Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **465,** 569–587 (2017)
Advance Access publication 2016 October 17

doi:10.1093/mnras/stw2670

What does the Bullet Cluster tell us about self-interacting dark matter?

Andrew Robertson,* Richard Massey and Vincent Eke
Institute for Computational Cosmology, Durham University, South Road, Durham DH1 3LE, UK

Accepted 2016 October 13. Received 2016 October 12; in original form 2016 May 18

ABSTRACT

We perform numerical simulations of the merging galaxy cluster 1E 0657−56 (the Bullet Cluster), including the effects of elastic dark matter scattering. In a similar manner to the stripping of gas by ram pressure, dark matter self-interactions would transfer momentum between the two galaxy-cluster dark matter haloes, causing them to lag behind the collisionless galaxies. The absence of an observed separation between the dark matter and stellar components in the Bullet Cluster has been used to place upper limits on the cross-section for dark matter scattering. We emphasize the importance of analysing simulations in an observationally motivated manner, finding that the way in which the positions of the various components are measured can have a larger impact on derived constraints on dark matter's self-interaction cross-section than reasonable changes to the initial conditions for the merger. In particular, we find that the methods used in previous studies to place some of the tightest constraints on this cross-section do not reflect what is done observationally, and overstate the Bullet Cluster's ability to constrain the particle properties of dark matter. We introduce the first simulations of the Bullet Cluster including both self-interacting dark matter and gas. We find that as the gas is stripped it introduces radially dependent asymmetries into the stellar and dark matter distributions. As the techniques used to determine the positions of the dark matter and galaxies are sensitive to different radial scales, these asymmetries can lead to erroneously measured offsets between dark matter and galaxies even when they are spatially coincident.

Key words: astroparticle physics – galaxies: clusters: individual: The Bullet Cluster – cosmology: theory – dark matter.

1 INTRODUCTION

The massive galaxy cluster 1E 0657−56 (the 'Bullet Cluster') acts as a dark matter (DM) particle collider, potentially allowing for discrimination between different particle physics models of DM. In particular, limits on the offset between the galaxies and DM associated with the smaller DM halo (the 'bullet') as well as limits on the loss of DM mass from the bullet have been used to place constraints on the DM–DM elastic scattering cross-section (Randall et al. 2008; Kahlhoefer et al. 2014, hereafter R08 and K14, respectively).

Of the myriad of possible DM candidates, the most favoured candidates for the DM particle (e.g. the lightest neutralino in the minimal supersymmetric standard model) typically have only weak non-gravitational interactions. This is not a necessary property of DM, and it was first noted by Spergel & Steinhardt (2000) that as well as being allowed from a particle physics perspective, DM with a significant cross-section for elastic scattering could have interesting astrophysical consequences. In particular, self-interacting dark matter (SIDM) could alleviate discrepancies between the results of *N*-body simulations of collisionless cold dark matter (CDM) and observations of dwarf galaxies (for a review see Weinberg et al. 2015).

While SIDM may not be unique in offering a solution to these 'small-scale problems' (Governato et al. 2012; Pontzen & Governato 2012; Sawala et al. 2016), there are numerous DM particle candidates that give rise to scattering between DM particles (Carlson, Machacek & Hall 1992; Bento et al. 2000; Kusenko & Steinhardt 2001; Mohapatra, Nussinov & Teplitz 2002; Feng et al. 2009; Tulin, Yu & Zurek 2013b; Boddy et al. 2014; Kaplinghat, Tulin & Yu 2014; Wang, Zhang & Zhao 2016), so it is an important challenge to try and constrain the cross-section for DM–DM scattering from astrophysical observations, in a bid to constrain the allowed parameter space for DM particle models. If it was found that DM must have a significant self-interaction cross-section, this would have a profound effect on particle physics theories of DM, ruling out many of the favoured (and most searched for) DM candidates.

Additional motivation for studying SIDM comes from the detection of separations between the distribution of stars and DM in galaxy clusters (Williams & Saha 2011; Mohammed et al. 2014; Massey et al. 2015). If the offset observed in Abell 3827 (Massey et al. 2015) is interpreted as resulting from SIDM, then it corresponds to an isotropic scattering cross-section of $\sigma/m \sim 1.5\,\mathrm{cm}^2\,\mathrm{g}^{-1}$

* E-mail: andrew.robertson@durham.ac.uk

(Kahlhoefer et al. 2015). While such an offset could potentially arise from an out of equilibrium system, or dynamical effects such as tides or dynamical friction acting differently on the differently distributed stars and DM, offsets of this size appear to be rare in a ΛCDM universe (Schaller et al. 2015).

With this motivation, we revisit (R08, K14) the use of galaxy-cluster collisions to constrain the nature of DM. Clusters are useful, as their distribution of DM can be probed by both strong and weak gravitational lensing. The relative velocities of DM particles within clusters is also of order 1000 km s^{-1}, two orders of magnitude larger than in dwarf galaxies. Velocity-dependent cross-sections can arise naturally in models for SIDM (Ackerman et al. 2009; Buckley & Fox 2010; Loeb & Weiner 2011; van den Aarssen, Bringmann & Pfrommer 2012; Tulin, Yu & Zurek 2013a), and constraining such models requires a handle on the cross-section at different velocity scales (Kaplinghat, Tulin & Yu 2016).

The first attempt to use colliding galaxy clusters to constrain the collisional nature of DM (Markevitch et al. 2004) found that $\sigma/m < 5\,\mathrm{cm^2\,g^{-1}}$ from limits on the size of any potential offset between the DM and stars in the Bullet Cluster. This constraint, derived from analytical toy models, was improved by R08 who ran *N*-body simulations of Bullet Cluster-like systems with SIDM. Combined with tighter constraints on any DM–galaxy separation (Bradač et al. 2006), they found $\sigma/m < 1.25\,\mathrm{cm^2\,g^{-1}}$.

Owing to the high relative velocity of the DM haloes in the Bullet Cluster, DM particles from the bullet that scatter with particles from the main cluster will typically have sufficient energy to escape the potential of the bullet halo, and so the bullet halo would evaporate due to DM self-interactions. The mass-to-light ratio of the bullet halo is similar to that of the main halo, and if one assumes that this similarity means that less than 23 per cent of the DM in the inner regions of the bullet halo could have scattered with particles from the main halo, then the R08 simulations suggest that $\sigma/m < 0.7\,\mathrm{cm^2\,g^{-1}}$. However, observations of over 200 galaxy clusters (Popesso et al. 2007) have shown that there is significant scatter in the luminosity–mass relation for clusters. Specifically, Popesso et al. (2007) found that the *r*-band luminosity of clusters was tightly related to the number of galaxies with an *r*-band absolute magnitude of $M_\mathrm{r} \leq -20$, but that from the number of galaxies the mass of the cluster could only be predicted with an accuracy of 55 per cent. This suggests that the significance of the $\sigma/m < 0.7\,\mathrm{cm^2\,g^{-1}}$ result derived in R08 is overstated, as it assumes little intrinsic scatter in the mass-to-light ratios of clusters.

Since the discovery of the Bullet Cluster, other colliding cluster systems have been found, and used to constrain the cross-section for DM scattering. Similar analysis to that performed on the Bullet Cluster places limits of $\sigma/m < 4\,\mathrm{cm^2\,g^{-1}}$ from MACS J0025.4−1222 (Bradač et al. 2008), $\sigma/m < 3\,\mathrm{cm^2\,g^{-1}}$ from Abell 2744 (Merten et al. 2011), and $\sigma/m < 7\,\mathrm{cm^2\,g^{-1}}$ from DLSCL J0916.2+2951, the 'Musket Ball Cluster' (Dawson et al. 2012).

K14 pointed out that during galaxy-cluster collisions, DM particles preferentially collide along the merger axis, and that these systems could be used to determine not just the cross-section for DM scattering, but its angular dependence. In particular, they showed that the resulting distribution of DM is different for the case of short-range, contact interactions (for which the scattering is isotropic), compared to long-range interactions, where there is a preference for low scattering angles, and particles can undergo many small momentum transfer collisions.

Compared with systems undergoing major mergers, clusters undergoing minor mergers with large mass ratios are ubiquitous. Harvey et al. (2015) found 30 such clusters, with a total of 72 pieces of substructure. By looking at the position of the DM substructure relative to the position of the corresponding stars and gas, they placed limits of $\sigma/m < 0.47\,\mathrm{cm^2\,g^{-1}}$ for the DM elastic scattering cross-section.

In this work we choose to focus on the Bullet Cluster, as the gas morphology and the lack of line-of-sight velocity difference between galaxies from the two clusters implies that the collision has taken place with little impact parameter and in the plane of the sky (Barrena et al. 2002). In addition, to this simple geometry, X-ray observations of the shock front leading the gaseous bullet allow the relative velocity between the two merging clusters to be estimated (Markevitch 2006). We limit our study to the case of isotropic and velocity-independent cross-sections, focusing on the importance of the method used to extract position estimates from the simulations.

Our paper is structured as follows. In Section 2, we discuss our implementation of DM scattering within an *N*-body code, as well as the initial conditions we use for our simulations. In Section 3, we discuss different methods for measuring the positions of different components within a merging galaxy cluster, before applying these different methods to our simulations in Section 4. Finally, we give our conclusions in Section 5. We use $\Omega_\mathrm{m} = 0.3$, $\Omega_\Lambda = 0.7$, and $H_0 = 70\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$. At the redshift of the Bullet Cluster ($z = 0.296$) 1 kpc corresponds to 0.23 arcsec.

2 SIMULATIONS

2.1 Implementation of DM scattering

We implemented DM scattering on top of the GADGET-3 TREE-PM *N*-body code, which is an updated version of the publicly available GADGET-2 code (Springel 2005). The scattering was done stochastically using the same algorithm as in Rocha et al. (2013) that they derive from the Boltzmann equation, although we use a top hat kernel rather than a spline kernel. At each time-step, the probability for each pair of nearby particles to scatter is calculated, and a random number is drawn to see which particles do actually scatter. This algorithm is similar to that used in other SIDM simulations (Kochanek & White 2000; Yoshida et al. 2000; Davé et al. 2001; Koda & Shapiro 2011; Vogelsberger, Zavala & Loeb 2012; Fry et al. 2015), with these algorithms differing in the number of neighbours (or search volume) used to find potential scattering pairs.

2.1.1 Assumed DM interaction model

We assume that the particle interactions are fully described by an azimuthally symmetric differential cross-section, defined in the centre of momentum frame of the two particles, which could have some velocity dependence, $\frac{\mathrm{d}^2\sigma}{\mathrm{d}\Omega\,\mathrm{d}v}$. Assuming that the angular and velocity dependences of the cross-section are separable, we can consider only an angularly dependent cross-section. The velocity dependence then enters as a normalization of the total cross-section between pairs of particles, a function of their relative velocity.

From the differential cross-section, $\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}$, we can calculate the total cross-section as

$$\sigma = 2\pi \int_0^\pi \sin\theta \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\,\mathrm{d}\theta. \tag{1}$$

We can then define the probability that a scattered particle changes direction by an angle in the range $[\theta, \theta + \mathrm{d}\theta]$ as

$$P(\theta)\,\mathrm{d}\theta = \frac{1}{\sigma} 2\pi \sin\theta \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\,\mathrm{d}\theta. \tag{2}$$

The code can implement velocity and angular-dependent scattering, but for the rest of this paper we will assume that DM scattering is velocity independent and isotropic, for which $\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \sigma/4\pi$. Point-like interactions that lead to isotropic scattering result from scattering mediated by a heavy particle, when the momentum exchange in the scattering is significantly larger than the mass of the mediator particle, m_ϕ, i.e. when $m_\phi \gg (v/c)m_{\mathrm{DM}}$.

2.1.2 Scattering rate

The scattering rate of an individual DM particle i, with velocity $\boldsymbol{v}_i$, is

$$\Gamma_i = \int f(\boldsymbol{v}')\,\rho\,\frac{\sigma}{m}\,|\boldsymbol{v}_i - \boldsymbol{v}'|\,\mathrm{d}^3\boldsymbol{v}', \tag{3}$$

where f is the velocity distribution function,[1] ρ the local density, and σ/m the cross-section for DM–DM scattering (which could depend on $|\boldsymbol{v}_i - \boldsymbol{v}'|$) divided by the DM particle mass). To calculate scattering probabilities in the simulations, f and ρ are estimated from the volume within a distance h of a particle's position. This leads to the scattering rate

$$\Gamma_i = \sum_j \frac{\sigma_{\mathrm{p}}|\boldsymbol{v}_i - \boldsymbol{v}_j|}{\frac{4\pi}{3}h^3}, \tag{4}$$

where the sum is over all simulation particles in the volume defined by h, and $\sigma_{\mathrm{p}} \equiv (\sigma/m)\,m_{\mathrm{p}}$ with m_{p} the mass of the simulation particles. Throughout this work, σ and m will be the cross-section and mass of individual DM particles, while σ_{p} and m_{p} will be the cross-section and mass of the simulation particles. Astrophysical observables, such as core sizes, are determined by the fraction of particles that are scattered during a process, and so relate to the scattering rate for individual particles. As the scattering rate depends on the product of the number density of particles and the cross-section, and the number density of particles scales inversely as the particle mass, the cross-section of our simulation particles must scale with the simulation particle mass, such that

$$\sigma_{\mathrm{p}} = \left(\frac{\sigma}{m}\right) m_{\mathrm{p}}. \tag{5}$$

From equation (4), the probability of two particles, i and j, separated by a distance less than h, scattering within the next time-step, Δt, is given by

$$P_{ij} = \frac{\sigma_{\mathrm{p}}|\boldsymbol{v}_i - \boldsymbol{v}_j|\,\Delta t}{\frac{4\pi}{3}h^3}, \tag{6}$$

where for velocity-dependent cross-sections, σ_{p} would be a function of $|\boldsymbol{v}_i - \boldsymbol{v}_j|$.

In this scattering procedure, h is a numerical parameter that has to be chosen. In Section 2.3, we investigate the effects of changing h, using both a fixed h for all particles, as well as a variable h that adapts to the local density. We also test adding a kernel weighting to the scattering probability in equation (6), such that the probability of nearby pairs of particles (separation much less than h) is larger than that for particles further apart. We find that using a fixed value of h equal to the gravitational softening length provides correct results, and that kernel weighting has little effect. This is in contrast to smoothed-particle hydrodynamics (for which kernel weighting and adaptive smoothing lengths are necessary) as the scattering is inherently stochastic, and so it is not important that the scattering probability varies smoothly.

[1] Here f is normalized such that $\int f(\boldsymbol{v})\,\mathrm{d}^3\boldsymbol{v} = 1$.

2.1.3 Scattering kinematics

If two particles with identical mass, and velocities $\boldsymbol{v}_i$ and $\boldsymbol{v}_j$, are to scatter, then we must first move into the centre of momentum frame, in which the velocities are $\boldsymbol{v}_i'$ and $\boldsymbol{v}_j' = -\boldsymbol{v}_i'$. We use the direction of $\boldsymbol{v}_i'$ to define the z-axis, from which θ is measured. Then we draw a random θ from $P(\theta)$ to determine the polar angle at which the particles scatter, as well as drawing a random number that we convert into an azimuthal angle. With these two angles, the scattering kinematics in the centre of mass frame is completely determined. Finding the momentum transfer in the centre of mass frame, we can then move back to the simulation frame, and apply these momentum kicks.

For the case of isotopic scattering that we consider in this work, the scattering kinematics is particularly simple. The post-scatter velocities are

$$\begin{aligned}\boldsymbol{u}_i &= \boldsymbol{V} + w\,\hat{\boldsymbol{e}}\\ \boldsymbol{u}_j &= \boldsymbol{V} - w\,\hat{\boldsymbol{e}},\end{aligned} \tag{7}$$

where the $\boldsymbol{u}$ are the post-scatter velocities, $\boldsymbol{V} = (\boldsymbol{v}_i + \boldsymbol{v}_j)/2$, $w = |\boldsymbol{v}_i - \boldsymbol{v}_j|/2$, and $\hat{\boldsymbol{e}}$ is a randomly oriented unit vector.

2.1.4 Multiple scatters within a time-step

As particle scattering is implemented on a particle by particle basis, it is possible for a particle to scatter more than once in a single time-step. While the low rate of particle scattering[2] results in these multiple scatters being infrequent, it is important that they are dealt with in an appropriate way. Because the momentum kick from one scattering event alters the velocities of the particles for any future scattering event; we cannot allow a particle to scatter twice in one time-step with the same initial velocity. Instead we arbitrarily order all pairs of particles, and update the particle velocities when we decide two particles will scatter. In this way, any future scattering events involving these particles will use the updated velocities, essentially time-ordering the scattering events within one simulation time-step.

A further complication arises when running simulations on multiple processors. In order for particles that reside on different processors (which have access to different memory) to scatter, the properties of one of the particles must be exported to the processor on which the other particle resides. To increase parallel efficiency, all of these exports take place simultaneously, and then processors determine if any of their imported particles scatter with their own particles. During this step, it would be possible for a particle that is currently exported to scatter off an imported particle on its own processor. As such, the particle could scatter simultaneously on different processors, which would lead to both scattering events taking place with the same initial velocity for the particle in question. Conserving energy and momentum in each of the two scattering events, and then later combining the two momentum kicks, does not in general conserve energy, and this process could lead to the production of additional kinetic energy in the simulation.

To prevent this, we assign a direction between each pair of processors, and only allow particles to be exported in this direction. We use a constant search radius for all particles, so that the search is

[2] For $\sigma/m = 1\,\mathrm{cm}^2\,\mathrm{g}^{-1}$, Robertson et al. (2015) showed that the average number of scattering events per particle is $\mathcal{O}(1)$ by redshift zero, and so the frequency of particles scattering twice within a single simulation time-step is very low.

symmetric, and particles that would have been exported in the other direction will still have a chance to scatter with particles that they would find there, when those particles are imported. Essentially, particles with a separation less than h that reside on different processors meet each other once and scatter with the probability given in equation (6), while particles on the same processor meet twice, but have half this probability of scattering on each occasion. Assigning a directionality to the particle send/receive process does not completely prevent particles scattering simultaneously on different processors, as it is still possible where particles on three or more domains[3] are within a distance h of each other. However, the rate of scattering within a time-step is low, and the size of the domains compared to the size of h is large, such that these events are highly unlikely.

We keep a log of all scattering events that allows us to detect these problematic encounters. For three particles drawn from a Maxwell–Boltzmann velocity distribution, the mean change in energy when one particle scatters 'badly' from the two other particles is $\langle \Delta E \rangle = \frac{1}{2} \langle \mathrm{KE} \rangle$, where $\langle \mathrm{KE} \rangle$ is the mean kinetic energy of individual particles. This rises to $\langle \Delta E \rangle \approx 0.87 \langle \mathrm{KE} \rangle$, when we weight the triplets of particles by the probability of them scattering, as particles with higher relative velocities are more likely to scatter (equation 6).

In the simulations used in the rest of this work, only one bad scattering event happened. As the expected change in energy due to a bad scattering is of the order of the kinetic energy per particle in the simulation, a bad scattering event changes the total energy by $\sim$1 part in N_{DM}, where N_{DM} is the number of DM particles in the simulation. In our simulations this corresponds to 1 part in 10^7, making it inconsequential compared to the non-conservation of energy from gravitational forces. With variable time-steps, manifest energy conservation is lost (Dehnen & Read 2011), and we find that in our simulations the typical level of energy conservation over the course of a simulation is $\sim$1 part in 10^4.

2.2 Initial conditions

In order to draw meaningful conclusions on the properties of DM from a comparison of our simulations to observations, it is important that the simulations do a reasonable job of recreating the Bullet Cluster's observed properties. Lage & Farrar (2014) performed a large suite of magnetohydrodynamic simulations of the Bullet Cluster, hoping to match a wide range of observational data sets. In order to constrain the 34 parameters required to generate their initial conditions, they required over 1000 simulations, which in our case (looking at the effect of changing the DM–DM scattering cross-section) would have to be done for each cross-section that we investigate. This would be an exceptionally computationally demanding task, and although complicated, the initial conditions generated are still idealized models for the two clusters, ignoring the effects of mass accretion prior to or during the merger, and without substructure that could be important for matching to the lensing data. Instead of attempting the demanding task of finding optimal initial conditions for each cross-section we investigate, we choose to take a simple idealized model for the system, which provides a reasonable match to key data sets. We then investigate how observables (in particular the offset between DM and galaxies) change as the DM cross-section is varied.

[3] A domain is a region of simulation space that is stored on one processor.

2.2.1 *Density profiles*

The main constraints on the total density profiles of the two clusters come from lensing observations. As a first model, we take the best-fitting values from fitting two spherically symmetric Navarro, Frenk & White (1997, hereafter NFW) mass distributions to weak-lensing data, as done in Springel & Farrar (2007, hereafter SF07). With our assumed cosmology, the best-fitting values are $r_{200} = 2136\,\mathrm{kpc}$, $c = 1.94$, and $r_{200} = 995\,\mathrm{kpc}$, $c = 7.12$, for the main cluster and bullet cluster, respectively. Given the redshift of the system at $z = 0.296$, the masses of the two haloes are then $M_{200} \approx 1.5 \times 10^{15}\,\mathrm{M}_\odot$ for the main cluster and $M_{200} \approx 1.5 \times 10^{14}\,\mathrm{M}_\odot$ for the bullet cluster.[4]

The concentration of the main halo derived from weak lensing would place this halo well below the concentration–mass relation derived from observations of galaxy clusters (Merten et al. 2015), or from numerical (Prada et al. 2012; Dutton & Macciò 2014; Diemer & Kravtsov 2015) or analytical (Correa et al. 2015) work. SF07 found that with $c = 2$, the ram pressure on the gas bullet is not sufficient to strip it away from its DM halo. The observed gas bullet trails its DM by $\sim$100 kpc, which they could match by increasing the concentration of the main halo to $c = 3$. Making the main halo even more concentrated than this resulted in overpredicting the gas–DM separation, and also lead to the morphology of the bow shock differing from what is observed. We therefore choose to use $c = 3$ rather than the weak-lensing derived $c = 1.94$ for the main halo in our fiducial model for the collision.

We model the total matter distribution of each cluster with a Hernquist profile (Hernquist 1990),

$$\rho(r) = \frac{M}{2\pi} \frac{a}{r} \frac{1}{(r+a)^3}. \tag{8}$$

These are used because unlike NFW profiles, they have a finite mass and so do not need to be truncated. They also have analytical distribution functions, which allow equilibrium initial conditions to be easily generated, and quantities such as the expected scattering rate within a halo to be calculated analytically.

In order to match an NFW profile to a Hernquist profile, we need to define two matching criteria to fix the Hernquist profile's two free parameters. The first of these we take to be matching the normalization of the density in the central regions, for which $\rho \propto r^{-1}$ for both NFW and Hernquist profiles. We also then match the mass within a radius of r_{200} for the Hernquist profile to that of the NFW profile, making use of the mass within a radius r for a Hernquist profile,

$$M(<r) = M \frac{r^2}{(r+a)^2}. \tag{9}$$

Enforcing these matching criteria, we can find the relationship between the Hernquist parameters, a and M, and the NFW parameters, M_{200}, r_{200}, and c:

$$M = M_{200} \frac{(r_{200}+a)^2}{r_{200}^2} \tag{10}$$

$$a = \frac{r_{200}}{\sqrt{\frac{c^2}{2\left[\ln(1+c) - \frac{c}{1+c}\right]}} - 1}. \tag{11}$$

[4] We define r_{200} as the radius at which the mean enclosed DM density is 200 times the critical density, and M_{200} as the mass enclosed within r_{200}. The concentration, c, is then r_{200}/r_{s}, where r_{s} is the NFW scale radius.

We note that a similar matching procedure is described in the text of Springel, Di Matteo & Hernquist (2005), but that they match M_{200} of the NFW profile to the total mass, M, of the Hernquist profile, resulting in a slightly different formula for a.

2.2.2 *Relative velocity of the DM haloes*

The relative velocity between the two DM haloes in the Bullet Cluster was originally estimated to be 4700 km s^{-1}, as this corresponded to the 'shock velocity', the velocity of the shock front relative to the pre-shocked gas (Markevitch 2006). This large relative velocity would be rare within the context of ΛCDM (Hayashi & White 2006), leading to the suggestion of a long-range fifth-force that would result in additional acceleration (Farrar & Rosen 2007).

Simulations including gas have since shown that the shock velocity can be considerably larger than the relative velocity between the DM haloes. The pre-shocked gas, which belongs to the main halo, is not at rest with respect to its halo, but is instead moving towards the bullet halo. Additionally, the shock front is not at rest with respect to the bullet-halo DM, but moves ahead of it. A discussion of the mechanisms responsible for these effects is available in SF07.

SF07 find that the observed shock velocity can be matched by haloes that collide with a velocity corresponding to infall from infinity. We therefore start our simulations with the cluster centres separated by 4 Mpc, and with a relative velocity that corresponds to the velocity they would obtain if falling from rest at infinite separation, assuming each halo acts like a point mass.

2.2.3 *Summary of initial conditions*

Our fiducial model for Bullet Cluster-like initial conditions is two Hernquist profiles, separated by 4 Mpc, and with a relative velocity of 2970 km s^{-1} along the line joining the two cluster centres. The main halo corresponds to an NFW profile with $M_{200} = 1.5 \times 10^{15}$ M$_\odot$ and $c = 3$, while the bullet halo has $M_{200} = 1.5 \times 10^{14}$ M$_\odot$ and $c = 7.12$. When converted into matched Hernquist profiles (following the method in Section 2.2.1), the masses and scale radii are $M = 3.85 \times 10^{15}$ M$_\odot$, $a = 1290$ kpc, and $M = 2.46 \times 10^{14}$ M$_\odot$, $a = 279$ kpc for the main and bullet halo, respectively.

The mass within each halo is 99 per cent DM, and 1 per cent stars, though we use an equal number of DM and star particles (10^7 of each). The star particles are distributed as a smooth halo following the DM density. While this is not the case in real galaxy clusters, where stars reside within galaxies, we do this to allow us to more easily identify the location of the stellar component. We also run some simulations including non-radiative gas, which are discussed in Section 4.4. The gas initially follows the same density profile as the DM and stars, with the halo mass being 83 per cent DM, 16 per cent gas, and 1 per cent stars. The gas temperature was set so that the gas was in hydrostatic equilibrium, which for the main halo in our fiducial mass model gave a maximum gas temperature of 8.4 keV, in agreement with the temperature of the pre-shocked gas in the Bullet Cluster (Markevitch 2006).

2.2.4 *Comparison to other SIDM studies*

In Fig. 1, we show the density distribution of the main halo and bullet halo from different simulations of the Bullet Cluster. As we are interested in the offset between stars and DM within the bullet halo, the fraction of DM particles from the bullet halo that scatter from a particle in the main halo is an important quantity. We therefore plot the density distributions of the two haloes in a manner that allows us to estimate this fraction. For the main halo, we plot the projected density of DM at different radii, which can be multiplied by the cross-section to get an optical depth for DM scattering. For example, an SIDM particle with $\sigma/m = 1$ cm^2 g^{-1} passing through the main halo of our fiducial model at a projected radius of 200 kpc, where the projected surface density is $\sim$0.15 g cm^{-2}, would have a $\sim$15 per cent chance of scattering off a particle in the main halo.

In the right-hand panel of Fig. 1, we plot the fraction of particles at different projected radii within the bullet halo. As the two haloes collide head-on, this is the distribution of projected radii of the main halo through which bullet-halo particles will pass (if we ignore the motion of DM particles within their own halo). We can then use the two panels of Fig. 1 to calculate the fraction of particles in the bullet halo that scatter with a particle from the main halo. For our fiducial model with $\sigma/m = 1$ cm^2 g^{-1}, we expect $\sim$23 (33) per cent of particles from the inner 400 (150) kpc of the bullet halo to scatter with a particle from the main halo, while for R08 and K14 the numbers are 21 (33) and 28 (36) per cent, respectively. Considering all particles in the bullet halo, the number goes down to 12 per cent for our fiducial model, in good agreement with the value of 13 per cent that we get in our simulations (see Section 4.3).

2.2.5 *Stability of an isolated halo*

In Fig. 2, we show the density of an isolated Hernquist profile, evolved both with and without DM scattering. The halo shown is the same as the smaller halo in our fiducial model for the Bullet Cluster. With collisionless DM, the halo forms a small core with a size $\sim 2\epsilon$, where ϵ is the Plummer-equivalent gravitational softening length. The gravitational force between pairs of particles is Newtonian when they are separated by more than 2.8ϵ, but is reduced below this when they are closer, resulting in the formation of small numerical cores in otherwise cuspy haloes.

With SIDM, the haloes form much larger cores, due to particles being preferentially scattered out of high-density regions. These cores form quickly, and settle to a size that is independent of the DM cross-section, in agreement with Kochanek & White (2000).

Starting our simulations with the cluster centres separated by 4 Mpc results in core passage taking place $\sim$1.1 Gyr after the simulations begin. During this time, the density profiles of the SIDM haloes evolve due to DM scattering, beginning to form constant density cores at their centres. To check that the extent of core formation does not have a large impact on our results, we experimented with different initial separations between the two haloes. Starting the haloes with a separation of 9 Mpc, the haloes have evolved for 3.4 Gyr before they collide. We found this only had a small impact on our results, changing the best-fitting separation between stars and DM (with a scattering cross-section of 1 cm^2 g^{-1}) from 9.2 to 8.4 kpc at the time of the observed Bullet Cluster in our fiducial model. This change is small compared to the effects discussed in Section 4.

2.3 Testing the SIDM implementation

In Fig. 3, we plot the scattering rate per particle in an isolated DM halo with a Hernquist density profile. The halo has a total mass of $M = 10^{15}$ M$_\odot$ and a scale radius $a = 1000$ kpc. The simulation was run for 2.5 Gyr with 10^6 particles, each with a mass $m_p = 10^9$ M$_\odot$ and a Plummer-equivalent gravitational softening length $\epsilon = 12$ kpc.

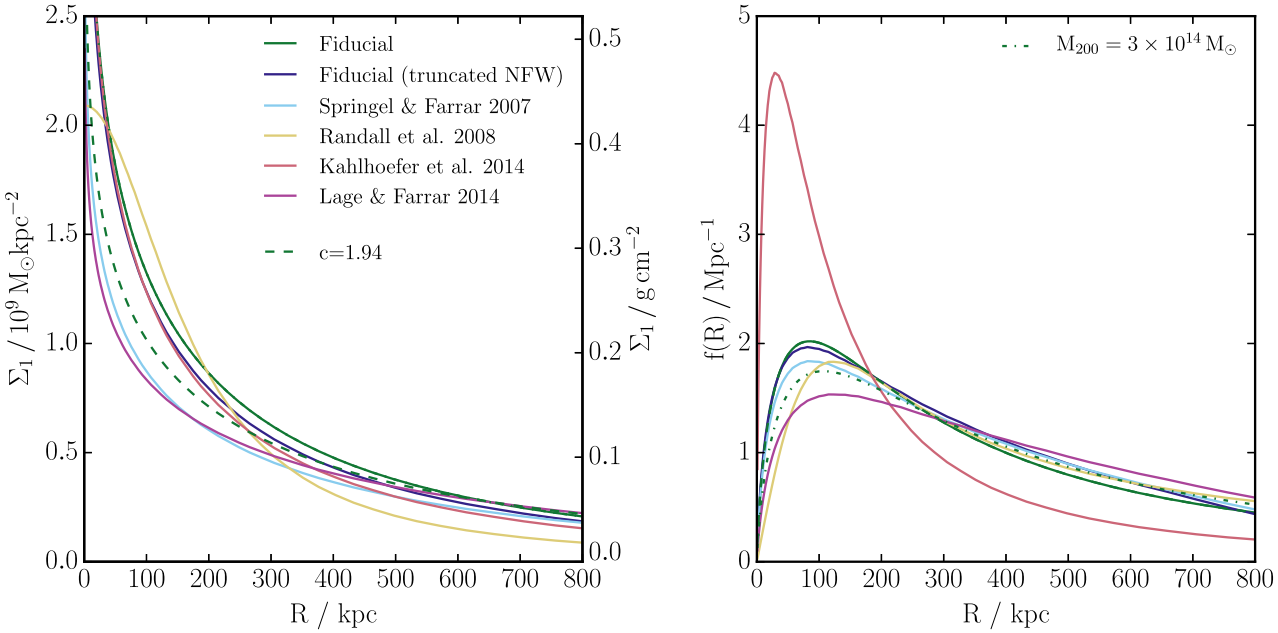

Figure 1. Left-hand panel: the projected mass density through the main halo of the Bullet Cluster, as a function of projected radius. Right-hand panel: the distribution of DM mass at different projected radii in the subcluster of the Bullet Cluster, normalized so that $\int_0^{1\,\mathrm{Mpc}} f(R)\mathrm{d}R = 1$. Different line styles and colours correspond to different choices for the density profiles. Our fiducial model is described in Section 2.2.3 while two variations to our fiducial model (c =1.94 and $M_{200} = 3 \times 10^{14}\,\mathrm{M}_\odot$) are described in Section 4.2. The fiducial (truncated NFW) lines are for the underlying NFW profiles that our fiducial model (which uses Hernquist profiles) are matched to, truncated so there is no mass outside of r_{200}. SF07 was the fiducial model used in that paper, while R08 is for the density profiles used in their simulation with σ/m =1.25 cm^2 g^{-1} (their initial conditions were changed slightly for different cross-sections). K14 only simulated one model for the Bullet Cluster, which had a particularly concentrated bullet halo as evident in the right-hand panel. Lage & Farrar (2014) was the best-fitting model found from running over a thousand simulations with different initial conditions and comparing the results to several observational data sets.

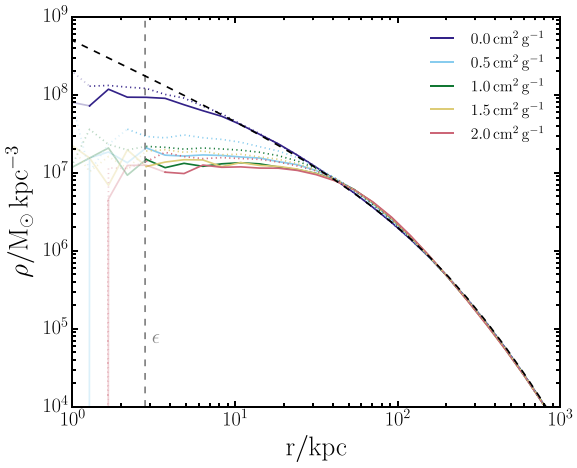

Figure 2. The radial density profile of an isolated halo with collisionless DM as well as SIDM with isotropic cross-sections ranging from 0.5 to 2 cm^2 g^{-1}. The dotted lines show the average profile between 1 and 2 Gyr after the start of the simulation, while the solid lines show the period 5–6 Gyr after the start of the simulation. Lines are semitransparent when the density corresponds to fewer than five particles in a radial bin. The vertical line corresponds to the Plummer-equivalent gravitational softening length, ϵ. For the collisionless DM, the initial Hernquist profile (dashed line) is stable, except for the formation of a numerical core with size $\sim 2\epsilon$ due to gravitational softening.

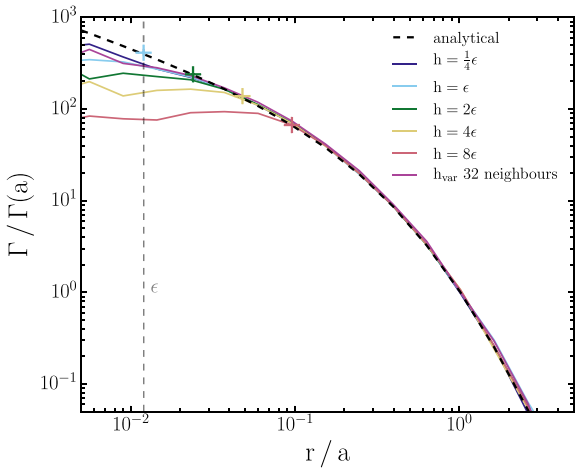

Figure 3. The scattering rate per particle in a Hernquist profile DM halo, plotted as a function of radius. Increasing the size of the search radius used for DM scattering leads to a decrease of the scattering rate in the inner regions of the halo. The results converge for $h < \epsilon$ as the density profile in the simulation forms a numerical core with radius $\sim\epsilon$ due to gravitational softening. All lines used a fixed h except for the h_{var} line for which h is varied for each particle to keep 32 neighbours within the search region. For the fixed h lines there are corresponding crosses plotted along the analytical curve at the radius equal to h, showing that for $r \lesssim h$ the scattering rate falls below the analytic result.

The scattering rate per particle as a function of radius was extracted from the simulations by taking the location of all scatters during the simulation and binning them in logarithmically spaced radial bins. This was then divided by the time-averaged number of particles within the same radial bins to get the scattering rate per particle.

2.3.1 *Analytical expectation for scattering rates in haloes*

For the analytical calculation of the expected scattering rate per particle, the density and the mean pairwise velocity need to be known. The density distribution is given in equation (8), while the pairwise velocities can be calculated from the velocity dispersion. Given isotropic velocities following a Maxwell–Boltzmann distribution, the mean pairwise velocity is given by $\langle v_{\rm pair} \rangle = (4/\sqrt{\pi})\sigma_{\rm 1D}$, where $\sigma_{\rm 1D}$ is the one-dimensional velocity dispersion. This can be calculated from the density profile and the Jeans equation, which (again assuming an isotropic velocity distribution) gives

$$\sigma_{\rm 1D}^2 = \frac{GM}{12a}\left\{\frac{12r(r+a)^3}{a^4}\ln\left(\frac{r+a}{r}\right) - \frac{r}{r+a}\left[25+52\frac{r}{a}+42\left(\frac{r}{a}\right)^2+12\left(\frac{r}{a}\right)^3\right]\right\}, \quad (12)$$

for a Hernquist profile.

Integrating over the velocity distribution function in equation (3) gives the average scattering rate for particles at position $\boldsymbol{r}$,

$$\Gamma(\boldsymbol{r}) = \frac{\langle \sigma\, v_{\rm pair}\rangle(\boldsymbol{r})\rho(\boldsymbol{r})}{m}. \quad (13)$$

If the DM cross-section is velocity independent then $\langle \sigma\, v_{\rm pair}\rangle = \sigma\langle v_{\rm pair}\rangle = \sigma(4/\sqrt{\pi})\sigma_{\rm 1D}$, and we can calculate the expected scattering rate per particle at different radii in the halo from equations (8) and (12). This is shown as the dashed line in Fig. 3.

2.3.2 *Scattering rates in simulated haloes*

As DM scattering leads to the formation of a cored density profile and also changes the velocity distribution, the scattering rate as a function of radius would not follow the analytical relation once the system has evolved due to self-interactions. To allow for a direct comparison to the analytical result we turn off the momentum kicks from scattering, such that the scattering algorithm is used to find particles that scatter, but does not actually change the particles' momenta as a result of scattering.

Fig. 3 demonstrates that our code reproduces the correct scattering rate within the halo at all but the smallest radii – where the scattering rate falls below the analytic prediction. This behaviour is easily understood by noting that the search radius for finding neighbours from which to scatter, h, acts as a scale on which the density and velocity distribution are smoothed in the calculation of scattering probabilities. The search radius therefore smooths away the density cusp in the scattering rate calculation leading to decreased scattering rates compared to the true unsmoothed rate. The scattering rate in the simulations drops significantly below the analytic rate only for radii less than h, so using a small h is preferred to capture the scattering rate in small high-density regions.

For h smaller than the gravitational softening length, ϵ, the radius within which Γ falls below the analytical result ceases to change. This is because there is a core formed in the particle distribution due to gravitational softening, with the core size of the order of ϵ. Pushing h to smaller values than ϵ therefore cannot recover the analytical result, because the particle distribution is already smoothed on the scale of the gravitational softening.

As using a small h leads to larger probabilities for pairs of particles to scatter (equation 6), smaller time-steps must be used to keep these probabilities below 1. We find that setting $h = \epsilon$ allows the usual dynamical time-steps to be used, while not excessively smoothing the density in the calculation of scattering rates.

Rocha et al. (2013) found that their scattering algorithm underpredicted scattering rates for small values of h in low-density regions. Specifically, they found the scattering rate dropped below the correct rate when $h\,(\rho/m_{\rm p})^{1/3} \lesssim 0.2$, i.e. when h is less than 20 per cent of the mean interparticle separation. For $h = 0.1$ kpc in Fig. 3, h is $\sim$4 per cent of the mean interparticle separation at $r = 2a$, but the scattering rate still matches the analytical prediction. The result found by Rocha et al. (2013) may be a result of allowing scattering probabilities within a time-step to go above 1. This is discussed further, along with some more tests of the SIDM implementation, in Appendix A.

3 MEASURING POSITIONS

In order to measure the offsets between different components, we first need a definition of position for each of the components. Observationally, the methods used to find the positions of the gas, galaxies, and DM are typically all different, and may also be different from the methods used to find the positions in associated simulations. It is therefore important that we understand the effects of changing the method used to find the positions of the various components, in a bid to understand how to best analyse the simulations in order to compare the results with observations.

3.1 Shrinking circles

The *shrinking-circles* approach to finding the position associated with a set of discrete points (the simulation particles) is the 2D analogue of the *shrinking-spheres* approach often used to find density peaks in N-body simulations (see e.g. Power et al. 2003). All of the particles under consideration are first projected along one axis. Then a circle is drawn. centred on the mean position of all particles, with radius chosen to be the distance between this centre and the most distant particle. The radius is then shrunk by a factor f and a new centre is calculated from the mean position of all particles within the current circle. The radius is shrunk again, and the process continues until the radius of the circle is $R_{\rm min}$. The mean position of all particles within this final circle gives the position of this set of particles.

This was the method employed by R08 who used $R_{\rm min} = 200$ kpc. This method clearly only gives one position for a distribution of particles, and so to get the position of both DM haloes from a simulation of the Bullet Cluster, the method needs to be run separately on particles belonging to the different haloes, or be started with the circles already shrunk to a size where they only contain one DM peak.

3.2 Parametric fits to 2D density maps

As an alternative to using shrinking circles to find the positions of the two haloes, we simultaneously fit the projected-density map with two profiles that have analytical projected densities. We use 2D projections of Pseudo Isothermal Elliptical Mass Distributions (PIEMDs), which have a 3D density profile

$$\rho(r) = \frac{\rho_0}{(1+r^2/r_{\rm core}^2)(1+r^2/r_{\rm cut}^2)}; \quad r_{\rm cut} > r_{\rm core}. \quad (14)$$

This profile has a core with central density ρ_0 and size $r_{\rm core}$, outside of which $\rho \propto r^{-2}$ as for an isothermal sphere, until $r \gtrsim r_{\rm cut}$ for which the density falls off as r^{-4}. This density profile is useful in these SIDM simulations, where the additional free parameter over an NFW or Hernquist profile, allows the cores produced by DM scattering to be well-fitted. The 3D potential and projected-potential are also analytical for this model, making it popular in gravitational-lensing analyses where deflection angles, shears, and convergence depend on gradients of the projected potential.

The projected density for a PIEMD is

$$\Sigma(R) = 2\int_R^\infty \frac{\rho(r) r}{\sqrt{r^2 - R^2}}\,{\rm d}r = \Sigma_0 \frac{r_{\rm core}\, r_{\rm cut}}{r_{\rm cut} - r_{\rm core}} \left(\frac{1}{\sqrt{r_{\rm core}^2 + R^2}} + \frac{1}{\sqrt{r_{\rm cut}^2 + R^2}} \right), \quad (15)$$

where R is the projected radius from the centre of the halo, and

$$\Sigma_0 = \pi \rho_0 \frac{r_{\rm core}\, r_{\rm cut}}{r_{\rm cut} + r_{\rm core}}. \quad (16)$$

As described in Kassiola & Kovner (1993), the axially symmetric projected-density profile in equation (15) can be made elliptical by substituting $R \rightarrow \tilde{R}$, where

$$\tilde{R}^2 = \frac{\tilde{x}^2}{(1+\epsilon)^2} + \frac{\tilde{y}^2}{(1-\epsilon)^2}, \quad (17)$$

and $\tilde{x}$ and $\tilde{y}$ are the spatial coordinates from the centre of the halo, along the major and minor projected axes of the halo, respectively. The ellipticity of the halo is defined as $\epsilon = (a-b)/(a+b)$, where a and b are the semimajor and semiminor axes. Along with ϵ, there is an additional parameter ϕ that describes the angle between the $\tilde{x}$-axis and the x-axis, i.e. the position angle of the major axis of the halo relative to our coordinate system (x, y).

We find the 2D positions of the DM haloes by simultaneously fitting two PIEMDs to the total projected DM density. We first discuss the case of fitting the distribution to a single isolated halo, the progression to two haloes then being relatively straightforward.

Each halo is described by seven parameters: the coordinates of the centre (X, Y), the central density ρ_0, the core radius $r_{\rm core}$, the outer radius $r_{\rm cut}$, the ellipticity ϵ, and the position angle ϕ. The distribution of simulation particles is split into evenly sized bins, generating the data map, d_{ij}, to which we find the best-fitting parametric model. Given values for the seven parameters that describe a PIEMD, the surface density can be calculated at each bin using $\Sigma(R)$ from equation (15) and using $R \rightarrow \tilde{R}$ calculated as the distance between the centre of each bin and the halo centre (X, Y) transformed according to equation (17). This would more accurately be done by integrating $\Sigma(x, y)$ over the area of the bin. As the density is roughly constant for $\tilde{R} < r_{\rm core}$, and our bin size used is smaller than the core radii found, the variation of Σ across any individual bin is small, and the mean surface density within a bin is well approximated by the surface density at the bin centre.

The model map, m_i, is the expected number of particles in each bin given the current parameter values, θ. This is simply the surface density at the bin position multiplied by the bin area, and divided by the mass of the simulation particles.

Once we have a data map and a model map, we can calculate the probability of getting our data map given the model map (i.e. the likelihood). For a bin with a given model value, we expect the data value to be Poisson distributed with the expectation value equal to the model value. The likelihood is the product over all map bins of the probabilities of obtaining each data value given the model value:

$$\mathcal{L}(\theta = \{X, Y, \rho_0, r_{\rm core}, r_{\rm cut}, \phi, \epsilon\}) = \prod_i \frac{m_i^{d_i}\, {\rm e}^{-{\rm m_i}}}{d_i!}. \quad (18)$$

We can combine this likelihood function with a set of priors to calculate posterior probabilities for the parameters. This is done using EMCEE (Foreman-Mackey et al. 2013), a PYTHON implementation of the affine-invariant ensemble sampler for Markov chain Monte Carlo proposed by Goodman & Weare (2010). We choose flat priors for X, Y, $r_{\rm core}$, and $r_{\rm cut}$, with a prior on ρ_0 that is flat in log-space.

In Fig. 4, we show the results of fitting two PIEMDs to a synthetic density map. The synthetic map was generated by taking the projected-density profile of two PIEMDs, here chosen to have parameters similar to that of the Bullet Cluster at the time at which it is observed, and then drawing a number of particles in each bin from a Poisson distribution with mean equal to the number of particles expected from the analytic profiles, assuming a particle mass equal to that used in our fiducial simulations.

The map in the top-right of Fig. 4 shows visually the level of noise associated with having a discrete set of particles and using 20 kpc bins, while the main corner plot shows that the fitting procedure recovers the input model within the error contours of the 2D projected posterior distributions. As the likelihood function in equation (18) is based upon Poisson statistics in each bin, the width of the posterior distributions shows the uncertainty in model parameters due to having a finite number of simulation particles. Of particular interest is the width of the posterior of the halo position along the collision axis (X), as it is the separation of different components along this axis that can be used to infer the DM cross-section. Using the particle mass used in our simulations, the width of the X posterior distribution is $\sim$2 kpc.

3.3 Parametric fits to shear maps

3.3.1 Generating shear maps

Although the projected density is technically observable through size and flux magnifications (as recently done in Duncan et al. 2016), weak lensing is usually done using the gravitational shear field. While the intrinsic ellipticities of galaxies are typically larger than the ellipticity from gravitational shear, with a large number of lensed galaxies the projected mass distribution of the lensing object can be determined.

The magnification of sources is described by the convergence, κ, while the distortion to the shape of galaxies is described by the shear (γ_1, γ_2). Here γ_1 describes stretching and squashing along the x-axis, while γ_2 describes these at 45° to the x-axis. In fact, the effect of lensing on galaxy ellipticities is described by the reduced shear, $g = \gamma/(1-\kappa)$. The quantities κ, γ_1, and γ_2 can all be related to the effective lensing potential, Ψ, through

$$\kappa = \frac{1}{2}\left(\frac{\partial^2 \Psi}{\partial x^2} + \frac{\partial^2 \Psi}{\partial y^2} \right), \quad (19)$$

$$\gamma_1 = \frac{1}{2}\left(\frac{\partial^2 \Psi}{\partial x^2} - \frac{\partial^2 \Psi}{\partial y^2} \right), \quad (20)$$

and

$$\gamma_2 = \frac{\partial^2 \Psi}{\partial x\, \partial y}. \quad (21)$$

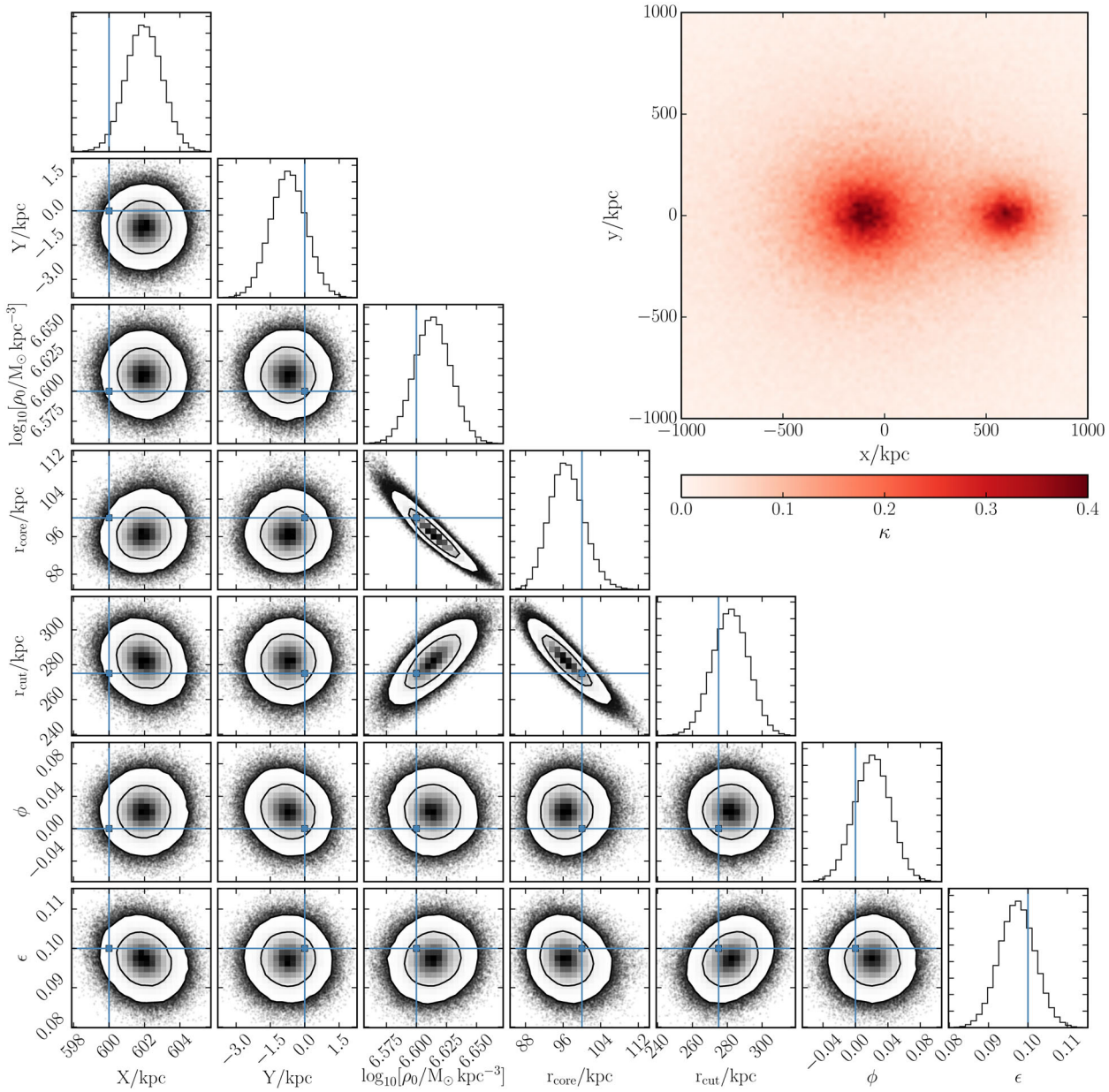

Figure 4. The posterior distributions for the model parameters of the smaller halo (on the right in the convergence map), found from simultaneously fitting two PIEMDs to the projected density generated from two model PIEMDs. The contours show 68 and 95 per cent confidence intervals. The model values used to generate the projected density are shown by the blue lines, and are recovered within the posterior distributions returned by the fitting procedure. The model values for the larger halo were also recovered, but are not shown here for clarity. The plot was made using CORNER.PY (Foreman-Mackey 2016).

The convergence is also given by the scaled projected density

$$\kappa(x, y) = \frac{\Sigma(x, y)}{\Sigma_{\rm crit}}, \tag{22}$$

where the critical surface density, $\Sigma_{\rm crit}$, is dependent on the geometry of the source, observer, and lens through

$$\Sigma_{\rm crit} = \frac{c^2}{4\pi G}\frac{D_{\rm s}}{D_{\rm l} D_{\rm ls}}, \tag{23}$$

where $D_{\rm s}$, $D_{\rm l}$, and $D_{\rm ls}$ are the angular diameter distances between the observer and the source, observer and lens, and lens and source, respectively.

Using equations (22) and (23), we can generate a κ map from a simulation snapshot by binning the 2D particle distribution, having projected along the third dimension. Using the number of particles in a bin, the particle mass, and the bin area, we can calculate a projected density, $\Sigma(x, y)$. Then, given a redshift for the lens (here, $z = 0.296$ as for the observed bullet cluster) and a redshift for the source galaxies (here, we use $z_{\rm s} = 1$), we can calculate the critical surface density, which for our choice of cosmology was $\Sigma_{\rm crit} = 2.85 \times 10^9$ M$_\odot$ kpc^{-2}.

Once we have a κ map, we can generate maps of γ_1 and γ_2 by making use of equations (19)–(21). Taking the Fourier transform of these equations, we find

$$\hat{\kappa} = -\frac{1}{2}(k_x^2 + k_y^2)\hat{\Psi}, \tag{24}$$

$$\hat{\gamma_1} = -\frac{1}{2}(k_x^2 - k_y^2)\hat{\Psi}, \tag{25}$$

and

$$\hat{\gamma_2} = -k_x k_y \hat{\Psi}, \tag{26}$$

where $\boldsymbol{k} = (k_x, k_y)$ is the wave vector conjugate to $\boldsymbol{x} = (x, y)$. These can be rearranged to give

$$\hat{\gamma_1} = \frac{k_x^2 - k_y^2}{k_x^2 + k_y^2}\hat{\kappa}, \tag{27}$$

and

$$\hat{\gamma_2} = \frac{2k_x k_y}{k_x^2 + k_y^2}\hat{\kappa}. \tag{28}$$

Finding γ_1 and γ_2 is then simply a case of taking the Fourier transform of κ, multiplying by the appropriate function of k_x and k_y, and taking the inverse Fourier transform to return the desired shear component. The two components of g are then given by these shear components divided by $1 - \kappa$.

3.3.2 Shear-map likelihood function

Given maps of the two reduced shear components generated from a simulation snapshot, $g^{\rm d}_{1,i}$ and $g^{\rm d}_{2,i}$, we can calculate a likelihood function

$$\mathcal{L}(\theta) = \prod_i \exp\left(\frac{(g^{\rm d}_{1,i} - g^{\rm m}_{1,i})^2}{2\sigma_\gamma^2}\right) \exp\left(\frac{(g^{\rm d}_{2,i} - g^{\rm m}_{2,i})^2}{2\sigma_\gamma^2}\right), \tag{29}$$

where $g^{\rm m}_{1,i}$ and $g^{\rm m}_{2,i}$ are the maps generated from the parametric model described by θ. When reconstructing a shear field from the ellipticities of lensed galaxies, the variance of each component of the shear field at a pixel, $\sigma_\gamma^2 = \sigma_{\rm int}^2 + \sigma_{\rm meas}^2$, comes from the intrinsic ellipticities of galaxies as well as shape-measurement errors. Shape-measurement errors depend on the quality of the data, as well as the method used to measure shapes, while the intrinsic ellipticities of galaxies are an unavoidable limitation to lensing measurements using gravitational shear. We thus set $\sigma_{\rm meas} = 0$ in this work, and assume that the only limitation to reconstructing a mass model using weak lensing comes from the number density of galaxies and the width of their intrinsic ellipticity distribution. Leauthaud et al. (2007) found that for each galaxy $\sigma_{\rm int} \sim 0.26$ across a wide range of sizes, magnitudes, and redshifts. Thus, given a number of lensed galaxies, N, within a pixel of a shear map, the contribution of intrinsic ellipticities to the average ellipticity of galaxies in that bin will be normally distributed with zero mean and standard deviation $\sigma_\gamma = 0.26/\sqrt{N}$.

In this work, we use a square shear map with a side length of 3 Mpc, centred on the centre of mass of the two haloes. We first produce a convergence map of this same area, and then generate a shear map from this, following the procedure described in Section 3.3.1. In order to avoid wraparound errors, the convergence map is zero-padded up to a side length of 10 Mpc. The posterior distribution for parameters describing two elliptical PIEMDs can then be calculated as for the projected density, using the likelihood function in equation (29), where σ_γ is calculated assuming a source-galaxy density of 80 galaxies arcmin^{-2}. We also mask out any pixels where $\kappa > 0.6$, as in these regions the reduced shear can become very large and then individual pixels dominate the likelihood, these regions are approaching or in the strong-lensing regime, and would not typically enter a weak-lensing analysis.

The result of fitting to a shear map generated from the projected-density profile in Fig. 4 is shown in Fig. 5. Unlike the case of fitting to the projected density, the width of the posterior distribution is no longer driven by the number of simulation particles, but by our greater uncertainty on the shear map from the intrinsic shapes of lensed galaxies. The synthetic shear map generated (and shown in the top-right of Fig. 5) did not include any shape noise, and so the posterior distributions returned are centred on the true model values. The width of the posterior describes the range of results one would expect to derive had there been shape noise, as demonstrated by the red dots which show the maximum likelihood parameter values for 20 different realizations of maps where Gaussian noise was added to the synthetic shear map, with the variance of the noise corresponding to σ_γ^2.

The width of the posterior distributions in Fig. 5 suggest that using gravitational shear with 80 galaxies arcmin^{-2}, we cannot determine the position of the bullet DM halo to better than $\sim\pm$40 kpc. This is consistent with Harvey et al. (2015) who found a typical 1σ error of 60 kpc on the DM halo positions determined from weak gravitational lensing with $\sim$60 galaxies arcmin^{-2} (Harvey, private communication).

4 RESULTS

4.1 Offsets with different cross-sections

As our fiducial method to measure the positions of both stars and DM, we fit two PIEMDs to the projected surface density as described in Section 3.2. Doing this independently for the DM and stellar component, we can then measure the offset along the collision axis between the two components. This was done for collisionless DM as SIDM with four different cross-sections. The offset between the stars and DM of the smaller bullet halo is shown as a function of the position of this halo in Fig. 6. This position was measured along the collision (x) axis, relative to the centre of mass of the two haloes. As the main halo is substantially more massive than the bullet halo, this position is similar to the separation between the two DM haloes. For collisionless DM, the observed DM halo separation of 720 kpc occurs when the bullet halo is at $X_{\rm DM} \approx 600$ kpc.

The offsets scale linearly with cross-section, in agreement with R08 and K14, but the size of the offsets for a given cross-section are considerably smaller than those found in R08, and about 40 per cent smaller than in K14. For $\sigma/m = 1$ cm^2 g^{-1}, the offset at the time of the observed Bullet Cluster is $\sim$10 kpc, whereas R08 find that a similar cross-section leads to the DM trailing the galaxies by almost 40 kpc. From the observed trailing of galaxies by DM of 25 ± 29 kpc, R08 placed constraints on the DM cross-section of $\sigma/m < 1.25$ cm^2 g^{-1}, whereas all of our simulated cross-sections would be consistent with this observation. This discrepancy is investigated in the following two sections, where we vary our initial conditions, and then the method used to measure positions.

4.2 Sensitivity to varying initial conditions

The offsets for different cross-sections depend on the initial conditions used, as changing the masses and concentrations of the haloes changes the rate of DM scattering as well as the gravitational forces that dominate the dynamics of the merger. In this section, we investigate changing the initial conditions. We vary one parameter at a time from its value in our fiducial model, and change the parameters in a way that has been used in previous simulations or has been hinted at by previous results.

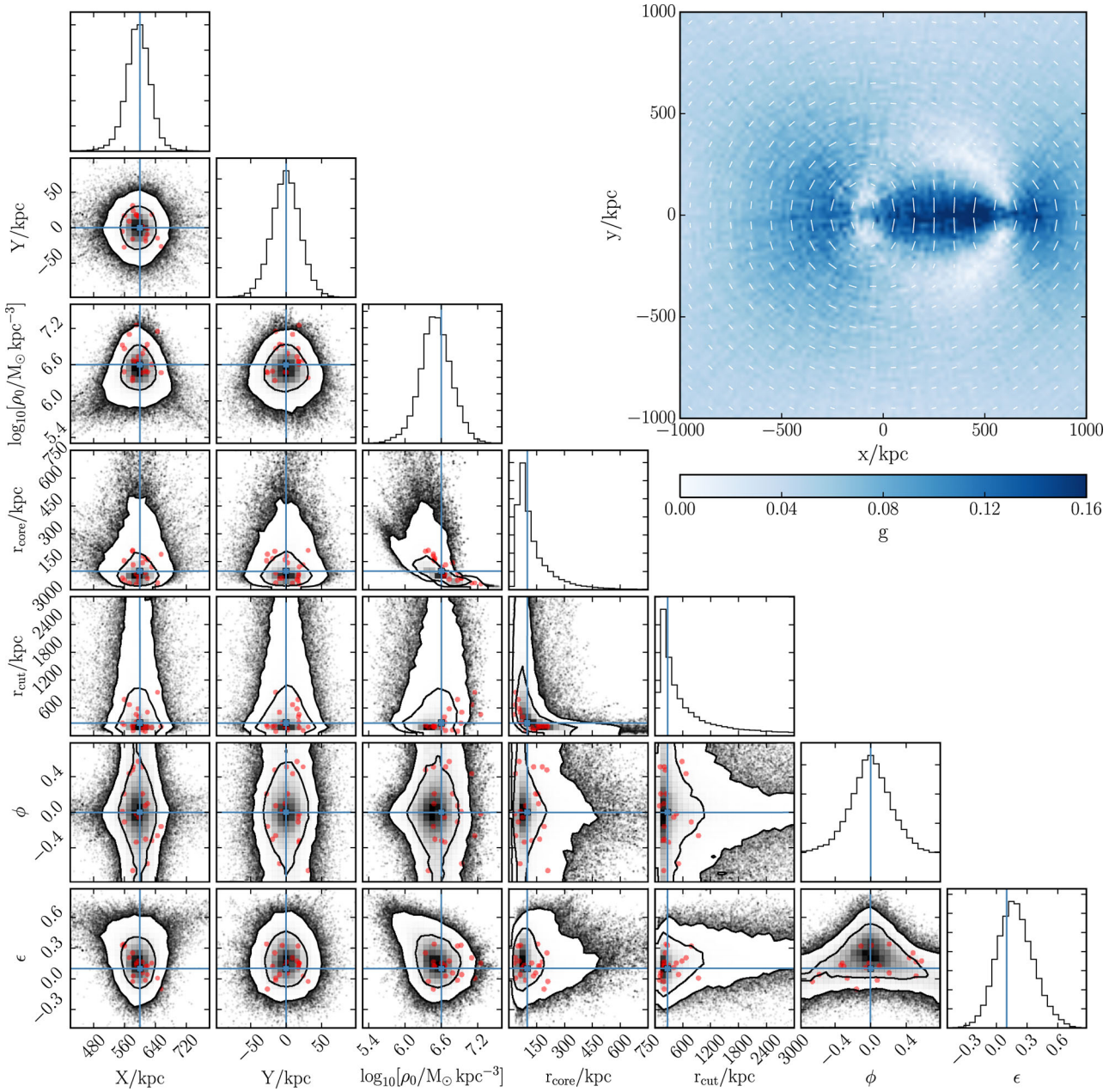

Figure 5. Similar to Fig. 4, but for the case of fitting to reduced gravitational shear. The red points indicate the maximum likelihood parameter values found from fitting to the underlying shear map from the model with the addition of 20 different realizations of noise from source-galaxy intrinsic ellipticities. In the shear map in the top-right the colour represents the value of the reduced shear, while the white lines show the direction.

4.2.1 Main-halo concentration

The first parameter we vary is the concentration of the main halo. Our fiducial model used $c = 3$, as this was found to be required by SF07 to reproduce the observed offset between the bright X-ray bullet and the associated DM halo. This result used the rather limiting assumption (as used in this work) that the gas density initially follows the DM density. Lage & Farrar (2014) used more complicated models for their initial conditions, with triaxial DM haloes, and a seven-parameter model for the gas profile of each halo. They found their best-fitting model to have $c = 1.17 \pm 0.14$ for the main halo, which would put this halo well below the median concentration–mass relation. To investigate how a low initial concentration for the main halo affects our results, we ran simulations with an initial concentration for the main halo of $c = 1.94$, which was the best-fitting concentration for the main halo measured after the collision via weak lensing.

The resulting offsets with collisionless DM and SIDM with $\sigma/m = 1\,\mathrm{cm^2\,g^{-1}}$ are shown in Fig. 7. The offset with SIDM is reduced relative to the fiducial model, which is to be expected given that with a lower concentration, the projected density through the centre of the halo is reduced. This means that particles in the bullet halo, which has a zero impact parameter and passes through the centre of the main halo, pass through less DM and are less likely to scatter from particles in the main halo. In fact, the fraction of DM particles from the bullet halo, which scatter with particles from

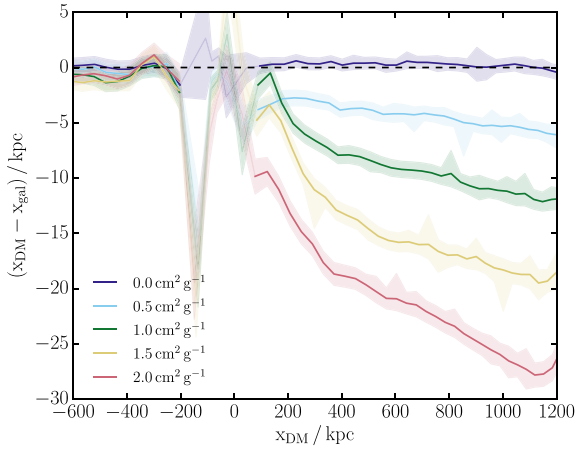

Figure 6. The offset between the stellar (galaxy) and DM component of the bullet halo for different SIDM cross-sections, where both the stellar and DM positions were determined by simultaneously fitting two PIEMDs to the respective projected maps. The offsets scale linearly with DM cross-section, and at the time of the observed bullet cluster the DM trails the galaxies by $\sim$10 kpc for $\sigma/m = 1\,\mathrm{cm^2\,g^{-1}}$. Lines are semitransparent around the time of core passage (which due to tidal forces happens at $X_{DM} \approx -80$ kpc before the centres of mass of the two haloes meet) due to a degeneracy in the positions of the two haloes leading to spurious offsets.

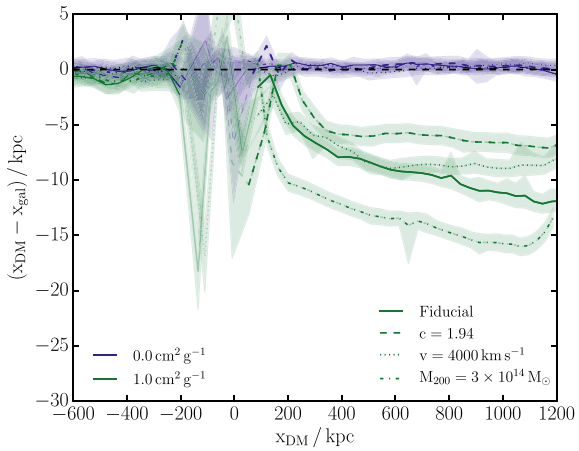

Figure 7. DM–galaxy offsets as in Fig. 6, but with four different sets of initial conditions, each run with collisionless DM and SIDM with $\sigma/m = 1\,\mathrm{cm^2\,g^{-1}}$. These initial conditions are described in Section 4.2, but in summary are as follows: compared to our fiducial model, '$c = 1.94$' has a lower concentration for the main halo, '$v = 4000\,\mathrm{km\,s^{-1}}$' has an increased relative velocity between the two haloes, and '$M_{200} = 3 \times 10^{14}\,M_\odot$' has a more massive bullet halo.

the main halo, drops from 13 per cent for our fiducial model to 10 per cent, in broad agreement with estimates of the scattering fraction that can be made from Fig. 1.

4.2.2 *Relative velocity between haloes*

As discussed in Section 2.2.2, the shock velocity in the observed Bullet Cluster is $4700 \pm 600\,\mathrm{km\,s^{-1}}$. In previous works using the Bullet Cluster to constrain SIDM (R08, K14), this has been used as the relative velocity between the two DM haloes, despite hydrodynamical simulations showing that the relative velocity of the shock front and pre-shocked gas in Bullet Cluster-like simulations

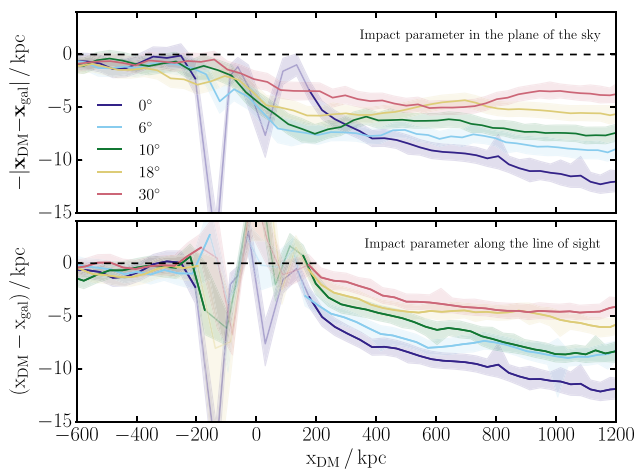

Figure 8. DM–galaxy offsets with $\sigma/m = 1\,\mathrm{cm^2\,g^{-1}}$ and four different impact parameters, as well as a head-on collision. The runs are labelled by the angle between the separation of the two haloes and the velocity of the bullet halo measured at the time of the observed Bullet Cluster, θ_{obs}. The fitting was done by simultaneously fitting two PIEMDs to the projected mass distribution. In the top panel the impact parameter was in the plane of the sky, while in the bottom panel it was along the line of sight.

is significantly higher than the relative velocity of the DM haloes (Milosavljević et al. 2007; SF07; Lage & Farrar 2014).

In Fig. 7, we show how the offset between the bullet DM halo and galaxies changes when the collision velocity is increased. We start the haloes with a relative velocity of $4000\,\mathrm{km\,s^{-1}}$ at a separation of 4 Mpc, which leads to a relative velocity of $4700\,\mathrm{km\,s^{-1}}$ at the time of the observed Bullet Cluster. This is in contrast with our fiducial model, where haloes start with the velocity corresponding to falling from rest at infinite distance, and the relative velocity between DM haloes is $3900\,\mathrm{km\,s^{-1}}$ at the time of the observed Bullet Cluster. We find that the offsets are not very sensitive to this change in relative velocity.

4.2.3 *Mass of bullet halo*

The weak-lensing-derived mass for the bullet halo of $M_{200} = 1.5 \times 10^{14}\,M_\odot$ is low in comparison with the strong-lensing results (Bradač et al. 2006) that calculate the mass in a 300 kpc radius cylinder centred on the bullet halo's brightest cluster galaxy to be $3 \times 10^{14}\,M_\odot$. While this is the total mass in this region, and includes a contribution from the main halo, this is still suggestive that the weak-lensing mass may be an underestimate. Simulations that have looked to reproduce the gas morphology and luminosity have also found best-fitting mass ratios for the merger between 7:1 and 5:1 (Milosavljević et al. 2007; Mastropietro & Burkert 2008; Lage & Farrar 2014).

For these reasons, we run simulations with an increased mass for the bullet halo of $M_{200} = 3 \times 10^{14}\,M_\odot$, keeping the concentration the same as in our fiducial model. This leads to a significant increase in the separation between DM and galaxies in the bullet, with the offset at the time of the observed Bullet Cluster increasing from 10 kpc for our fiducial model to 14 kpc.

4.2.4 *Impact parameter*

While the gas morphology implies a collision that was close to head-on, the bright gas bullet is not located precisely along the

Table 1. Summary of non-zero impact parameter simulations.

r_* (kpc)	$\theta_{\rm init}$ (°)	$b_{4\,\rm Mpc}$ (kpc)	$r_{\rm min}$ (kpc)	$\theta_{\rm obs}$ (°)
0	0	0	0	0
12.5	3.2	224	102	6
25	4.5	316	153	10
50	6.4	447	236	18
100	9.1	632	354	30

line connecting the centres of the two cluster haloes, suggesting a small non-zero impact parameter. We therefore run simulations with off-centre collisions, and investigate how sensitive the DM–galaxy offsets are to this change.

We continue to start the simulations with the two haloes separated by 4 Mpc and on a zero-energy orbit, but rotate the velocities of the haloes by $\theta_{\rm init}$ with respect to the x-axis that connects the two halo centres (keeping the velocities of the two haloes antiparallel). We choose $\theta_{\rm init}$ such that the two haloes would have a closest approach of r_* if they behaved as point masses throughout the merger. The force between the two haloes is reduced (compared to the case of point masses) when their mass distributions overlap, so that the actual minimum separation between the halo centres, $r_{\rm min}$, is significantly larger than r_*.

We summarize our different impact parameter runs in Table 1. Along with r_*, $\theta_{\rm init}$, and $r_{\rm min}$, we include the perpendicular distance between the two haloes' velocities when they are separated by 4 Mpc, $b_{4\,\rm Mpc}$, and the angle between the halo–halo separation and the bullet-halo velocity at the time of the observed Bullet Cluster, $\theta_{\rm obs}$. Assuming that gas is stripped in the opposite direction to the direction of motion, $\theta_{\rm obs}$ should roughly correspond to the angle between the DM–gas separation in the bullet halo and the DM–DM separation between the two haloes.

In Fig. 8, we plot the separation between the DM and galaxies with different impact parameters. In the top panel, the impact parameter is in the plane of the sky, while in the bottom panel it is along the line of sight, and the collision appears as if head on. Note that in the top panel we measure the 2D offset between the DM and galaxies, as this offset is no longer along the x-axis.

We find that moderate impact parameters only have a small effect on the DM–galaxy offsets. SF07 found that $r_* < 12.5$ kpc to avoid a gas distribution that is more asymmetric than that observed, while Mastropietro & Burkert (2008) found that an impact parameter, $b = 150$ kpc, gave the best match to the gas morphology and relative X-ray brightness of the two gaseous haloes. This means that even our smallest non-zero impact parameter is large compared to that used for the best-fitting results from other simulations of the Bullet Cluster, and so we expect any impact parameter consistent with the observed Bullet Cluster to decrease the DM–galaxy offset by less than 20 per cent.

4.3 Offsets with different position measures

Having found that the offset results are reasonably insensitive to the choice of initial conditions, in Fig. 9 we show the effect of using different methods to measure positions. For both collisionless DM and SIDM with $\sigma/m = 1\,{\rm cm^2\,g^{-1}}$, we measured the separation between the stellar and DM components of the bullet halo using the methods described in Section 3. For the shrinking circles and projected-density measurements, the same method was used for finding the position of both the stars and DM, while for the shear measurement the separation is that between the stellar halo

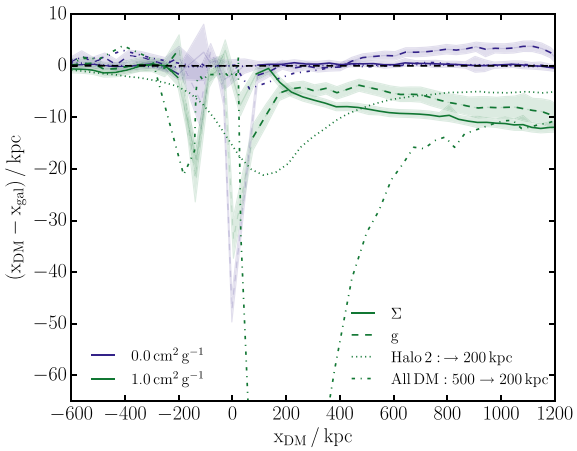

Figure 9. DM–galaxy offsets as in Fig. 6, but measured using different methods: fitting to the projected surface density (Σ), the reduced gravitational shear (g), and two different shrinking-circles techniques. For all methods but g, the same method was used to find the position of both the stars and the DM, while for g it was only the DM measured using shear with the stars being measured using Σ. The two shrinking-circles techniques are described in Section 4.3.

measured by fitting to the projected density and the DM halo measured using reduced gravitational shear. Fig. 5 demonstrates that with 80 galaxies arcmin^{-2}, the position of the bullet halo can only be determined to ± 40 kpc. As this uncertainty is larger than the offsets for any of our simulated cross-sections, detecting SIDM using weak lensing and the Bullet Cluster alone would not be possible. In Fig. 9, the lines derived from reduced gravitational shear used 8000 galaxies arcmin^{-2}, giving errors indicative of what could be achieved with ∼100 Bullet Cluster-like systems.

As discussed in Section 3.1, the shrinking-circles procedure will only return one position for a distribution of particles. We therefore use two different approaches to return the position of the bullet halo, both of which shrink the circle down to a final size of $R_{\rm min} = 200$ kpc as used by R08. The first method (*Halo 2*) is to apply the shrinking-circles procedure to only the particles that were originally part of the bullet halo. The second method (*All DM*) is to apply the shrinking-circles procedure to all of the DM, but starting with a circle centred on the second halo, as determined by Halo 2, with a starting radius of 500 kpc.

The different methods for measuring positions give very different results for the same SIDM cross-section, highlighting the importance of matching the analysis to what is done observationally. The offsets measured for $\sigma/m = 1\,{\rm cm^2\,g^{-1}}$ using different methods can be as different as the offsets for the different cross-sections shown in Fig. 6, particularly soon after core passage. Of particular note is the large offsets measured using shrinking circles on all of the DM. This method was also highly sensitive to the choice of starting position and starting radius, suggesting it is not a robust way to measure offsets from simulations. As a method similar to this was used by R08, this explains the large offsets and tight constraints on the DM cross-section that they found.

The sensitivity to the method used to measure the positions can be understood when one considers that after core passage of the two haloes, there are three distinct sets of DM particles: those originally from the main halo that have not interacted with any particles from the sub halo, those originally from the sub halo that have not interacted with any particles from the main halo, and particles from one halo that have scattered with a particle from

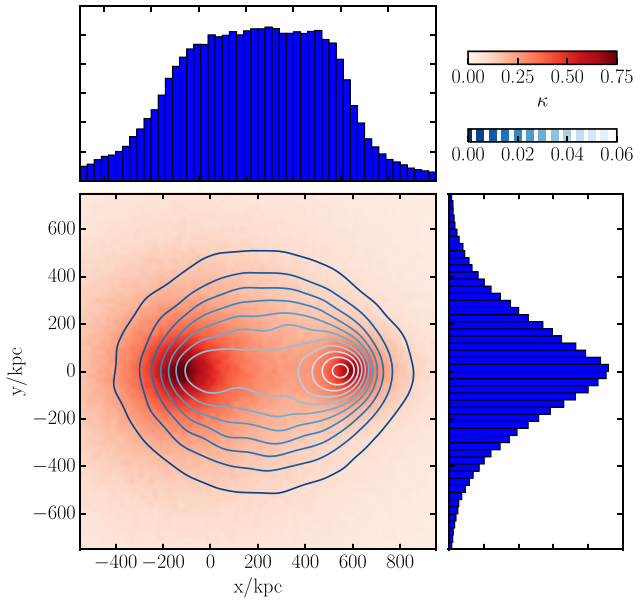

Figure 10. For SIDM with $\sigma/m = 1\,\mathrm{cm^2\,g^{-1}}$, the projected density of all DM in red, with the projected density of DM particles that have scattered with a particle originally from the other DM halo shown in blue contours, and also projected along the axes and shown as 1D histograms. The total mass in these scattered particles is $6.5 \times 10^{13}\,\mathrm{M_\odot}$, corresponding to 13 per cent of particles from the bullet halo scattering with particles in the main halo.

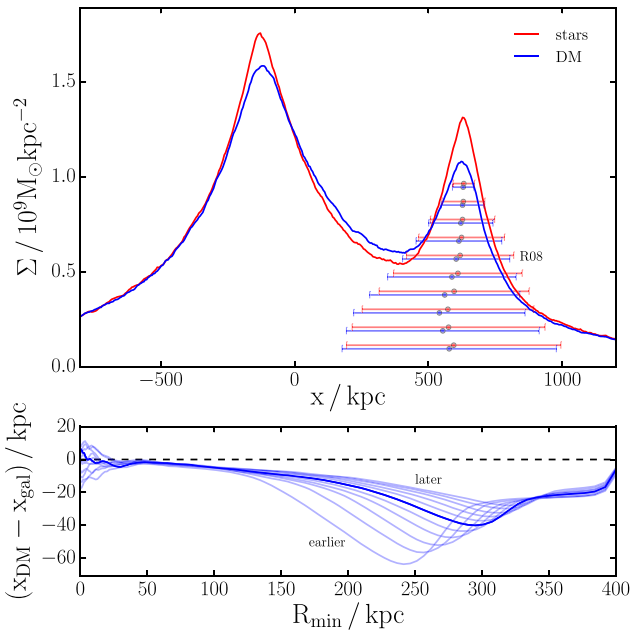

Figure 11. Top panel: the average projected surface density measured in a 400 kpc strip centred on the collision axis, for both the stars and DM (the stellar surface densities have been scaled up so that the mass in stars matches that in DM). The points show the position of the bullet halo returned by the shrinking circles procedure with different R_{min}, with the width of horizontal bars being twice R_{min}. Bottom panel: the DM–galaxy offset as a function of R_{min}. The dark line corresponds to the top panel (when the two haloes are separated by $\sim$720 kpc, and with $\sigma/m = 1\,\mathrm{cm^2\,g^{-1}}$), while the lighter lines are for successive snapshots separated by 10 Myr.

the other halo.[5] The momentum transfer between the two haloes caused by isotropic DM elastic scattering acts differently to the stripping of gas due to hydrodynamical forces, as only a subset of DM particles receive a momentum kick. These particles then lag behind the halo from which they came, gravitationally pulling it back, but they do this equally to unscattered DM particles and galaxies, and so do not lead to an offset between unscattered DM particles and the collisionless galaxies. Any offset found between the DM and galaxies is a result of fitting the wake of scattered particles and so depends sensitively on how positions are measured.

For $\sigma/m = 1\,\mathrm{cm^2\,g^{-1}}$, we show the projected DM density at the time of the observed Bullet Cluster in Fig. 10, along with the distribution of particles that have scattered with a particle from the other halo. For this cross-section and our fiducial initial conditions, 13 per cent of particles from the bullet halo scatter with particles from the main halo. The distribution of these scattered particles is quite broad, with the highest projected density of scattered particles being only 10 per cent of the total projected density at the same location.

4.3.1 The position returned by shrinking circles to different final radii

To illustrate the problems with using a shrinking-circles procedure to measure the positions of stars and DM, we show an example in Fig. 11, run on our fiducial simulation with $\sigma/m = 1\,\mathrm{cm^2\,g^{-1}}$ at the time of the observed Bullet Cluster (the same snapshot used for Fig. 10). The position returned for both the stars and DM varies as a function of R_{min}, with the offset between the stars and DM also depending sensitively on R_{min}.

An initial position for each of the stellar and DM components of the bullet halo is made by running shrinking circles on only particles that were part of the bullet halo in the initial conditions. The initial radius used was 400 kpc, a bit over half of the separation between the two DM haloes. Initially, as the circles are shrunk and recentred, they shift left due to the gradient in density coming from the main halo. As this gradient is steeper closer to the main halo, the DM position (that initially lies to the left of the stellar position) is affected more by the presence of the main halo, which in turn leads to spuriously large offsets. As the circles are shrunk further, they centre in on a region dominated by the bullet halo, and the offsets decrease. Shrinking down to $R_{\mathrm{min}} \lesssim 50$ kpc, the results become noisy as the number of particles involved in the position estimate decreases, and there is no clear density peak (with $\sigma/m = 1\,\mathrm{cm^2\,g^{-1}}$, the core size of the bullet halo is $\sim$100 kpc, though this is less obvious in the top of Fig. 11 due to projection effects).

Even before the results become noisy, the offsets between the stellar and DM peaks become very small, in agreement with K14 who found that the peaks in stellar and DM projected density were perfectly coincident when DM scattering was isotropic. This raises the question of whether any constraints can be placed on *isotropic* SIDM from looking at separations between local galaxy and DM peaks in colliding clusters. That being said, most studies that look for offsets between galaxies and peaks in free-form lensing reconstructions, either bin lensed galaxies (Ragozzine et al. 2012; King et al. 2016) effectively smoothing the DM distribution on some scale, or use a regularization scheme (e.g. Bradač et al. 2006), such

[5] For particles that are involved in an interhalo scattering event, particles from the two haloes are indistinguishable when the scattering cross-section is isotropic.

that the diffuse cloud of scattered particles (Fig. 10) could shift the derived DM peak back and lead to a measurable offset.

From the bottom panel of Fig. 11, it is clear that $R_{\rm min} = 200\,{\rm kpc}$ can give misleadingly large offsets, which explains the tight constraints on the DM cross-section found by R08. What is also clear is that there is no good choice for $R_{\rm min}$, as the results do not converge as $R_{\rm min}$ is decreased. For these reasons, we fit parametric models to our haloes in this paper, as is often done in gravitational-lensing analyses (Smith et al. 2005; Richard et al. 2010; George et al. 2012; Harvey et al. 2015; Massey et al. 2015; Shu et al. 2016). While this does not directly relate to what was done in Bradač et al. (2006), where strong and weak lensing were combined to produce a non-parametric mass model of the Bullet Cluster, a mock strong-lensing analysis is beyond the scope of this paper. We cannot do strong lensing with our simulations, as the surface density of our simulated bullet halo does not exceed the critical surface density for a lens at the Bullet Cluster's redshift. The absence of strong lensing with SIDM was noted by Meneghetti et al. (2001), who found that with moderate cross-sections of 0.1–1 $\rm cm^2\,g^{-1}$, the number of radial and giant-tangential arcs would fall well below what is observed. However, they point out that even with a collisionless DM simulation, the number of strong-lensing features falls below what is observed, and that bright central galaxies probably play an important role in generating strong-lensing features. While this is certainly an interesting avenue to constrain SIDM, without including the effects of galaxy-formation physics in our simulations, and with these simulations starting from idealized initial conditions, our work is not suited to testing whether the presence of strong-lensing features can constrain the DM cross-section.

4.4 Offsets including gas

So far, the results have been from simulations without any gas. However, real galaxy clusters have significant gas fractions. While there is less gas than DM, the additional hydrodynamic forces that act on the gas can alter the dynamics of merging clusters. In this section, we look at the changes from the previous results when each halo contains an adiabatic gas component making up 16 per cent of the total halo mass.

The resulting offsets between stars and DM are shown in Fig. 12. The offsets measured for $\sigma/m = 1\,{\rm cm^2\,g^{-1}}$ remain largely unchanged, with a small decrease (compared with the gas-free case) in the offset measured by fitting to the projected surface density. This results from the decreased optical depth for scattering as particles pass through the main halo, owing to ∼16 per cent of the DM mass now being in the form of gas. Most strikingly, there is now a significant offset measured with collisionless DM when measuring the DM position using gravitational shear. This is surprising, particularly as our DM and stars have the same phase space distribution at all times in our collisionless DM run, so this offset is a result of different fitting methods returning substantially different position estimates.

4.4.1 Explaining offsets with collisionless DM

In the top row of Fig. 13, we plot the projected DM distribution, and resulting shear field, at the time of the observed Bullet Cluster, but only using particles that were part of the bullet halo in the initial conditions. What is clear from the projected density is that the mass distribution is not elliptically symmetric, with the peak on small scales being shifted to the left of (i.e. lagging behind) the centre

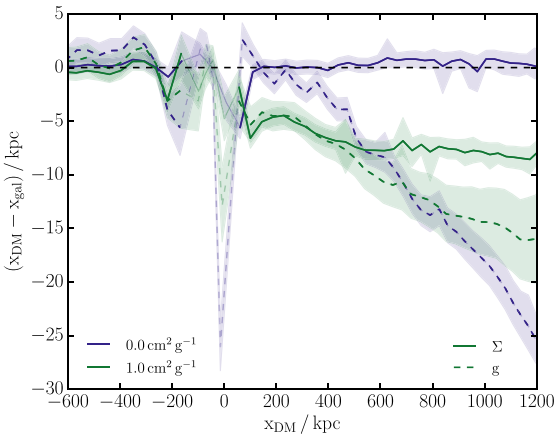

Figure 12. The offset between the stars and DM from simulations including adiabatic gas. As in Fig. 9, the g measurement is the offset between the DM position measured using reduced shear and the stellar position measured by fitting to the projected density of stars. As such, the large offset with collisionless DM which is not seen when both the stellar and DM positions are measured from their projected densities, means that fitting to the projected density or reduced shear of the same mass distribution, can lead to strongly differing results.

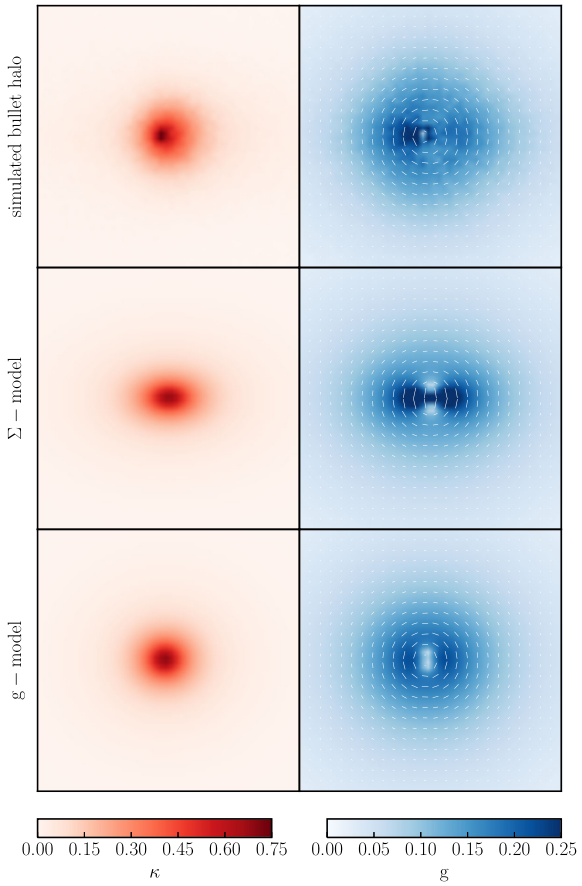

Figure 13. The convergence (left-hand column) and reduced shear (right-hand column) due to the bullet-halo DM, for a simulation with collisionless DM and non-radiative gas. The top row shows the simulation output (only including DM particles that are part of the bullet halo in the initial conditions), while the middle and bottom rows show the best-fitting maps generated by fitting to the projected surface density and reduced gravitational shear, respectively. Each panel is 1 Mpc across.

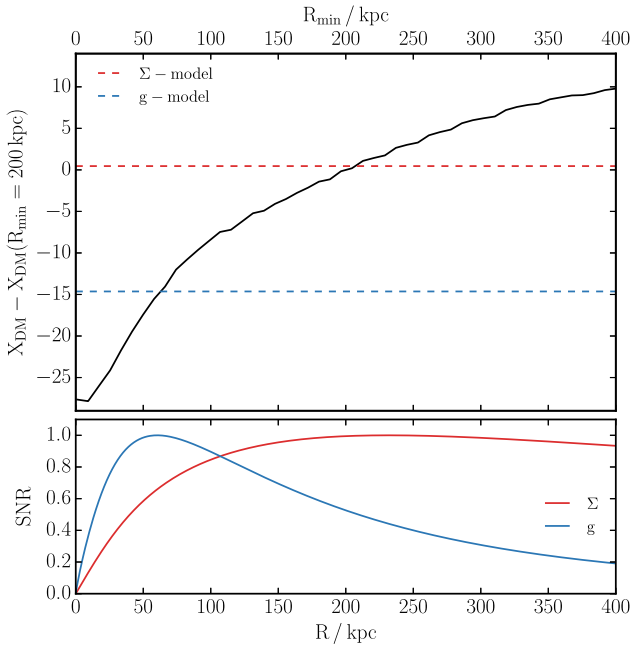

Figure 14. Top panel: the x-position of the bullet DM halo measured using shrinking circles on only the DM particles that are part of the bullet halo in the initial conditions. The position is shown as a function of the radius to which the circle is shrunk, with the DM halo shifting to the left as the measurement is made on smaller scales. The best-fitting positions of the bullet DM halo from fitting to the projected density and reduced gravitational shear are also shown. Bottom panel: the signal-to-noise integrated over the projected density or reduced shear map, due to mass within an annulus of fixed width at radius R. This was calculated using the projected density as a function of R from the best-fitting model to the projected density. Further details are in Appendix B.

of mass measured on larger scales. This is quantified in the top panel of Fig. 14 where we show the position returned by applying the shrinking circles algorithm on the DM particles from the bullet halo, shrinking down to different final radii, $R_{\rm min}$.

The middle and bottom rows of Fig. 12 show the best-fitting maps from fitting to the projected surface density and reduced gravitational shear, respectively. The projected-surface-density fit favours a more elliptical halo, centred further to the right, than the shear fit. This, combined with the fact that the halo position shifts left when measuring on smaller scales, suggests that reduced shear is more sensitive to the inner regions of the halo, whereas the projected-density fit is more sensitive to larger scales. In the bottom panel of Fig. 14, we show that this is what is expected, plotting (for both Σ and g) the sum of the signal-to-noise ratio over the whole map, due to annuli of mass at different radii. The details of this are explained in Appendix B. We find that this quantity peaks at $R \sim$60 kpc for reduced shear and $R \sim$230 kpc for the projected density, in rough agreement with the shrinking circles $R_{\rm min}$ that returns the same position as the respective fitting procedure.

Asymmetry in the DM distribution, and consequent differences in the positions returned by different fitting methods is most pronounced for the collisionless DM case as the cuspy halo is tightly bound to the gas. The formation of DM cores with SIDM reduces the strength of this gravitational binding, such that when the gas is stripped with SIDM, it does not drag back the central regions of the DM halo as strongly as with collisionless DM. The stripping of gas is just one mechanism that could cause an asymmetric DM profile, but serves as a cautionary tale for attempts to use offsets between different cluster components to constrain DM's collisional properties. The general result that an asymmetric profile can lead to a measured offset between spatially coincident components, due to them being measured using techniques sensitive to different scales, is an important systematic to consider in future studies.

4.4.2 *Changes to the gas morphology*

Aside from its effect on the DM distribution, the gas itself could potentially be used as a probe of DM self-interactions. Unfortunately, changes to the gas morphology as the DM cross-section is changed are fairly small, with the largest differences being the width and temperature of the shocked region. Increasing the DM cross-section lowers the luminosity-weighted projected temperature in the shocked-gas region, from 30 keV with collisionless DM, to 25 keV with $\sigma/m = 2\,{\rm cm}^2\,{\rm g}^{-1}$, both well within the quoted observational error (Markevitch 2006).

This decrease in temperature also comes with an increase in the width of the shocked region. The distance between the shock front and the contact discontinuity connecting the shocked gas to the cold gas bullet, increases from 70 kpc with collisionless DM to 110 kpc with $\sigma/m = 2\,{\rm cm}^2\,{\rm g}^{-1}$. While this latter value is in better agreement with the observed distance between the shock front and contact discontinuity ($\sim$140 kpc), we find (in agreement with SF07) that this distance is highly sensitive to the concentration of the main halo, making this measurement unsuitable for constraining the DM cross-section.

5 CONCLUSIONS

We have presented modifications to the GADGET-3 code to include elastic DM scattering, allowing us to run simulations with SIDM. We have shown that this code performs as expected when used for simple test cases where the correct behaviour can be predicted analytically. We have also discussed the choice of the numerical parameter, h, which is the radius within which particles look for neighbours to scatter with. We have shown that the choice $h \sim \epsilon$ leads to the correct scattering rate within a DM halo, without the computational overhead associated with having an environmentally dependent h.

We have used this code to perform idealized simulations of Bullet Cluster-like systems. With SIDM, the momentum transfer from particles in the main halo to particles in the bullet halo with which they scatter, leads to a tail of scattered particles in the bullet halo that shifts the measured position of this halo relative to the collisionless stars. Our fiducial model for the Bullet Cluster was derived from fits to weak-lensing data. Changes to this fiducial model led to changes in the measured offsets between stars and DM, although these changes were small and in a predictable manner.

Our primary conclusion is that the method used to measure the positions of the different components can have a larger effect than using a different model for the Bullet Cluster. In particular, shrinking-circles methods similar to those used by R08 give substantially larger DM–galaxy offsets than more observationally motivated methods, such as parametric fits to the projected density or reduced gravitational shear. This suggests that the $\sigma/m < 1.25\,{\rm cm}^2\,{\rm g}^{-1}$ constraint placed on the cross-section for DM scattering by R08 is strongly overstated. In fact, for our fiducial model of the Bullet Cluster with $\sigma/m = 2\,{\rm cm}^2\,{\rm g}^{-1}$, the DM–galaxy offset at the time of the observed Bullet Cluster is $\sim$20 kpc, which is allowed by the

25 ± 29 kpc observed offset used by R08 to place their constraint. We produce more robust results by fitting parametric models to the haloes – which can be done observationally (Smith et al. 2005; Richard et al. 2010; George et al. 2012; Harvey et al. 2015; Massey et al. 2015; Shu et al. 2016). We recommend that future simulation efforts adopt this, or similarly motivated techniques, to enable a better comparison to observations.

We went on to show results from the first simulations of merging clusters to include both SIDM and gas. The gas does not have much effect on the offset between the stellar and DM components. However, as the gas is stripped it introduces asymmetries into the stellar and DM components, with the central regions of the bullet halo lagging behind the larger-scale centre. This is strongest with collisionless DM where the cuspy halo is tightly bound to the gas. As the methods used observationally to measure the positions of the galaxies and DM will be different, they are likely to be sensitive to different scales. We showed that this can result in a measured offset between these two components, even if they have an identical spatial distribution. These asymmetric halo shapes could also be produced by tidal forces or dynamical friction, and these asymmetries are an important potential systematic that could lead to the false detection of SIDM.

ACKNOWLEDGEMENTS

This work was supported by the Science and Technology Facilities Council grant numbers ST/K501979/1 and ST/L00075X/1. RM was supported by the Royal Society. The authors wish to thank Richard Bower and David Harvey for useful discussions.

This work used the DiRAC Data Centric system at Durham University, operated by the Institute for Computational Cosmology on behalf of the STFC DiRAC HPC Facility (www.dirac.ac.uk). This equipment was funded by BIS National E-infrastructure capital grant ST/K00042X/1, STFC capital grants ST/H008519/1 and ST/K00087X/1, STFC DiRAC Operations grant ST/K003267/1, and Durham University. DiRAC is part of the National E-Infrastructure.

REFERENCES

Ackerman L., Buckley M. R., Carroll S. M., Kamionkowski M., 2009, Phys. Rev. D, 79, 023519
Barrena R., Biviano A., Ramella M., Falco E. E., Seitz S., 2002, A&A, 386, 816
Bento M. C., Bertolami O., Rosenfeld R., Teodoro L., 2000, Phys. Rev. D, 62, 041302
Boddy K. K., Feng J. L., Kaplinghat M., Tait T. M. P., 2014, Phys. Rev. D, 89, 115017
Bradač M. et al., 2006, ApJ, 652, 937
Bradač M., Allen S. W., Treu T., Ebeling H., Massey R., Morris R. G., von der Linden A., Applegate D., 2008, ApJ, 687, 959
Buckley M. R., Fox P. J., 2010, Phys. Rev. D, 81, 083522
Carlson E. D., Machacek M. E., Hall L. J., 1992, ApJ, 398, 43
Correa C. A., Wyithe J. S. B., Schaye J., Duffy A. R., 2015, MNRAS, 452, 1217
Davé R., Spergel D. N., Steinhardt P. J., Wandelt B. D., 2001, ApJ, 547, 574
Dawson W. A. et al., 2012, ApJ, 747, L42
Dehnen W., Read J. I., 2011, Eur. Phys. J. Plus, 126, 55
Diemer B., Kravtsov A. V., 2015, ApJ, 799, 108
Duncan C. A. J., Heymans C., Heavens A. F., Joachimi B., 2016, MNRAS, 457, 764
Dutton A. A., Macciò A. V., 2014, MNRAS, 441, 3359
Farrar G. R., Rosen R. A., 2007, Phys. Rev. Lett., 98, 171302
Feng J. L., Kaplinghat M., Tu H., Yu H.-B., 2009, J. Cosmol. Astropart. Phys., 7, 004
Foreman-Mackey D., 2016, J. Open Source Softw., 24. Available at: http://dx.doi.org/10.5281/zenodo.45906
Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306
Fry A. B. et al., 2015, MNRAS, 452, 1468
George M. R. et al., 2012, ApJ, 757, 2
Goodman J., Weare J., 2010, Commun. Appl. Math. Comput. Sci., 5, 65
Governato F. et al., 2012, MNRAS, 422, 1231
Harvey D., Massey R., Kitching T., Taylor A., Tittley E., 2015, Science, 347, 1462
Hayashi E., White S. D. M., 2006, MNRAS, 370, L38
Hernquist L., 1990, ApJ, 356, 359
Kahlhoefer F., Schmidt-Hoberg K., Frandsen M. T., Sarkar S., 2014, MNRAS, 437, 2865 (K14)
Kahlhoefer F., Schmidt-Hoberg K., Kummer J., Sarkar S., 2015, MNRAS, 452, L54
Kaplinghat M., Tulin S., Yu H.-B., 2014, Phys. Rev. D, 89, 035009
Kaplinghat M., Tulin S., Yu H.-B., 2016, Phys. Rev. Lett., 116, 041302
Kassiola A., Kovner I., 1993, ApJ, 417, 450
King L. J. et al., 2016, MNRAS, 459, 517
Kochanek C. S., White M., 2000, ApJ, 543, 514
Koda J., Shapiro P. R., 2011, MNRAS, 415, 1125
Kusenko A., Steinhardt P. J., 2001, Phys. Rev. Lett., 87, 141301
Lage C., Farrar G., 2014, ApJ, 787, 144
Leauthaud A. et al., 2007, ApJS, 172, 219
Loeb A., Weiner N., 2011, Phys. Rev. Lett., 106, 171302
Markevitch M., 2006, in Wilson A., ed., ESA SP-604: The X-ray Universe 2005. ESA, Noordwijk, p. 723
Markevitch M., Gonzalez A. H., Clowe D., Vikhlinin A., Forman W., Jones C., Murray S., Tucker W., 2004, ApJ, 606, 819
Massey R. et al., 2015, MNRAS, 449, 3393
Mastropietro C., Burkert A., 2008, MNRAS, 389, 967
Meneghetti M., Yoshida N., Bartelmann M., Moscardini L., Springel V., Tormen G., White S. D. M., 2001, MNRAS, 325, 435
Merten J. et al., 2011, MNRAS, 417, 333
Merten J. et al., 2015, ApJ, 806, 4
Milosavljević M., Koda J., Nagai D., Nakar E., Shapiro P. R., 2007, ApJ, 661, L131
Mohammed I., Liesenborgs J., Saha P., Williams L. L. R., 2014, MNRAS, 439, 2651
Mohapatra R. N., Nussinov S., Teplitz V. L., 2002, Phys. Rev. D, 66, 063002
Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493 (NFW)
Pontzen A., Governato F., 2012, MNRAS, 421, 3464
Popesso P., Biviano A., Böhringer H., Romaniello M., 2007, A&A, 464, 451
Power C., Navarro J. F., Jenkins A., Frenk C. S., White S. D. M., Springel V., Stadel J., Quinn T., 2003, MNRAS, 338, 14
Prada F., Klypin A. A., Cuesta A. J., Betancort-Rijo J. E., Primack J., 2012, MNRAS, 423, 3018
Ragozzine B., Clowe D., Markevitch M., Gonzalez A. H., Bradač M., 2012, ApJ, 744, 94
Randall S. W., Markevitch M., Clowe D., Gonzalez A. H., Bradač M., 2008, ApJ, 679, 1173 (R08)
Richard J., Kneib J.-P., Limousin M., Edge A., Jullo E., 2010, MNRAS, 402, L44
Robertson A., Massey R., Eke V., Bower R., 2015, MNRAS, 453, 2267
Rocha M., Peter A. H. G., Bullock J. S., Kaplinghat M., Garrison-Kimmel S., Oñorbe J., Moustakas L. A., 2013, MNRAS, 430, 81
Sawala T. et al., 2016, MNRAS, 457, 1931
Schaller M., Robertson A., Massey R., Bower R. G., Eke V. R., 2015, MNRAS, 453, L58
Shu Y., Bolton A. S., Moustakas L. A., Stern D., Dey A., Brownstein J. R., Burles S., Spinrad H., 2016, ApJ, 820, 43
Smith G. P., Kneib J.-P., Smail I., Mazzotta P., Ebeling H., Czoske O., 2005, MNRAS, 359, 417
Spergel D. N., Steinhardt P. J., 2000, Phys. Rev. Lett., 84, 3760

Springel V., 2005, MNRAS, 364, 1105
Springel V., Farrar G. R., 2007, MNRAS, 380, 911 (SF07)
Springel V., Di Matteo T., Hernquist L., 2005, MNRAS, 361, 776
Tulin S., Yu H.-B., Zurek K. M., 2013a, Phys. Rev. D, 87, 115007
Tulin S., Yu H.-B., Zurek K. M., 2013b, Phys. Rev. Lett., 110, 111301
van den Aarssen L. G., Bringmann T., Pfrommer C., 2012, Phys. Rev. Lett., 109, 231301
Vogelsberger M., Zavala J., Loeb A., 2012, MNRAS, 423, 3740
Wang W., Zhang M., Zhao J., 2016, preprint (arXiv:1604.00123)
Weinberg D. H., Bullock J. S., Governato F., Kuzio de Naray R., Peter A. H. G., 2015, Proc. Natl. Acad. Sci., 112, 12249
Williams L. L. R., Saha P., 2011, MNRAS, 415, 448
Yoshida N., Springel V., White S. D. M., Tormen G., 2000, ApJ, 544, L87

APPENDIX A: TESTS OF SCATTERING KINEMATICS

A1 Scattering rates

To test our scattering algorithm, we modelled a uniform cube of particles moving through a constant density background of stationary particles. To allow for simple predictions to be made for the system, we did not allow particles to scatter more than once. All of the particles in the cube had initial velocities v_0 along the z-axis, and gravity was turned off.

The average rate of scattering is $\Gamma = N_c\, n_b\, v_0\, \sigma_p$, where N_c is the number of particles in the cube, and $n_b = (\rho_b/m_p)$ is the number density of background particles. This leads to the expected number of interactions after a time t:

$$N_{exp} = N_c\, \rho_b\, v_0\, (\sigma_p/m_p)\, t. \tag{A1}$$

The number of scattering events in these test simulations is plotted in Fig. A1 as a function of the search radius, h. For small values of h, the number of scattering events falls below that expected. This was noted by Rocha et al. (2013), who found that scattering was not correctly resolved for h less than 20 per cent of the mean background interparticle separation. By running the test case with different time-steps, we find that this 20 per cent is not an intrinsic property of simulating scattering using a Monte Carlo method. Instead we find that the minimum h for which scattering is correctly implemented is a function of the time-step, scattering cross-section, and the relative velocity of particles.

In general, the scattering rate is insensitive to h, as the number of neighbouring particles that a particle finds at each time-step is proportional to the volume searched ($\propto h^3$), but the probability of scattering from each of those particles follows equation (6) ($\propto h^{-3}$). The product of the number of neighbouring particles, and the probability of scattering with each of them, gives the total probability of a particle scattering, which does not depend on h. This breaks down when the probability to scatter from a neighbouring particle becomes greater than unity. At that point, the probability of a particle scattering in a time-step is just the probability of finding a neighbouring particle during that time-step, which goes as h^3. For this reason, in this 'probability-saturated' regime, the rate of scattering is proportional to h^3, as shown by the solid lines in the right-hand panel of Fig. A1. As the probability for a pair of neighbouring particles to scatter is proportional to $\Delta t/h^3$, a smaller h can be used when using shorter time-steps.

A2 Choosing h_{si}

From equation (6), we see that probabilities become greater than unity when

$$h < \left(\frac{3}{4\pi}\sigma_p\, |\boldsymbol{v}_i - \boldsymbol{v}_j|\, \Delta t\right)^{1/3}. \tag{A2}$$

The time-step criterion used for the DM particles in GADGET is

$$\Delta t = \min\left[\Delta t_{max}, \left(\frac{2\eta\epsilon}{a}\right)^{1/2}\right], \tag{A3}$$

where a is a particle's acceleration, ϵ the gravitational softening, and η a dimensionless constant.

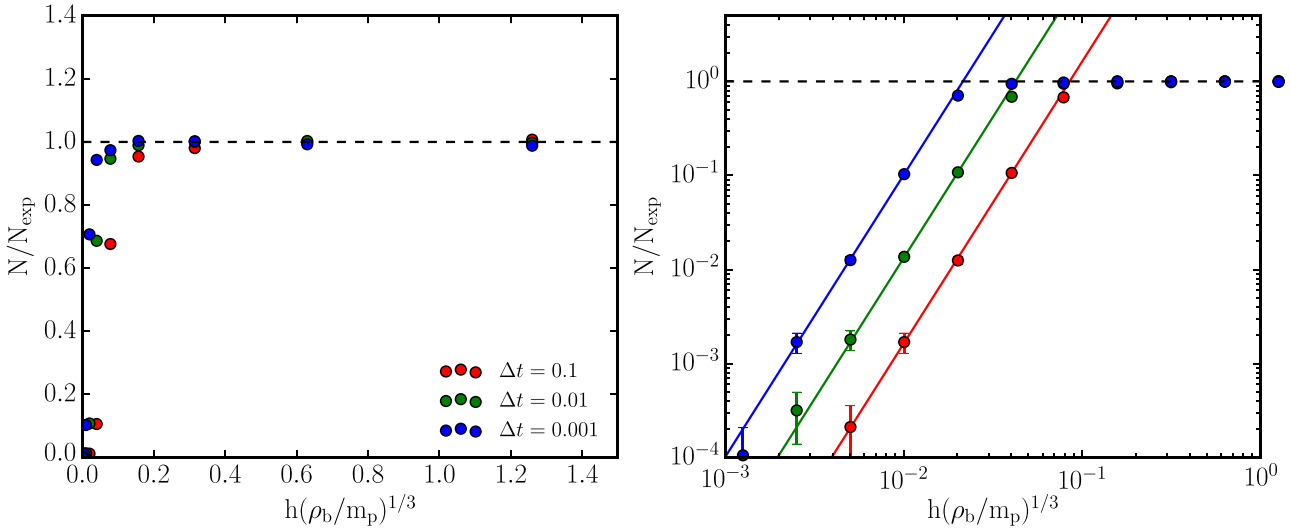

Figure A1. The number of scattering events in our test simulations as a function of neighbour-search radius, h. The left-hand panel is similar to fig. 1 in Rocha et al. (2013), and we find that we also see a decrease in the rate of scattering, below that expected, when using small h. While this happens for $h \lesssim 0.2(m_p/\rho_b)^{1/3}$ in agreement with the results in Rocha et al. (2013), the precise h at which the drop in scattering rate begins is a function of the simulation time-step, Δt. As discussed in the text, this is a result of the probabilities for pairs of particles to scatter within a time-step becoming greater than 1. These probabilities are $\propto h^{-3}$, and so in the right-hand panel we show the same data as in the left-hand panel, but plotted on logarithmic scales. The solid lines show $N \propto h^3$, the result one expects from probability saturation. For these test simulations, $N_{exp} \approx 10^4$, and the error bars show the 1σ uncertainty, assuming that N is Poisson distributed.

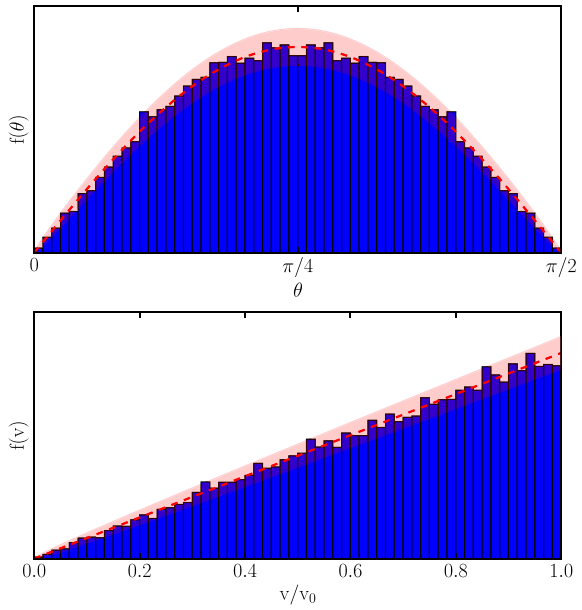

Figure A2. The distribution of polar angles and velocity magnitudes of scattered particles in one of our test simulations. The red dashed line shows the expected distribution, with the red shaded region showing the 2σ variation about this expectation, assuming the number of particles in each bin is Poisson distributed.

As the time-steps are dependent on the dynamics of the system being simulated, the constraint on h from equation (A2) depends on the content of the simulation. We found that when using the standard time-step criterion, the probabilities remained below 1 when h was set equal to the gravitational softening, ϵ. For example, the simulation with our fiducial initial conditions and $\sigma/m = 1\,\mathrm{cm}^2\,\mathrm{g}^{-1}$, had a maximum probability for a pair of particles to scatter within a time-step of 0.15, with 99 per cent of scattering events taking place with a probability <0.03. If running simulations with very large cross-sections, an additional time-step criterion could be added to prevent probabilities exceeding unity.

A3 Post-scatter kinematics

As well as the rate of scattering, the directions and velocities of scattered particles in our test case were compared to expectations. The expected distribution of velocities and directions is calculated by transforming the differential cross-section from the centre of mass frame of the collisions, into the frame of our simulations. For the case of isotropic scattering, these distributions take on simple forms, with $f(\theta) \propto \sin 2\theta$, and $f(v) \propto v$ for $v \leq v_0$, with no particles with velocities greater than v_0.[6] These results are shown in Fig. A2, with results that match expectations.

APPENDIX B: THE CONTRIBUTION OF MASS AT DIFFERENT RADII TO THE PROJECTED DENSITY AND SHEAR SIGNALS

In Fig. 14, we see that with an asymmetric DM profile, the positions of the halo returned from fitting to the projected density and reduced gravitational shear differ, being similar to the shrinking-circles positions of the halo when shrinking to $R_{\mathrm{min}} = 200$ and 60 kpc, respectively. This implies that shear is more sensitive to the central regions of the halo, which appears to be at odds with the maps in Fig. 13 showing that the projected density (and equivalently convergence) increases towards the centre of the halo, while the shear has a flatter profile. This can be explained by noting that shear is a non-local quantity, and that for a circularly symmetric projected mass distribution the shear at radius R depends on all of the mass within R. In fact, the tangential shear (γ_{t}) from a circularly symmetric mass distribution can be written in terms of the 'excess surface density'

$$\Delta\Sigma = \bar{\Sigma}(<R) - \Sigma(R) = \Sigma_{\mathrm{crit}}\gamma_{\mathrm{t}}. \tag{B1}$$

For an annulus of mass at radius R_0 with mass M_0 the shear internal to R_0 is zero, while the shear at $R > R_0$ is simply the average surface density within R divided by the critical surface density. The average surface density is the enclosed mass divided by the area, so

$$\gamma_{\mathrm{t}} = \frac{1}{\Sigma_{\mathrm{crit}}}\frac{M_0}{\pi R^2}. \tag{B2}$$

The noise in the shear map is independent of position, so the signal-to-noise ratio in a particular pixel is just proportional to the shear there. The number of pixels in an annulus at radius R is proportional to $2\pi R\,\mathrm{d}R$. This implies that the sum of the signal to noise over all pixels in an annulus at R due to the mass M_0 at R_0 is proportional to $(M_0/R^2)R\,\mathrm{d}R$. Integrating this from $R = R_0$ outwards, we find that the sum of the signal to noise over all pixels in the map is proportional to $M_0 \ln(R_{\mathrm{max}}/R_0)$, where we have truncated the integration at a maximum radius R_{max}. For R_{max}, we use half the side length of the square regions used when fitting to shear. As the total signal to noise only grows logarithmically with R_{max}, this choice is not particularly important.

The mass M_0, which is the mass in an annulus at radius R_0, is the surface density at radius R_0 multiplied by the area of the annulus, so $M_0 \propto \Sigma(R_0)R_0$. As such, the sum of the signal-to-noise ratio of all pixels in the map due to mass at radius R_0 is

$$\mathrm{SNR}_g \propto R_0\Sigma(R_0)\ln\left(\frac{R_{\mathrm{max}}}{R_0}\right). \tag{B3}$$

The projected density is a local quantity, leading to the calculation being simpler than for the case of shear. Fitting to the projected density used Poisson statistics, which for large numbers of particles per bin can be approximated by Gaussian statistics. The signal-to-noise ratio of a single pixel is then $\sqrt{N} \propto \sqrt{\Sigma}$, where N is the number of particles in that pixel. The mass at radius R_0 only contributes to the signal at R_0, and the number of pixels at radius R_0 is proportional to R_0. Using this, the sum of the signal-to-noise ratio of all pixels in the map due to mass at radius R_0 is

$$\mathrm{SNR}_\Sigma \propto R_0\sqrt{\Sigma(R_0)}. \tag{B4}$$

SNR_g and SNR_Σ are the quantities plotted in the bottom panel of Fig. 14, where they have been normalized by their maximum value.

[6] For isotropic scattering, the distribution of scattered particles is the same for those originally part of the background or originally part of the moving cube.

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