steep gradients, can be used to mitigate this problem at the expense of making the method more diffusive. There are many different possible limiters, each trading stability and diffusion in a different way (e.g. Toro 2009).

High-order slope estimation would be particularly challenging in our galaxy formation simulations. We model stellar feedback by directly modifying gas properties in cells surrounding young star particles, often resulting in sharp discontinuities within the density, velocity, and energy fields. We have chosen the MinMod slope limiter (e.g. Toro 2009) which is the safest choice in terms of suppressing instabilities, but consequently also the most diffusive. We find that, if we adopt a less diffusive slope limiter, the solution becomes unphysical; with the current implementation of feedback, we are hence required to adopt the MinMod slope limiter in order to maintain stability. In contrast, galaxy simulations accounting for strong stellar feedback via source terms in the Riemann solver, e.g. using a non-thermal pressure (Agertz et al. 2013; Kretschmer, Agertz & Teyssier 2020), have successfully been carried out using slope limiters such as MonCen (Agertz & Kravtsov 2015). We leave a more systematic investigation of this topic for future work, and instead focus here on a completely new solution to the diffusion problem.

If neither the solver nor the cell size can easily be changed, the final alternative is to ensure the rate of flow relative to the grid is small. However, haloes with a naturally small net velocity are exceptionally rare. For example, in the region used by the EDGE project, 95 per cent of all haloes move at velocities greater than $40\,\mathrm{km\,s^{-1}}$ at the time of early star formation, $z = 16$. As we will see, this is unavoidable in cosmological volumes, and already enough to cause substantive numerical errors. Further, adding a spatially uniform velocity offset to the initial conditions cannot usefully reduce even a single halo's net speed: the cosmological expansion 'redshifts' the momentum of particles. Any uniform velocity offset that is not associated with a potential gradient therefore decays proportionally to $1 + z$; in contrast, the physical velocities grow (during matter domination in the linear regime) as $(1 + z)^{-1}$. Therefore, uniform velocity offsets rapidly become insignificant compared to the growing physical velocities.

Our solution is to use the genetic modification approach (Roth, Pontzen & Peiris 2016; Rey & Pontzen 2018; Stopyra et al. 2020) to post hoc modify the initial conditions such that the velocity of a chosen halo is minimized, while its environment, accretion history and other properties are minimally affected. In fact, we have already adopted this approach as a matter of course for published EDGE

results (Rey et al. 2019, 2020). In this paper, we examine the effects of *disabling* the modification in one of our galaxies in order to compare the AMR solution with and without rapid streaming.

In Section 2, we describe the simulation and modification techniques. Section 3 presents the comparison between simulated galaxies with and without the velocity-zeroing in the initial conditions. We summarize and provide recommendations for future investigations in Section 4, taking into account the speeds at which galaxies stream in typical cosmological simulations. Appendix A gives details of an estimate for the scale of numerical diffusion in our simulations, while Appendix B1 calculates the expected streaming speeds in cosmological simulations as a function of box size.

2 GALAXY FORMATION SIMULATIONS

Initial conditions are generated using the GENETIC code that permits modifications to be made to a Gaussian random field realized at increasing resolution in nested zoom boxes (Stopyra et al. 2020). The parent box has 512^3 particles across a box length of $L = 50\,\mathrm{Mpc}$ comoving and is generated to match the cosmological parameters $\Omega_M = 0.309$; $\Omega_\Lambda = 0.691$; $\Omega_b = 0.045$; $H_0 = 67.77\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ (Planck Collaboration XVI 2014). Two zoom grids are opened with side lengths 12.5 and 6.25 Mpc, respectively, with the deepest level attaining a resolution equivalent to $32\,768^3$ particles on the base grid.

The zooms are selected in two stages. First, we simulate the uniform-resolution box with dark matter only, and visually identify a void that we resimulate at the equivalent of 2048^3 resolution. We then identify all dark matter haloes within the void at $z = 0$ using HOP (Eisenstein & Hut 1998) and search for candidate dwarf galaxies with no massive neighbours and a suitable virial mass $10^9 < M_{200c}/M_\odot < 10^{10}$. Here, M_{200c} is defined as the mass inside a spherical volume encompassing 200 times the cosmic critical density and we will also use the corresponding virial radius, r_{200c}. We consider only isolated objects, by which we mean those without an equally or more massive counterpart within 10 virial radii at $z = 0$; for more details see Agertz et al. (2020). In this work, we will focus on a single example with $M_{200c} \simeq 1.6 \times 10^9\,M_\odot$, which is the 'Reference' galaxy from Rey et al. (2019).

We perform an initial dark-matter-only zoomed simulation of our object. All simulations are performed using RAMSES (Teyssier 2002) and analysed with the combination of PYNBODY (Pontzen et al. 2013) and TANGOS (Pontzen & Tremmel 2018). At each output, we measure the halo circular velocity v_{max}, defined as the maximum of v_c where $v_c^2 = GM(<r)/r$, and $M(<r)$ is the mass enclosed in spheres r around the halo centre. We also measure the mass-weighted mean velocity of the Lagrangian region v_{region}, defined by all dark matter particles that will fall into the halo by $z = 0$.

The top two panels of Fig. 1 show these quantities as a function of time. The blue lines denote results from the original initial conditions; we find that the region streams at $51\,\mathrm{km\,s^{-1}}$ at $z = 16$ (a time at which we will find our hydrodynamical simulations start to form stars) and grows to a peak of $90\,\mathrm{km\,s^{-1}}$ at $z = 1$, compared to a halo maximum circular velocity of $21\,\mathrm{km\,s^{-1}}$. The halo's potential well is therefore shallow compared to the flow rate across the grid.

We now discuss how to suppress the streaming in the initial conditions. A slice through the unmodified zoom initial conditions for the chosen object is shown in Fig. 2. In the linear regime, the velocity field is uniquely determined by the overdensity field (through the continuity equation). Consequently, we may zero the mean velocity in the target Lagrangian region through a suitable change to the overdensity field. We follow the procedure described by Roth et al. (2016), which is a variant of the Hoffman & Ribak (1991) algorithm.
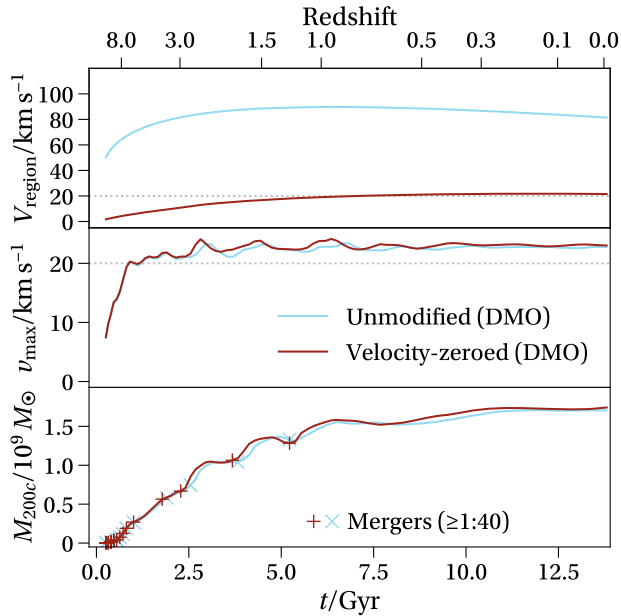
Figure 1. Lagrangian region streaming velocities relative to the simulation grid (top) contrasted with the maximum circular velocities (middle) for dark-matter-only simulations of the EDGE galaxy studied in this work. The horizontal dotted grey line on each panel shows $20\,\mathrm{km\,s^{-1}}$ for reference. The mean circular velocity of $21\,\mathrm{km\,s^{-1}}$ is more than four times smaller than the peak region streaming velocity of $90\,\mathrm{km\,s^{-1}}$ in the original run (dotted lines), raising our initial concern that numerical diffusion may have a significant effect on this object. Once our modification is activated, the region velocity is zeroed in the linear initial conditions. This leaves only a higher-order residual drift which averages $11\,\mathrm{km\,s^{-1}}$; it is even smaller at high redshift where the effects of diffusion are most important. The accretion history (bottom) remains unchanged – here we show, for both simulations, the virial mass M_{200c} as a function of time; we also mark the mergers with ratios at least 1:40 with crosses. The results from the two initial conditions are almost indistinguishable, because the modification to the velocity is by construction highly coherent across the scales from which the halo forms (Fig. 2).

Modifications made in this way change the field to the minimum extent necessary to obtain the desired alteration, while retaining a high likelihood within the cosmological Gaussian random ensemble. For more information, see van de Weygaert & Bertschinger (1996) for the relation between velocity and overdensity constraints; Roth et al. (2016) and Rey & Pontzen (2018) for the principles of modification and their close relationship to constraints; and Stopyra et al. (2020) for the details of our multigrid algorithm that implements the required modifications in practice.

The outcome of the operation is shown in the middle and right-hand panels of Fig. 2. The middle panel shows the change in the overdensity required to zero the linear velocity, exaggerated in colour scale by a factor 10 relative to the field in the left-hand panel. The right-hand panel shows the change in the *x*-component of the corresponding velocity field. These show how the algorithm generates a shift in the velocity by adding a very large-scale density gradient to the box. We refer to simulations with the modified initial conditions as 'velocity-zeroed', and we verified that the velocity of the Lagrangian region in the initial conditions is indeed precisely zero.

We perform a new dark-matter-only simulation, now starting from the updated initial conditions. Non-linear evolution implies that the halo is not perfectly stationary relative to the grid at late times, but the net motion is substantially suppressed. The typical streaming velocity in the modified simulation is $\simeq 10\,\mathrm{km\,s^{-1}}$ (peaking at $22\,\mathrm{km\,s^{-1}}$ by $z = 0$); see the red line in the upper panel of Fig. 1. It remains smaller than the circular velocity of the halo at all times, and is close to zero at high redshift.

The velocity modifications are extremely coherent across the Lagrangian patch of the target halo meaning that there is virtually no change in the formation history of our halo or the surrounding structures. The mass accretion histories in the original and velocity-zeroed cases are shown in the lower panel of Fig. 1; they match closely. We indicate all mergers more significant than 1:40 (using $\times$ symbols for mergers in the unmodified case, and $+$ in the velocity-zeroed case), and these also show good agreement in timing between the two versions of the simulation. The left-hand panels of Fig. 3 illustrate the large-scale structure, showing the projected dark matter density in the zoom region at $z = 15.7$; there is barely any change between the two cases. This lack of any major effects at the level of the dark matter simulation, coupled to the increased gas physics fidelity that we will describe in Section 3, has led us to adopt velocity-zeroed initial conditions as the default for other EDGE studies starting with Rey et al. (2019).

In addition to the numerical effects of grid-relative streaming, there is also a physical effect from streaming between baryons and dark matter (Tseliakhovich & Hirata 2010). This streaming can suppress the condensation of gas into small haloes at high redshift, which is a significant consideration in the formation of the first stars (Bromm 2013) but concerns earlier times and smaller haloes than we consider here. These physical relative velocities decay as the universe expands, and even at the start of our simulation ($z = 100$) have an r.m.s. of $\simeq 3\,\mathrm{km\,s^{-1}}$, far smaller than the velocities relevant to numerical diffusion; they have been neglected in this study.

We next perform three hydrodynamical simulations with velocity-zeroed and three with unmodified initial conditions. The purpose of repeating each three times is to ensure we can distinguish the improvement from velocity-zeroing from stochastic effects that may cause intrinsic variation between runs (Genel et al. 2019; Keller et al. 2019). The integration for both types of initial condition is performed using a HLLC Riemann solver (Toro, Spruce & Speares 1994) and the MinMod slope limiter to construct gas variables at cell interfaces. AMR follows a quasi-Lagrangian strategy, with cells being split once the number of dark matter particles exceeds 8. Refinement is only permitted for zoom particles, and is disabled beyond RAMSES level 24, i.e. the smallest cell permitted is 2^{24} times smaller than the simulation length, corresponding to a cell size of 3 pc. Because the code solves gravitational forces on the grid, this is also the smallest softening length in the refined region.

Finally, we undertake further hydrodynamical simulations to test for the effect of improved grid resolution that, as discussed in the Introduction, is guaranteed to reduce the effects of numerical diffusion. We performed two versions of such an enhanced-resolution simulation; the first reduces the threshold for refinement to 2 (rather than 8) dark matter particles, while the second maintains the same number threshold but increases the mass resolution of the dark matter particles themselves by a factor 8. The two approaches yield almost identical results, and we present results only from the first. The chief problem with forcing high resolution in either of these ways is the computational cost, as shown in Table 1. We were only able to continue the simulation to $z = 6.6$, at which point it had consumed almost 120 000 core hours (on 160 cores). By contrast, the unmodified runs were able to reach $z = 0$ within approximately 30 000 core hours. Velocity zeroing the initial conditions leads to a modest speed *improvement* compared with using unmodified initial conditions, possibly because the decreased velocities permit longer time-steps.
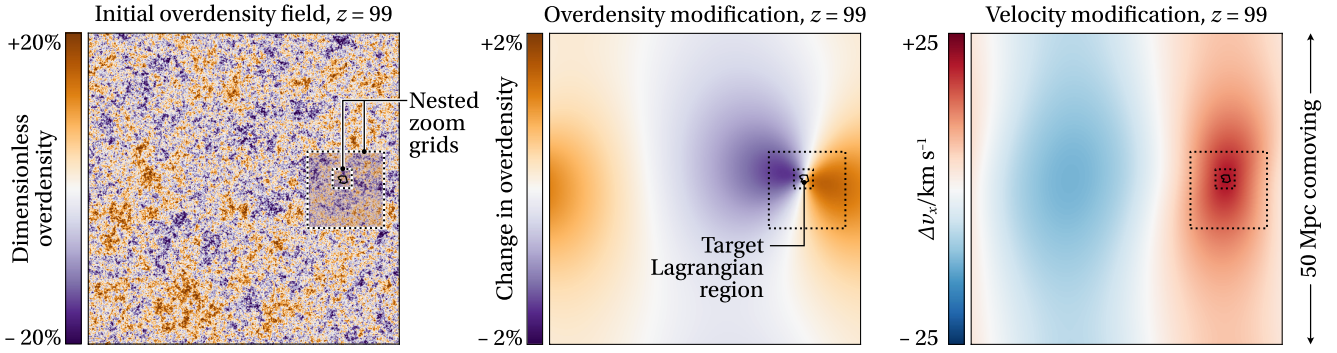
Figure 2. A slice through the overdensity field in our initial conditions at $z = 99$ (left-hand panel) and the genetic modification that we apply to make the linear-theory velocity of the Lagrangian region zero (centre panel, colour scale exaggerated by a factor 10 relative to the left-hand panel). The x-component of the resulting change in the linear velocity field is shown in the right-hand panel. This illustrates how, to effect a change in the velocity, GENETIC creates gradients in the potential by marginally altering very long-wavelength modes in the overdensity. These have negligible effect on the local large-scale structure from which the galaxy forms (see Fig. 3) because the Lagrangian region is tiny compared to the correlation length of the velocity. All panels show the location of our nested zoom grids that are required to achieve an effective resolution of $32\,768^3$.

All simulations implement radiative cooling, star formation, feedback, metal enrichment, and the cosmic UV background following the 'weak' scheme of Agertz et al. (2020). In brief, star formation obeys a Schmidt law, and is only allowed to proceed in cold gas ($T < 100\,\mathrm{K}$) with densities exceeding $300\,m_p\,\mathrm{cm}^{-3}$ where m_p is the proton mass. Feedback injects energy, momentum, mass, and heavy elements into the interstellar medium from SNIa, SNII, and stellar winds. The energy injection scheme includes a temperature ceiling $T < 10^8\,\mathrm{K}$ and velocity ceiling $v < 10^3\,\mathrm{km\,s^{-1}}$; these limitations are computationally efficient but unphysical and so our results should not be directly compared to observations. The focus of this work is comparisons between different simulations. In particular, we will see that numerical diffusion is most pronounced in the intergalactic medium surrounding our galaxies and therefore we would expect the issues discussed in this paper to be insensitive to uncertainties in the feedback prescription.

3 RESULTS

Fig. 3 shows snapshots of (from left to right) the dark matter, AMR structure, gas density in the large-scale structure, and a closer view of gas densities in the galaxy, all at $z = 15.7$. The top panels show the unmodified simulation, the middle panels show the effect of forcing additional mesh refinement, and the bottom panels show results from adopting velocity-zeroed initial conditions.[2] While the dark matter large-scale structure agrees well between all versions of the simulation, the gas in filaments is softened to differing degrees. Because the filament is poorly resolved when using the unmodified initial conditions, gas struggles to cool within the galaxy's shallow potential well. As a result, while star formation has commenced at $z = 15.7$ in the velocity-zeroed case, it is yet to do so in the other simulations. Differences in the gas structure persist until reionization heats the gas at $z \simeq 6$, introducing thermal pressure into the simulations and so smoothing the gas distribution on much larger scales than the filaments illustrated here. We will first verify that numerical diffusion accounts for the differences between simulations (Section 3.1), before giving an overview of the impact on the galaxy (Section 3.2).

3.1 Scale of numerical diffusion

In Appendix A, we outline a 1D advection test, supported by analytic arguments, to show that the expected scale σ of diffusion on a fixed grid with our adopted solver is given approximately by
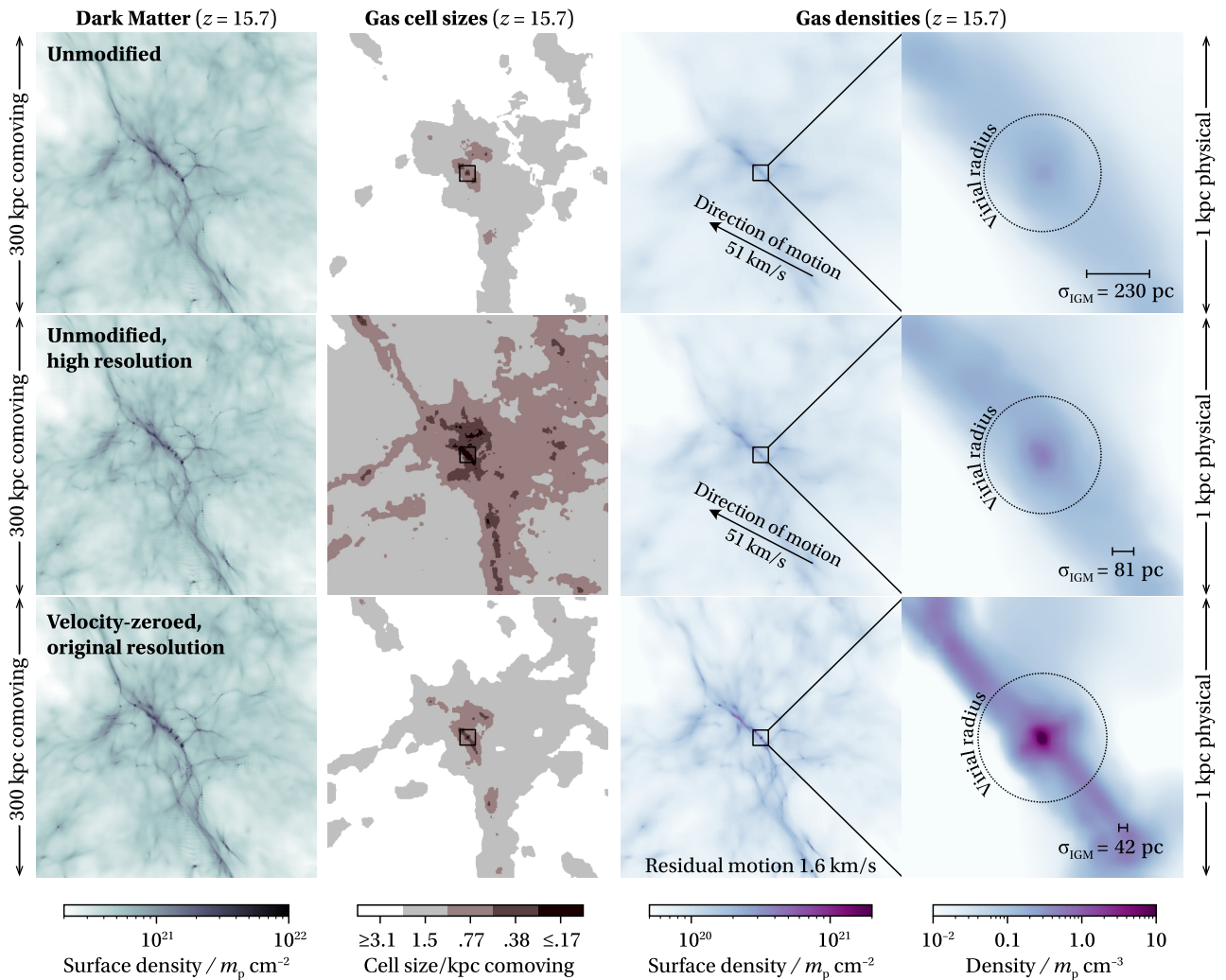
$$\sigma \approx 0.9 L^{0.29} \Delta^{0.71}, \tag{1}$$

where L is the distance travelled by the fluid across the grid and Δ is the size of a cell. The diffusion scale σ can be reduced by decreasing one or both of L and Δ, in agreement with expectations. We do not take into account refinement, cooling or self-gravity in this simple estimate, and σ should therefore be taken only as a rough indicator of the magnitude of numerical diffusion.

None the less, we use equation (1) to create a basic cross-check by estimating the scale of numerical diffusion in the intergalactic medium at $z = 15.7$, when star formation is about to begin in the velocity-zeroed run. The total comoving distance travelled by the protogalactic region between $z = 99$ and $z = 15.7$ can be directly measured from our outputs to be $L = 286$ and $5\,\mathrm{kpc}$ in the unmodified and velocity-zeroed runs, respectively. We can also study the refinement maps (Fig. 3, second panels from left) to establish the cell size in the intergalactic medium filaments. From top to bottom, for the unmodified, high-resolution and velocity-zeroed runs, respectively, we obtain $\Delta = 0.77, 0.17$, and $0.38\,\mathrm{kpc}$ comoving. With these values, we evaluate from equation (1) $\sigma_{\mathrm{IGM}} = 3.9, 1.3$, and $0.7\,\mathrm{kpc}$ comoving (230, 80, and 40 pc physical), respectively.

The right-hand panels of Fig. 3 show scale bars with these estimated diffusion scales. The minimum scale of structure in the intergalactic medium appears in each case to agree with these approximate smoothing lengths, from which we conclude the numerical diffusion effects are indeed a dominating factor in determining the intergalactic medium densities. We verified that, once reionization heats the intergalactic medium, the thermal pressure erases these differences.

In the case of the velocity-zeroed simulation, part of the reduction in diffusion scale comes from mesh refinement (i.e. its Δ is smaller than that of the original unmodified run). Without including this refinement, the diffusion scale would instead be 70 pc physical, which is still a major improvement on the unmodified simulation (230 pc). The modest initial reduction in diffusion allows for gravity to produce densities that are sufficient to trigger cooling and, in turn, refinement. Inside the galaxy, this virtuous cycle continues, leading

---

[2]For the unmodified and velocity-zeroed cases, we show the first of the three simulations that we performed to check for stochastic effects; we verified that there is a close visual agreement between the re-runs.

Figure 3. The effective resolution of a simulation depends on numerical diffusion, which is a function both of simulation cell size and net motion relative to the grid. Here, we compare, at $z = 15.7$, the effect on our simulations of starting from the original initial conditions (upper panels) with our standard refinement criterion; forcing an extra level of refinement (middle panels); and instead changing the velocity structure to perform the simulation near the galaxy's rest frame (bottom panels). The left-hand panels show the dark matter distribution that is nearly unaffected. This reflects the large-scale coherence of the modifications made to the velocity fields (right-hand panel of Fig. 2). The second set of panels from the left show the gas cell sizes, showing that the velocity-zeroed simulation has a similar refinement map to the unmodified simulation, while by construction the high-resolution simulation has much smaller cells throughout the region depicted (and so is more computationally expensive; see Table 1). The third set of panels shows the gas distribution, which is softened in the unmodified case due to numerical diffusion from advection. Increasing the refinement or suppressing the velocities both reduce the extent of numerical diffusion. The right-hand panels show a zoom-in to the region in which the galaxy is forming, with the virial radius r_{200c} indicated by a dotted circle. The scale bars show the estimated numerical diffusion scales in the intergalactic medium near the galaxy, σ_{IGM}, which we obtain with equation (1).

Table 1. Required number of computational core hours for our hydrodynamic simulations to reach different redshifts. From top to bottom, we report, respectively, for the median of three simulations with unmodified initial conditions; for the simulation with forced refinement (giving higher resolution throughout the intergalactic medium); and for the median of three simulations at the original resolution but with velocity zeroing in the initial conditions. The running total is given at $z = 15.7$ (as star formation is due to commence), at $z = 6.6$ (during the tail end of reionization) and at $z = 0$. The high-resolution run was too slow to continue beyond $z = 6.6$.

	$z = 15.7$	$z = 6.6$	$z = 0$
Unmodified	161	7098	32 236
Unmodified, high resolution	581	119 598	–
Velocity zeroed, original resolution	130	5381	28 708

to an even smaller diffusion scale. The central cooling time of the velocity-zeroed simulation is 10 Myr, compared to 300 Myr for the unmodified run and 200 Myr for the high-resolution case. Provided the cooling time-scale becomes short compared to the age of the Universe (250 Myr at $z = 15.7$), diffusion is no longer a barrier to star formation.

We can conclude that our velocity-zeroing approach has suppressed numerical errors to the point where they no longer hold up high-redshift star formation. We will now confirm this directly by measuring the star formation rates in the simulations.

3.2 Effects of numerical diffusion

We now contrast the properties of our dwarf galaxy simulations in the unmodified and velocity-zeroed cases. We will quote quantitative
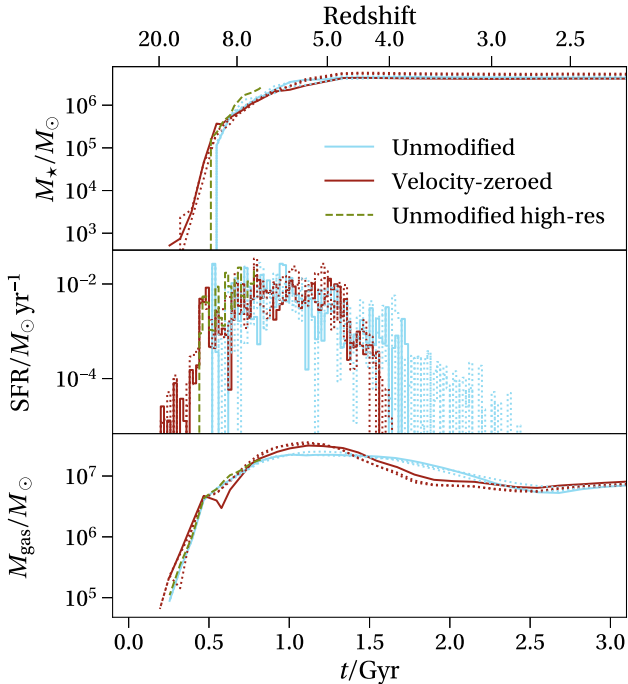
Figure 4. From top to bottom, the total mass of stars in the galaxy, the star formation rate, and the gas mass interior to the virial radius r_{200c} as a function of time for unmodified (blue) and velocity-zeroed (red) initial conditions. The dotted lines show reruns to check for stochastic effects, while the dashed green line shows results from the high-resolution simulation (which terminates at $z = 6.6$). The most physical outcome is obtained using velocity-zeroed initial conditions, with which star formation starts gradually at $z \simeq 16$. The unmodified simulations, which have rapid streaming relative to the grid, delay star formation by 320 and 240 Myr (for standard and high resolution, respectively) due to numerical diffusion. They then undergo a compensating burst so that the final stellar mass is comparable between all cases.

values based on the two means: of the three hydrodynamical simulations with the unmodified initial conditions, and of the three with velocity-zeroed initial conditions. This also allows us to quote uncertainties, given by the half-width of the full spread of results from the respective runs.

We have already established that numerical diffusion due to grid-relative streaming prevents dense gas from accumulating at early times. Fig. 4 illustrates the effect of this on the properties of our dwarf galaxies; from top to bottom it shows the total stellar mass, the star formation rate, and the mass of gas inside the virial radius as a function of time up to $t = 3.1$ Gyr ($z = 2.1$). No star formation activity takes place after this time, although the simulations run to $z = 0$.

The original initial conditions (blue lines) give rise to a galaxy that starts forming stars in an intense initial burst at $z = 9$, forming $1.1 \pm 0.1 \times 10^5$ M$_\odot$ in less than 80 Myr (the time between individual outputs). By contrast, when zeroing the initial velocity (red), star formation ramps up gradually between $z = 17$ and $z = 9$, which is the expected behaviour prior to reionization, since gravitational collapse should deliver a constant stream of gas to the centre of the proto-galaxy. The high-resolution simulation (the green-dashed line) has a star formation history that is intermediate between these two, as expected given the extent of numerical diffusion (Fig. 3). Overall, the onset of star formation is delayed by 320 and 240 Myr, respectively

in the unmodified and high-resolution cases, but the stellar mass rapidly catches up due to the strong initial bursts.

After $z = 9$, all simulations continue to form stars for a significant period. Even though reionization imposes an intergalactic medium thermal floor ($>10^4$ K), our prescription for self-shielding means that residual cool gas can be retained inside the galaxy and continues forming stars. Note also that our reionization completes relatively late, at around $z = 6$, corresponding to the expected thermal history in a void (see Rey et al. 2020). There is considerable scatter in the final time at which residual star formation ceases, truncating at $z = 4.0 \pm 0.2$ for the velocity-zeroed runs but as late as $z = 3.0$ when adopting the original initial conditions. The origin of the very late star formation is a dense knot of gas that is self-shielding and resistant to disruption by feedback. As discussed in Section 2, the feedback implemented in these simulations is somewhat weaker than the 'fiducial' recipe adopted elsewhere in our suite (Agertz et al. 2020), but the effect of rapid bulk flow appears to make the feedback weaker still, perhaps because high-temperature pockets of gas diffuse instead of expanding. We verified that such dense, persistent star-forming knots are not seen in any of our galaxies, irrespective of feedback details, once velocity-zeroed initial conditions are adopted.

The final stellar mass is $4.8 \pm 0.6 \times 10^6$ M$_\odot$ and $4.4 \pm 0.3 \times 10^6$ M$_\odot$ in the unmodified and velocity-zeroed cases, respectively, showing that the early differences in star formation rate do not significantly affect the final stellar mass. This can be understood by considering the overall supply of gas to the virial radius of the halo, which is near-independent of the initial conditions (see bottom panel of Fig. 4). The key effect of diffusion is to keep the gas at lower densities for longer, hence delaying star formation without changing the overall supply of fuel in the circumgalactic medium.

Despite the close agreement in stellar mass, the final distribution of stars is markedly different in the two cases. The top panel of Fig. 5 shows the spherically averaged stellar density between 20 and 2 kpc. The minimum softening scale of 3 pc implies that the profiles are well resolved throughout this range; to increase precision we average the density over three time-steps, and plot the *r.m.s.* error bands assuming uncorrelated Poisson noise. The 3D stellar half-light radius (in the V band) is 300 ± 20 pc with the original initial conditions and 220 ± 20 pc when velocity-zeroed; in other words, the final galaxy is more compact when it is allowed to form at rest with respect to the grid. Notwithstanding this, in two out of three runs with the unmodified initial conditions we find a dense central stellar knot that has been formed during the extended tail in star formation discussed above. The knot is prominent in the stellar density profiles but does not impact significantly on the half-light radius because the total mass in this feature, in the two simulations where it is present, is only 4×10^5 M$_\odot$ (less than a tenth of the total stellar mass).

The difference between the half-light radii from different initial conditions is driven by the ability of stellar feedback to redistribute collisionless orbits (for a review see Pontzen & Governato 2014). At $z = 6$, the half-light radii are 108 ± 40 pc and 115 ± 20 pc for the original and velocity-zeroed initial conditions respectively: they are in agreement. The tail of star formation after this time injects supernova energy which, through gravitational potential fluctuations, expands the orbits of stars and dark matter. The expansion is thus much more pronounced in the galaxy forming from the unmodified initial conditions. The lower panel of Fig. 5 shows that an even stronger discrepancy between unmodified and velocity-zeroed simulations is observed for the dark matter density, with a significant core extending around 200 pc in the former case. The velocity-zeroed run, by contrast, retains a density profile almost as steep as that of the dark-matter-only run (the black-dashed line).
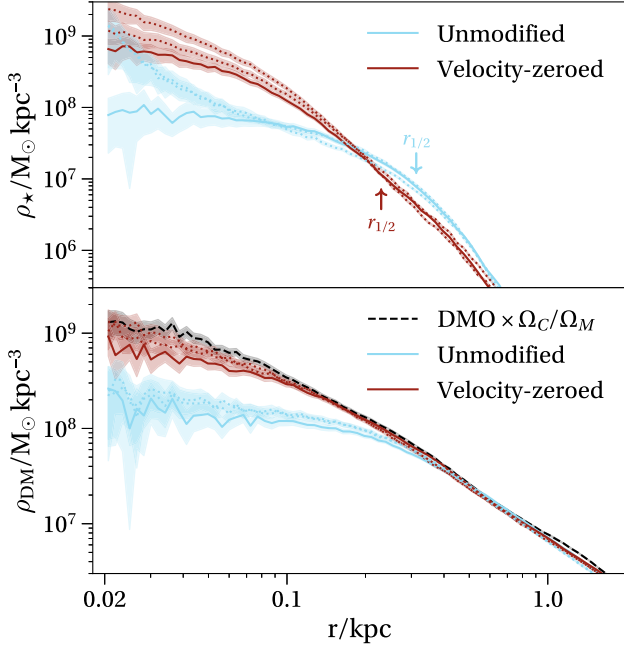
Figure 5. The stellar (top) and dark matter (bottom) spherically averaged density profiles at $z = 0$. The colour scheme follows that of Fig. 4, with an added black dashed line to show the density profile in the dark-matter-only simulation. The shaded regions show r.m.s. uncertainty due to Poisson noise on the profiles. Star formation bursts in the unmodified runs (blue) lead to expansion of orbits, and so a dark matter core and a more diffuse stellar distribution. In the velocity-zeroed runs (red), the dark matter cusp survives and the stellar distribution is more concentrated. Additionally, in two out of three runs with the unmodified initial conditions a dense central star cluster forms, constituting around 10 per cent of the total stellar mass.

We next quantify the burstiness in the star formation, to help understand why the two initial conditions give rise to such different dynamical properties. Fluctuations in the gravitational potential sourced by bursty feedback can pump energy into collisionless orbits, provided changes occur over periods comparable to or shorter than the dynamical time t_{dyn} (Pontzen & Governato 2012). Fluctuations on longer time-scales act adiabatically and thus cannot cause cumulative, long-term changes to stellar or dark matter distributions. To distinguish these regimes, we require time resolution substantially better than $t_{dyn} \equiv (G\rho)^{-1/2} \simeq 21$ Myr in our dwarf at 100 pc. Storing a sufficient number of snapshots (each measuring 9 GB) given that star formation extends over >1 Gyr is out of reach for our study. Instead, we reconstruct the star formation history from the birth time of each star particle, using this as a proxy for the gas density and potential fluctuations.

To compare the level of burstiness, we Fourier transform the star formation history then produce a binned estimate of its modulus squared. This produces a power spectrum $P_{SFR}(\omega)$ as a function of ω, the wavenumber in units of Myr^{-1}. We use 15 bins spaced equally in $\log_{10}\omega$ between -2.6 and 0.5. The result is plotted in Fig. 6, along with 1σ uncertainty bars estimated from the variance within each bin, assuming Gaussian statistics. To reduce these errors and produce a cleaner measurement, we average the power spectra from the three runs for each initial condition type. Larger power indicates greater star formation variability on the corresponding time-scales $\Delta t = 2\pi/\omega$.

The results in Fig. 6 show that the variability around the dynamical time-scale $\omega_{dyn} = 2\pi/t_{dyn}$ is larger in the unmodified initial conditions
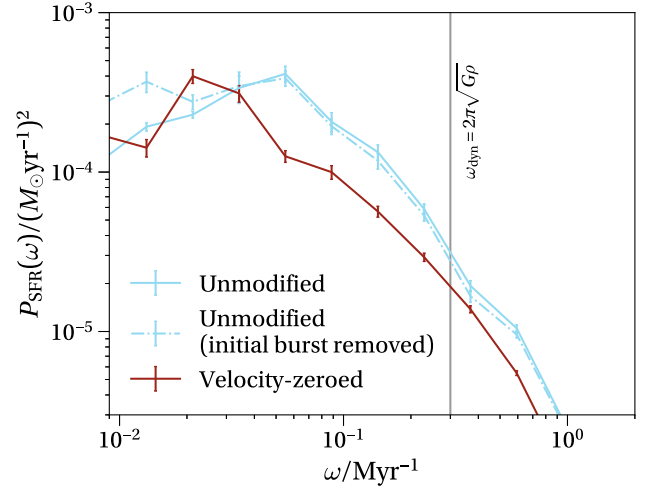
Figure 6. The power spectrum of the star formation rate, which gives a measure of time variability. The unmodified initial conditions (blue line) give rise to a burstier star formation rate, as quantified by the increased power compared to the velocity-zeroed simulation (red line). The dash–dotted line shows the star formation history of the unmodified initial conditions where the initial dramatic burst is excluded, showing that the enhanced burstiness is present throughout the star formation history. The error bars show the r.m.s. uncertainty in the power spectrum estimates. Fluctuations near the dynamical time ω_{dyn}, shown with a vertical grey line, are those which contribute to flattening dark matter cores and expanding stellar orbits.

(blue line) compared to the more physical solution with velocity-zeroing (red line). It is therefore plausible that differences in the burstiness directly give rise to the difference in dark matter and stellar profiles between unmodified and velocity-zeroed cases. Even if we exclude the initial large burst in the unmodified case (blue dash–dotted line), this remains true – star formation proceeds in an excessively bursty mode at all times. The likely cause is that, following a supernova explosion, numerical diffusion holds up the re-condensation of gas to star-forming densities. Just as with the initial burst, the delay causes gas to accumulate in the circumgalactic medium so that once condensation gets underway there is an unphysically catastrophic collapse and a new starburst. This cycle of alternating numerical diffusion and physical collapse ultimately gives rise to a galaxy with significantly altered dynamics.

The actual burstiness of star formation in such dwarf galaxies at high redshift will be very hard to determine observationally, and within simulations is subject to a wide range of subgrid and numerical uncertainties. Therefore, our results should not be taken to indicate that dark matter cores in dwarf spheroidal galaxies are necessarily unobtainable within ΛCDM (although energy conservation sets a size limit to such cores; e.g. Peñarrubia et al. 2012; Read, Agertz & Collins 2016). The results serve only to illustrate that numerical diffusion should be suppressed to the maximum extent possible when studying the cusp–core question.

4 DISCUSSION AND CONCLUSIONS

We have studied how numerical diffusion coupled to streaming relative to a static mesh can unphysically soften gas filaments in the early Universe. The numerical diffusion scale can be orders of magnitude larger than the cell size of the simulation. In the context of the EDGE project to study faint dwarf galaxies, we have solved the problem using genetic modification to zero the streaming velocities of the target region in the initial conditions. Genetic modification is

an ideal tool for this manipulation, since it maintains the galaxy's accretion history and nearby large-scale structure (Fig. 3). While non-linear structure formation still generates a net grid-relative motion, it is suppressed by an order of magnitude and shifted to lower redshift (Fig. 1).

Published EDGE results already use these 'velocity-zeroed' initial conditions (Fig. 2); in this paper we compared hydrodynamical simulations performed with and without the modification. We found that, without the corrective velocity-zeroing, numerical diffusion delays the onset of star formation by 320 Myr. Once star formation finally does initiate in the unmodified case, a dramatic burst brings the total mass of stars into close agreement across all simulations; none the less, the overall properties of the final dwarf galaxy are substantially affected. In particular, the altered star formation history (Fig. 4) leads to stronger potential fluctuations and hence an expanded stellar half-light radius and reduced central dark matter density (Fig. 5).

The extent to which any given mesh-based simulation suffers from these kind of effects will depend on the cooling and feedback physics, the refinement criteria, the hydrodynamic solver and slope limiter, the accretion history of the galaxy, and the nature of the observables under consideration. Advection errors are likely to be most problematic prior to reionization, when there is little physical pressure smoothing in the intergalactic medium; and in the smallest objects, where the gravitational potential is insufficient to compress the gas in the presence of the numerical damping. These conditions bring to mind recent AMR-based simulations of reionization by young galaxies (Xu et al. 2016; Trebitsch et al. 2017; Katz et al. 2018; Rosdahl et al. 2018; Trebitsch et al. 2020). Further investigation will be required to determine whether star formation efficiency and escape fractions in those works may have been affected by the effects of diffusion. Trebitsch et al. (2017), for example, have already noted that the time at which their first stars form, especially in their lowest mass halo, is dependent on resolution. This may be a reflection of the numerical diffusion's dependence on cell size Δ, as seen in equation (1), in agreement with our high-resolution test simulation (e.g. Fig. 4).

Large peculiar velocities are common in cosmological simulations. Typical velocities also increase with box size because they can be sourced by very long-wavelength modes. The magnitude of the problem and its dependence on box size is illustrated in Fig. 7, in which we plot the r.m.s. streaming velocities σ_v as a function of box size L at $z = 16$ (the time at which stars start to form in the EDGE simulations), along with estimated 95 per cent box-to-box variability due to sample variance. The linear theory calculations to obtain this result are given in Appendix B1. For example, the r.m.s. velocity for our box size (50 Mpc $\simeq$ 33.9 Mpc/h) is 79 km s^{-1}, meaning that the halo examined in this work (if unmodified, streaming at 51 km s^{-1}) actually suffers from diffusion somewhat *less* than a typical halo. As the box size increases and more long-wavelength modes start contributing, σ_v reaches 125 km s^{-1} for $L = 100$ Mpc/h; flattens to around 171 km s^{-1} for $L = 500$ Mpc/h; and saturates at 180 km s^{-1} by $L = 3$ Gpc/h. The box-to-box variations around this average are relatively small.

There are two consequences to this size dependence. As well as the obvious conclusion that larger AMR simulations may suffer proportionally more from diffusion effects, it also shows indirectly that velocities are always generated by some of the largest modes in the box; this is demonstrated more explicitly by defining and computing a velocity correlation length in Appendix B2. The practical conclusion is that velocity-zeroing should be achievable for objects of a wide range of masses, without affecting the merger history or local
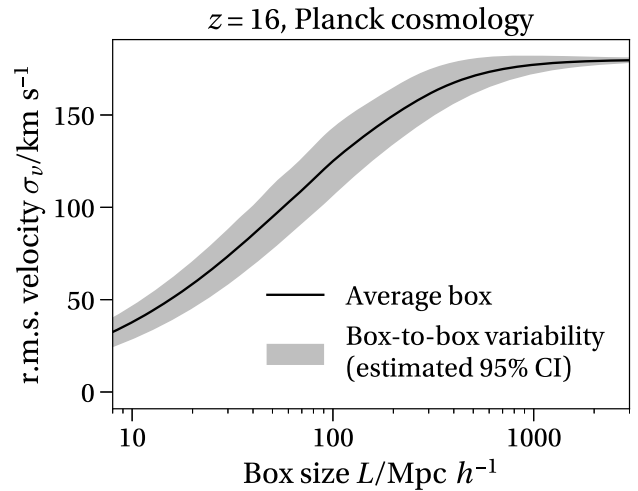
Figure 7. The expected linear r.m.s. velocity σ_v relative to the grid within a single simulation box as a function of its comoving size L, evaluated at $z = 16$ (the time at which stars start to form in our EDGE simulations). Typical streaming velocities increase with the simulation volume, reaching a limit of 180 km s^{-1} for a hypothetical simulation with 3 Gpc box size. Therefore, effects of grid-relative free-streaming are expected to be exacerbated in large cosmological volumes. (Individual boxes will not have precisely this r.m.s. velocity, due to sample variance; the 95 per cent confidence interval from the estimated box-to-box scatter is shown as a grey band.)

environment, just as we have seen for small galaxies in this paper. Provided the box size is very much larger than the Lagrangian region of the halo, we expect the modifications will remain strongly coherent across the entire patch, meaning that the dark matter accretion history and environment around the target object will be essentially unchanged. On the other hand, the main limitation of velocity-zeroing is that it is only achievable for zoom simulations; we cannot suppress the streaming throughout the simulation domain without invoking an unphysical modification to the power spectrum (to remove all long-wavelength modes).

The GENETIC code (Stopyra et al. 2020) allows the user to perform the velocity-zeroing straight forwardly. Similarly, any initial conditions generator that implements Hoffman–Ribak velocity constraints (e.g. Hahn & Abel 2011) can perform a similar manipulation, although we caution that at present we are not aware of codes (other than GENETIC) that correctly propagate constraints across different zoom levels. Given the simplicity of the operation, and the significance of its corrections, there is a strong case for incorporating the technique as standard for AMR zoom simulations.

ACKNOWLEDGEMENTS

DATA AVAILABILITY

The data underlying this article will be shared on reasonable request to the corresponding author.

REFERENCES

Agertz O., Kravtsov A. V., 2015, ApJ, 804, 18
Agertz O. et al., 2007, MNRAS, 380, 963
Agertz O., Kravtsov A. V., Leitner S. N., Gnedin N. Y., 2013, ApJ, 770, 25
Agertz O. et al., 2020, MNRAS, 491, 1656
Bromm V., 2013, Rep. Prog. Phys., 76, 112901
Chabanier S. et al., 2020, A&A, 643, L8
Chisari N. E., Zaldarriaga M., 2011, Phys. Rev. D, 83, 123505
Eisenstein D. J., Hut P., 1998, ApJ, 498, 137
Few C. G., Dobbs C., Pettitt A., Konstandin L., 2016, MNRAS, 460, 4382
Genel S. et al., 2019, ApJ, 871, 21
Hahn O., Abel T., 2011, MNRAS, 415, 2101
Hoffman Y., Ribak E., 1991, ApJ, 380, L5
Hopkins P. F., 2015, MNRAS, 450, 53
Katz H., Kimm T., Haehnelt M., Sijacki D., Rosdahl J., Blaizot J., 2018, MNRAS, 478, 4986
Keller B. W., Wadsley J. W., Wang L., Kruijssen J. M. D., 2019, MNRAS, 482, 2244
Kim J.-h. et al., 2016, ApJ, 833, 202
Kretschmer M., Agertz O., Teyssier R., 2020, MNRAS, 497, 4346
Lewis A., Challinor A., Lasenby A., 2000, ApJ, 538, 473
McNally C. P., Lyra W., Passy J.-C., 2012, ApJS, 201, 18
Muñoz D. J., Kratter K., Springel V., Hernquist L., 2014, MNRAS, 445, 3475
Peñarrubia J., Pontzen A., Walker M. G., Koposov S. E., 2012, ApJ, 759, L42
Planck Collaboration XVI, 2014, A&A, 571, A16
Pontzen A., Governato F., 2012, MNRAS, 421, 3464
Pontzen A., Governato F., 2014, Nature, 506, 171
Pontzen A., Tremmel M., 2018, ApJS, 237, 23
Pontzen A., Roškar R., Stinson G., Woods R., 2013, Astrophysics Source Code Library, record ascl:1305.002
Read J. I., Agertz O., Collins M. L. M., 2016, MNRAS, 459, 2573
Rey M. P., Pontzen A., 2018, MNRAS, 474, 45
Rey M. P., Pontzen A., Agertz O., Orkney M. D. A., Read J. I., Saintonge A., Pedersen C., 2019, ApJ, 886, L3
Rey M. P., Pontzen A., Agertz O., Orkney M. D. A., Read J. I., Rosdahl J., 2020, MNRAS, 497, 1508
Robertson B. E., Kravtsov A. V., Gnedin N. Y., Abel T., Rudd D. H., 2010, MNRAS, 401, 2463
Rosdahl J. et al., 2018, MNRAS, 479, 994
Roth N., Pontzen A., Peiris H. V., 2016, MNRAS, 455, 974
Scannapieco C. et al., 2012, MNRAS, 423, 1726
Somerville R. S., Davé R., 2015, ARA&A, 53, 51
Springel V., 2010a, ARA&A, 48, 391
Springel V., 2010b, MNRAS, 401, 791
Stopyra S., Pontzen A., Peiris H., Roth N., Rey M., 2020, ApJS, preprint (arXiv:2006.01841)
Teyssier R., 2002, A&A, 385, 337
Teyssier R., 2015, ARA&A, 53, 325
Toro E. F., 2009, Riemann Solvers and Numerical Methods for Fluid Dynamics. Springer, Berlin Heidelberg
Toro E. F., Spruce M., Speares W., 1994, Shock Waves, 4, 25
Trebitsch M. et al., 2020, A&A, preprint (arXiv:2002.04045)
Trebitsch M., Blaizot J., Rosdahl J., Devriendt J., Slyz A., 2017, MNRAS, 470, 224
Tseliakovich D., Hirata C., 2010, Phys. Rev. D, 82, 083520
van de Weygaert R., Bertschinger E., 1996, MNRAS, 281, 84
Weinberger R., Springel V., Pakmor R., 2020, ApJS, 248, 32
Xu H., Wise J. H., Norman M. L., Ahn K., O'Shea B. W., 2016, ApJ, 833, 84

APPENDIX A: EXPECTATIONS FOR NUMERICAL DIFFUSION

In order to confirm that differences in our simulations are accounted for by numerical diffusion, in Section 3 we estimated the expected smoothing using equation (1). In this Appendix, we provide the derivation of that relationship. We study an idealized 1D advection problem, where an initial density distribution $\rho(x, t = 0)$ moves at a constant speed v relative to the grid. We will first provide some analytic insights, and then describe a numerical test that leads to the desired estimate.

The diffusion–advection equation takes the form

$$\frac{\partial \rho}{\partial t} + v \frac{\partial \rho}{\partial x} = \nu_{\text{num}} \frac{\partial^2 \rho}{\partial x^2}. \tag{A1}$$

The term on the right-hand side is generated by inaccuracies in the first-order gradient estimates, and ν_{num} may be obtained by replacing derivatives in the ideal advection problem with Taylor expansions of the discrete numerical gradient estimates (e.g. Toro 2009). Provided that the time-step is sufficiently small compared with the Courant–Friedrichs–Lewy (CFL) time, the numerical diffusion coefficient for the first-order upwind scheme reduces to $\nu_{\text{num}} \simeq 2|v|\Delta$. By Fourier transforming equation (A1) in space, one obtains the solution

$$\tilde{\rho}(k, t) = \tilde{\rho}(k, 0) \, e^{-ikvt} \, e^{-2k^2 |v| t \Delta}, \tag{A2}$$

where $\tilde{\rho}(k, t)$ is the Fourier transform of $\rho(x, t)$. The first exponential in equation (A2) represents the physical advection, while the second represents a Gaussian filter, suppressing high-k modes due to the effect of numerical diffusion. Note that $|v|t = L$, the total distance travelled across the grid at time t. When the inverse Fourier transform is applied to equation (A2), one obtains the displaced original density distribution convolved with a Gaussian of width

$$\sigma_{\text{1st}} \approx L^{1/2} \Delta^{1/2}. \tag{A3}$$

However, our configuration of RAMSES for the EDGE suite does not use a purely first-order scheme; rather it uses a hybrid between first-order and second-order slope estimation known as the MinMod scheme. In regions of the intergalactic medium that are relatively smooth, the scheme should come close to the performance of a pure second-order solver. In such a case, Taylor expansion of the numerical derivatives reveals the errors are third order

$$\frac{\partial \rho}{\partial t} + v \frac{\partial \rho}{\partial x} = \alpha_{\text{num}} \frac{\partial^3 \rho}{\partial x^3}. \tag{A4}$$

where $\alpha_{\text{num}} = \beta |v| \Delta^2$, with a scheme-dependent numerical pre-factor β. We have again assumed that the time-step is much smaller than the CFL time. Instead of a pure Gaussian spread, the Fourier-space solution now involves a dispersion:

$$\tilde{\rho}(k, t) = \tilde{\rho}(k, 0) \, e^{-ikvt} \, e^{i\beta k^3 vt \Delta^2}. \tag{A5}$$

This is poorly behaved as k becomes large, giving rise to oscillatory downstream instabilities for pure second-order schemes. In practice, as stated above, second-order schemes are regularized using a slope limiter that cannot be represented in a linear analysis. All one can say from equation (A5) is that there is a critical wavenumber k_{crit} at which errors must become significant, $\beta \Delta^2 k_{\text{crit}}^3 |v| t \approx 1$. From this, we estimate the scaling of the diffusion in second-order schemes:

$$\sigma_{\text{2nd}} \propto k_{\text{crit}}^{-1} \propto L^{1/3} \Delta^{2/3}. \tag{A6}$$

The steeper scaling with cell size Δ compared to the first-order case (A3) implies that the effective numerical diffusion can be suppressed more efficiently with increasing resolution.

To test whether this argument holds, we used RAMSES to perform 1D simulations; the purpose is to estimate the width of numerical diffusion as an improvement on equation (A6). Specifically, we initialize a periodic box with 2^8 cells at uniform density and temperature. We then introduce by hand a 50 per cent enhancement in density to a single cell at the centre of the box. The perturbation is chosen to be in pressure equilibrium with its surroundings and moreover we do not include any gravity or cooling in our simulation, so that the problem remains one of scale-free advection. To measure the diffusion width as a function of time we impart a uniform velocity v to the entire grid so that we directly test the accuracy of the advection. We ensure our time-steps are much smaller than the CFL time.

For each snapshot, we fit a Gaussian with width σ; the fit remains qualitatively good even at late times, although analysis in Fourier space shows that the high-k behaviour departs from an exact Gaussian (as expected given the higher order solver). The width σ is fit extremely well by the relation

$$\sigma \approx 0.9 L^{0.29} \Delta^{0.71}, \tag{A7}$$

which is very close to the expected trend, equation (A6).

Our tests were run with the HLLC solver and MinMod slope limiter, to match the choices in the EDGE cosmological suite. We caution, however, that the estimated diffusion scales do not take into account the intrinsic effect of mesh adaptation, nor the related fact that the structure being advected forms from initial small fluctuations rather than being passively advected. None the less, these scales do permit a helpful qualitative interpretation of the difference between unmodified and velocity-zeroed simulations. In Section 3, we showed that equation (A7) gives consistent estimates with the scales on which smoothing is seen in our simulations.

APPENDIX B: CALCULATING STATISTICAL PROPERTIES OF THE VELOCITY FIELD IN COSMOLOGICAL INITIAL CONDITIONS

B1 Typical grid-relative velocities as a function of box size

In concluding the paper, we drew attention to the steep rise in r.m.s. velocity σ_v as a function of computational box size (Fig. 7). This Appendix provides a derivation of that trend.

To compute the typical velocity with which galaxies will be streaming relative to the grid, we assume that large-scale cosmic flows are dominated by the linear field, at least at the high redshifts of interest. We assume that the real-space overdensity field $\delta(\boldsymbol{x}, z)$ (at a given comoving position $\boldsymbol{x}$ and redshift z) is linked to the Fourier-space overdensities $\delta(\boldsymbol{k}, z)$ with the convention

$$\delta(\boldsymbol{x}, z) = \frac{1}{(2\pi)^3} \int d^3 \boldsymbol{k} \, e^{i\boldsymbol{k}\cdot\boldsymbol{x}} \delta(\boldsymbol{k}, z). \tag{B1}$$

We can then define the power spectrum of density fluctuations $P(k, z)$ via

$$\langle \delta(\boldsymbol{k}, z)\delta(\boldsymbol{k}', z) \rangle = (2\pi)^3 \delta_{\mathrm{D}}(\boldsymbol{k} + \boldsymbol{k}') P(k, z), \tag{B2}$$

where δ_{D} denotes the Dirac delta function, and $k = |\boldsymbol{k}|$. The one-point r.m.s. density fluctuations σ_δ are given by

$$\sigma_\delta^2 = \langle |\delta(\boldsymbol{x})|^2 \rangle = \frac{1}{2\pi^2} \int k^2 P(k, z) \, dk. \tag{B3}$$

Typically, values of σ are only quoted after a filter has been applied to the field to smooth small-scale fluctuations. However, for velocity (as opposed to density) fluctuations, the small-scale fluctuations are

a subdominant contribution as we shall see, and no smoothing will be required to obtain a converged result.

We wish to find the equivalent expression to (B3) but for r.m.s. fluctuations in velocity. As we will consider box sizes of several Gpc, relativistic gauge freedom must be considered. While numerical simulations by construction compute density fields in synchronous gauge, their velocity fields reside in conformal Newtonian gauge (e.g. Chisari & Zaldarriaga 2011). During matter domination, when potentials are constant, the physical velocities and overdensities are then related by

$$v_i(\boldsymbol{k}, z) = \frac{-i H(z) k_i}{(1+z)k^2} \delta(\boldsymbol{k}, z), \tag{B4}$$

where $v_i(\boldsymbol{k}, z)$ is the Fourier transform of the velocity field components for wavenumber $\boldsymbol{k}$, $\delta(\boldsymbol{k}, z)$ is the corresponding overdensity, z is the redshift for which we are computing the fields, and $H(z)$ is the Hubble parameter. Consequently, the expected variance on the velocity within a cube of comoving size L is given by

$$\langle \sigma_v^2 \rangle = \langle |\boldsymbol{v}(\boldsymbol{x})|^2 \rangle = \frac{H(z)^2}{2\pi^2(1+z)^2} \int_{2\pi/L}^{\infty} P(k, z) \, dk. \tag{B5}$$

In the physical Universe, the limits of the integral are between $k = 0$ and $k = \infty$ but in the numerical Universe there are cut-offs at both ends, provided by the particle spacing (high k) and box size (low k); well-resolved simulations have sufficient resolution to ignore the former, but the latter must be given by $2\pi/L$. The missing factor k^2 compared to the density case (B3) means that the velocity variance integral (B5) is typically dominated by low-k modes. Therefore, in our estimate we ignore any high-k cut-off or filtering scale and integrate from $k_{\mathrm{min}} = 2\pi/L$ where L is the box size. We evaluate $P(k)$ for our adopted ΛCDM cosmology (see Section 2) using CAMB (Lewis, Challinor & Lasenby 2000). This gives rise to the function $\sigma_v(L)$, which is plotted in Fig. 7.

The steep rise in σ_v as a function of L up to several hundred Mpc/h comoving implies that flows will be sourced by the fundamental or near-fundamental modes in cosmological simulations. In turn, this suggests that there may be a significant box-to-box variability in σ_v since the sampling of such modes is sparse, exacerbating sample variance. To quantify the scatter, we must drop the assumption of the continuum limit and instead write

$$\delta(\boldsymbol{x}) = \sum_{\boldsymbol{k}} e^{i\boldsymbol{k}\cdot\boldsymbol{x}} \delta_{\boldsymbol{k}}, \tag{B6}$$

where the sum is over all modes between the fundamental and Nyquist frequencies. We do not consider the effect of the Nyquist cut-off here, since for practical simulation geometries it has little effect on the magnitude of velocity fluctuations, for the reasons given immediately below equation (B5). The discrete $\delta_{\boldsymbol{k}}$ obey Gaussian statistics with

$$\langle \delta_{\boldsymbol{k}}(z)\delta_{\boldsymbol{k}'}(z)^* \rangle = \frac{1}{L^3} \delta_{\boldsymbol{k}\boldsymbol{k}'} P(k, z). \tag{B7}$$

where the Kronecker delta $\delta_{\boldsymbol{k}\boldsymbol{k}'} = 1$ if $\boldsymbol{k} = \boldsymbol{k}'$ and 0 otherwise; and the wavenumber k is defined by $k \equiv |\boldsymbol{k}|$. Starting from equation (B6) and applying Wick's theorem alongside equation (B7), one may now show that

$$\langle |\delta(x)^4| \rangle - \langle |\delta(x)^2| \rangle^2 = 2 \sum_{\boldsymbol{k}} \frac{P(k, z)^2}{L^6}. \tag{B8}$$

When dealing with the velocity magnitude instead of the overdensity we pick up four factors of the conversion ratio (B4) and so, returning to the continuum approximation, we can quantify the expected box-

to-box scatter in σ_v as

$$\langle(\sigma_v^2)^2\rangle - \langle\sigma_v^2\rangle^2 = \frac{H(z)^4}{\pi^2 L^3 (1+z)^4} \int_{2\pi/L}^{\infty} \mathrm{d}k \, k^{-2} P(k,z)^2. \qquad (B9)$$

This final expression retains two dependences on L. The dependence in the integration limits can be traced to the cut-off in power above the box scale; the pre-factor L^{-3}, on the other hand, can be traced to the density of sampling in Fourier space as a function of box size. To turn the expression into the confidence interval shown in Fig. 7, we assume that the distribution in σ_v is Gaussian, with second and fourth moments as expressed in equations (B5) and (B9). This does not describe the precise distribution of velocities, which is in fact a generalized chi-squared distribution in σ_v^2. None the less, we verified by comparison to 1000 realizations of the velocity field with a box size of $L = 60\,\mathrm{Mpc}/h$ that the Gaussian limit constitutes a good approximation, with the main difference being that the true distribution has a slight skew towards a high σ_v tail.

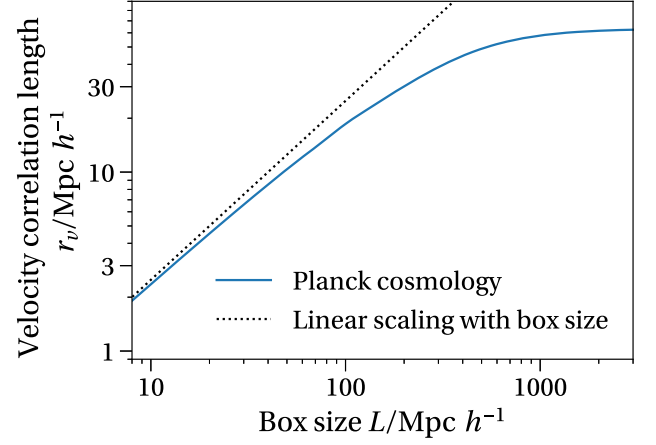## B2 Velocity correlation length as a function of box size

Fig. 2 shows that, in our specific case, the modification is coherent over a region far larger than the Lagrangian patch of our simulated galaxy. This in turn ensures that its halo accretion history and environment are unchanged by the velocity-zeroing operation. In the conclusions (Section 4), we claimed that such coherence can be arranged even for larger haloes, by sufficiently increasing the simulation box size L. We now provide evidence for that claim, by calculating the distance over which the initial velocity field is correlated.

Consider the velocity cross-correlation between two points separated by a distance r, which we define by $\xi_v(r) \equiv \langle \boldsymbol{v}(\boldsymbol{0}) \cdot \boldsymbol{v}(\boldsymbol{r})\rangle$; by statistical homogeneity we take the first point to be the origin, and by statistical isotropy the displacement $\boldsymbol{r}$ can be chosen to lie along any direction. Following the same arguments leading to equation (B5), we obtain

$$\xi_v(r) = \frac{H(z)^2}{2\pi^2 (1+z)^2} \int_{2\pi/L}^{\infty} P(k,z) \frac{\sin kr}{kr} \, \mathrm{d}k. \qquad (B10)$$

The correlation function ξ_v correctly boils down to equation (B5) in the limit $r \to 0$. As $r \to \infty$, $\xi_v(r)$ oscillates and dies away to zero.

We now define the velocity correlation length r_v (in comoving units) to be the value of r at which the correlation has declined by a factor of 4, i.e. $\xi_v(r_v) \equiv \xi_v(0)/4$. The factor 4 is somewhat arbitrary; adopting another value rescales the quantitative values for r_v, but does not lead to qualitatively different conclusions. Note that r_v is independent of redshift in the linear approximation.

Figure B1. Representative comoving distance r_v over which the velocity remains correlated, as a function of simulation comoving box size L. For small box sizes, r_v scales proportionally to L because the velocity power is dominated by waves near the fundamental mode. (The dotted line shows a linear scaling for comparison.) For $L > 100\,\mathrm{Mpc}/h$, the correlation length begins to flatten until it reaches its cosmological value $r_v \simeq 63\,\mathrm{Mpc}/h$.

We evaluate $r_v(L)$ numerically using the ΛCDM cosmological power spectrum $P(k)$ as above, and plot the result in Fig. B1. For $L < 100\,\mathrm{Mpc}/h$, r_v increases linearly with L because the velocities are strongly dominated by the near-fundamental modes. The velocity modifications are coherent over a significant fraction of the box, which is consistent with the visual impression given by Fig. 2. At larger L, as the overall r.m.s. velocity begins to saturate (Fig. 7), the correlation length grows more slowly. For very large boxes ($L = 3\,\mathrm{Gpc}/h$) patches as large as $r_v \simeq 63\,\mathrm{Mpc}/h$ are well correlated in velocity. There is no benefit in further increasing L since r_v is now fully determined by the cosmological model rather than the finite numerical domain.

Given the large values of r_v attainable, we conclude that for zoom galaxy formation simulations, it should be possible to make velocity modifications without altering the accretion history and local large-scale structure in any significant way. For example, a galaxy of mass $10^{12}\,\mathrm{M}_\odot/h$ collapses from a region of comoving radius $\simeq 1.4\,\mathrm{Mpc}/h$; for a cluster of mass $10^{14}\,\mathrm{M}_\odot/h$, the corresponding value is $\simeq 6.5\,\mathrm{Mpc}/h$, still a factor 10 lower than the limiting value of r_v.

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