The Cosmological Implications of Self-Interacting Dark Matter

ROBERTSON, ANDREW

The Cosmological Implications of Self-Interacting Dark Matter

Andrew Alexander Robertson

Abstract

In this thesis I study how dark matter particles that interact through forces other than just gravity would affect the formation of structure in the Universe. This begins with a theoretical calculation of the location and rate at which these interactions take place throughout cosmic history. Giant galaxy clusters are expected to have the highest rates of dark matter interactions, at least for the simplest dark matter particle models.

Predicting the formation of structure with non-standard dark matter requires the use of N-body simulations. I therefore introduce and test a set of modifications to the GADGET code that allow it to simulate a class of dark matter models known as self-interacting dark matter (SIDM). I focus particular attention on rarely discussed aspects of simulating SIDM; including how to handle particles scattering multiple times within a single time-step and how to implement scattering across processors. I also discuss how best to choose numerical parameters associated with the SIDM implementation and the range of numerical parameters that produce converged results.

Because galaxy clusters should have particularly high rates of dark matter interactions, I use this code to perform simulations of a pair of merging galaxy clusters known as the 'Bullet Cluster'. At first these employ simple SIDM particle physics models for the dark matter. I demonstrate the importance of analysing simulations in an observationally motivated manner, finding that the way in which simulation outputs are compared with observations can have a significant impact on the derived constraints upon dark matter's properties. I then look at what happens to these constraints for more complicated particle physics models of SIDM. In isolated systems, the effects of a complicated scattering cross-sections can be modelled using an appropriately-matched simple cross-section, while in systems like the Bullet Cluster, complicated cross-sections lead to phenomenology not seen with simpler particle models. Overall I find that SIDM remains a viable class of dark matter models, consistent with current observations.

The Cosmological Implications of Self-Interacting Dark Matter

Andrew Alexander Robertson

A thesis presented in accordance with
the regulations for admittance to the degree
of Doctor of Philosophy

Institute for Computational Cosmology
Department of Physics
University of Durham
April 2017

Contents

List of Figures

Declaration

The work described in this thesis was undertaken between October 2013 and April 2017 while the author was a research student under the supervision of Dr Richard Massey and Dr Vincent Eke in the Department of Physics at the University of Durham. This work has not been submitted for any other degree at the University of Durham or any other University.

Chapters 2, 5 and 6 have all been published as papers in Monthly Notices of the Royal Astronomical Society (MNRAS):

- *Self-interacting dark matter scattering rates through cosmic time*
  Andrew Robertson, Richard Massey, Vincent Eke and Richard Bower
  MNRAS 2015, Volume 453, Issue 3, pp. 2267-2276,
  ARXIV:1505.02046

- *What does the Bullet Cluster tell us about self-interacting dark matter?*
  Andrew Robertson, Richard Massey and Vincent Eke
  MNRAS 2017, Volume 465, Issue 1, pp. 569-587,
  ARXIV:1605.04307

- *Cosmic particle colliders: simulations of self-interacting dark matter with anisotropic scattering*
  Andrew Robertson, Richard Massey and Vincent Eke
  MNRAS 2017, Volume 467, Issue 4, pp. 4719-4730,
  ARXIV:1612.03906

All figures in this thesis have been produced by the author, or have been properly attributed in the figure caption.

Acknowledgements

Firstly, I would like to thank my supervisors, Richard Massey and Vincent Eke, without whom this work would certainly not have been possible. Richard's enthusiasm and knowledge of observations, combined with Vince's levelheadedness and detailed understanding of simulations, was a perfect combination, and I cannot imagine better supervisors. I particularly want to thank them for being so generous with their time, and also with their words; on the second front I am sure that it was not always merited. Thanks also to Richard Bower for significant input in my first year, and to Lydia Heck for all her help with computing.

I would like to thank some of my teachers from school and university. Linda Fleming was particularly influential in igniting my early love of mathematics, later cultivated by Nevil Hopley and many of the other Watson's maths teachers. Both Keith Grainge and Rosie Bolton played crucial roles in convincing me that I would enjoy doing a PhD, especially one in astrophysics. Thankfully it turned out that they were both correct!

I am thankful for my friends, both from physics as well as the wider University. Richard, Sam, Katie, Bennett, Natasha, Lauren, Rhi, Hans, John, Abi, Eleonore and all the other Hatfielders – I certainly should not be thanking you for helping me to get work done, but you all made sure that I had a fun time in Durham. Tim, Helen and Paddy, I would like to thank you not just for being great housemates, but also great friends. Gabriel, Tim and Clara, you were three of a kind, and I sorely miss our 'formal team' evenings. I have also been fortunate to share an office with many interesting people; Hannah, Ton, Tom, Peter, Oliver, Alex and Stefan. There are many more people who I was never lucky enough to share an office with, but whose company I always enjoyed, with Sownak, James, Saavi and Matthieu deserving a special mention.

I would of course like to thank my family, who have always supported my interest in mathematics and physics. They somehow never tired of hearing me tell them about dark matter, or if they did, were much too kind to tell me. Finally, Hannah, I am truly grateful to you for always being there. I am sorry for having dragged you away from your natural habitat to the frozen north, but thoroughly enjoyed our time living together in Durham.

This thesis was supported by the Science and Technology Facilities Council grant ST/K501979/1.

Chapter 1

Introduction

Few questions can have as rich a history of human enquiry as the origin, evolution, and eventual fate of the Universe. The understanding of the laws that govern our Universe, as well as the ingredients that those laws act upon, is a branch of science now known as *physical cosmology*.

1.1 Cosmology

The framework within which we study cosmology usually rests upon two key assumptions. The first is that Einstein's General Theory of Relativity (GR) is the correct description of gravity, and the second is something known as the cosmological principle.[1] This principle states that on large enough scales the Universe appears to be homogeneous and isotropic. That this is not true on small scales is evident from the rocky planet on which we live, orbiting a star about $8\,\mathrm{kpc}$ away from the centre of a spiral galaxy; but if we look at spherical volumes with a radius of $100\,\mathrm{Mpc}$, then the typical density within these spheres only fluctuates by around 1% from the cosmic average (Wu et al., 1999).

GR allows us to determine what the Universe looked like in the past, or will look like in the future, provided we can measure the state of it now. Our first real insight into what this present state looks like came from Edwin Hubble's observation that galaxies were moving away from us, with velocities that were, on average, proportional to their distance from us

[1] Also known as the Copernican principle after Nicolaus Copernicus

(Hubble, 1929). This observation,

$$\mathbf{v} = H_0 \, \mathbf{r}, \tag{1.1}$$

known as Hubble's law, may appear to put us in a special place in the Universe, violating the cosmological principle. This is not the case, because the distance-velocity relation in equation (1.1) implies that all other observers would observe the same law. This is only true for a linear relationship between position and velocity; if, for example, we had observed $v \propto r^2$, then we would be at a unique position in the 'centre' of the Universe, and other observers would measure anisotropic distance–velocity relations. The Hubble law therefore fits in naturally with our cosmological principle.

1.1.1 Hot Big Bang Model

The constant H_0 is known as the Hubble constant, and describes the rate of expansion of the Universe at the present time ($t = t_0$). In general the constant in the distance–velocity relationship is a function of time, $H(t)$, which we can calculate from its current value and the contents of the Universe, using solutions to the equations of GR for an isotropic and homogeneous universe. Working backwards, the picture we get is of a Universe that had an infinite temperature and density at a finite time in the past. GR is not able to adequately describe physics in this regime, but can take it on from a short while later, where it predicts a hot and dense Universe that is expanding and cooling. This model for the early Universe is known as the *Hot Big Bang*, and it has successfully explained many cosmological observations.

As well as explaining the current expansion of the Universe, a Hot Bing Bang naturally explains the origin of the Cosmic Microwave Background (CMB), as well as the nucleosynthesis of the light elements. The former relies on the fact that when the Universe was sufficiently hot that hydrogen atoms were unstable, radiation was efficiently scattered by free electrons, and the Universe was opaque to photons. Once the temperature had dropped enough to allow electrons and protons to form atomic hydrogen, the cross-section for photon scattering dropped, and the Universe became transparent to radiation. Rather oddly,

this point on the cosmic timeline is known as *recombination*,[2] which takes place around 400,000 years after the Big Bang (Planck Collaboration et al., 2016). The radiation from this time has propagated freely since then, shifting in energy as its wavelength was stretched by the expansion of the Universe. This radiation, sometimes dubbed 'the afterglow of the big bang' was first detected as microwaves with a temperature $\sim 3\,\mathrm{K}$ by Penzias & Wilson (1965), providing strong evidence for the Hot Big Bang model.

The nucleosynthesis of the light elements can also be predicted within the Hot Big Bang paradigm. In the very early Universe, weak-force interactions keep protons and neutrons in thermal equilibrium. When the Universe was just one second old, the density and energies of particles had dropped such that the weak force was no longer able to maintain this equilibrium. This is a process known as freeze out, and led to a Universe with a proton to neutron ratio of about 6:1, in favour of protons because they are slightly lighter than neutrons. The build-up of nuclei from protons and neutrons can then be calculated building upon the method outlined in Alpher, Bethe & Gamow (1948). The rates of production and destruction of the different light elements depend in different ways on the density of baryons, so the fact that a single value for the baryon density can adequately describe the observed primordial abundances of the light elements (Boesgaard & Steigman, 1985) provides further strong evidence for a Hot Big Bang. These calculations rely on our ability to model the expansion history of the Universe, so we will show how to do this now.

1.1.2 The expansion history of the Universe

As discussed above, the equations that describe the evolution of the Universe are those of Einstein's GR. However, it turns out that a Newtonian approach can be used to derive equations that are almost identical to those in GR, and so we do that here. We begin by considering a sphere of mass M and radius R, with the surface of the sphere expanding at a rate $\dot{R}$ and the mass within the sphere expanding such that at all times, the density is constant as a function of radius. The edge of the sphere is decelerated by all of the mass within the sphere, while any symmetric mass distribution outside of the sphere produces

[2] This is odd because protons and electrons had not combined at any previous epoch.

no net acceleration due to Newton's shell theorem. With Newtonian gravity, this implies

$$\ddot{R} = -\frac{GM}{R^2} = -\frac{4\pi G\rho(t)R^3}{3R^2}, \tag{1.2}$$

where $\rho(t)$ is the mass density inside the sphere, and dots correspond to derivatives with respect to time. Conserving mass, the density scales as $1/R^3$, which we can write as

$$\rho = \rho_0 \left(\frac{R_0}{R}\right)^3, \tag{1.3}$$

with the density equal to ρ_0 at the time when $R = R_0$.

Multiplying equation (1.2) by $2\dot{R}$ and integrating with respect to time

$$\dot{R}^2 = \frac{8\pi G\rho_0 R_0^3}{3R} + \text{constant} \tag{1.4}$$

$$\left(\frac{\dot{R}}{R}\right)^2 = \frac{8\pi G\rho(t)}{3} + \frac{\text{constant}}{R^2}. \tag{1.5}$$

When considering the expansion of the Universe, it is useful to replace the radius of the sphere with the scale-factor, $a(t)$. This can be related to R through $R(t)/R(t_0) = a(t)$, which defines the scale factor as unity at the current time, t_0. When doing the full GR calculation, an additional constant term appears in equation (1.5), known as the cosmological constant, Λ. We can therefore write the first Friedmann equation, which governs the expansion history of the Universe, as

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G\rho(t)}{3} - \frac{kc^2}{a^2} + \frac{\Lambda}{3}, \tag{1.6}$$

where we can identify $\dot{a}/a = \dot{R}/R$ with the Hubble parameter introduced in equation (1.1), so $\dot{a}/a = H(t)$. The constant, k, can be positive, negative or zero, and relates to the curvature of the Universe.

In order to solve equation (1.6), we need to know how the density, ρ, varies with time (or scale factor). The matter density of the Universe evolves $\propto 1/a^3$, as mass is conserved and the volume of the Universe $\propto a^3$. If we have a matter-only universe, with zero cosmological constant, then using $\rho = \rho_0/a^3$ equation (1.6) becomes

$$\dot{a}^2 = \frac{8\pi G\rho_0}{3a} - kc^2. \tag{1.7}$$

For a universe that is expanding, the first term on the right hand side of equation (1.7) decreases in magnitude. If k is positive, then at finite a, the right hand side goes to zero,

and the expansion of the Universe halts. The Universe would then proceed to collapse. If k is negative, then $\dot{a}$ decreases with time, but eventually tends to a constant value, while if $k = 0$ the expansion of the Universe continues to slow, but does not halt in a finite time.

The density in equation (1.6), is not just the density of matter, but the energy-density in all forms. For the case of radiation, the number density of photons is $\propto 1/a^3$, but as the photon wavelength increases $\propto a$, and the energy of a photon is inversely proportional to its wavelength, the energy density in radiation $\propto 1/a^4$. The curvature term and cosmological constant can also be related to an energy density, such that equation (1.6) becomes

$$H^2 = \frac{8\pi G}{3}\left(\rho_{\mathrm{m},0}/a^3 + \rho_{\mathrm{r},0}/a^4 + \rho_{\mathrm{k},0}/a^2 + \rho_{\Lambda,0}\right), \tag{1.8}$$

with $\rho_{\mathrm{x},0}$ the energy density in species x at $t = t_0$ (the present day), with m, r, k, and Λ corresponding to matter, radiation, curvature and a cosmological constant respectively.

Looking at equation (1.8) at $t = t_0$ (when $a = 1$), we can see that the sum of all the $\rho_{\mathrm{x},0}$ is equal to $3H_0^2/8\pi G$, which we call the critical density, $\rho_{\mathrm{crit},0}$. Defining $\Omega_{\mathrm{x},0} = \rho_{\mathrm{x},0}/\rho_{\mathrm{crit},0}$, we can write

$$H^2 = H_0^2\left(\Omega_{\mathrm{m},0}/a^3 + \Omega_{\mathrm{r},0}/a^4 + \Omega_{\mathrm{k},0}/a^2 + \Omega_{\Lambda,0}\right), \tag{1.9}$$

with $\Omega_{\mathrm{k},0} = 1-(\Omega_{\mathrm{m},0}+\Omega_{\mathrm{r},0}+\Omega_{\Lambda,0})$. Note that the subscript 0 here is often dropped, so that (for example) Ω_{m} is used to designate the density of matter relative to the critical density at t_0. We could however ask about the value of $\Omega_{\mathrm{m}}(t)$, which need not be constant as the density of matter can evolve in a different way to the critical density. A large amount of current work in cosmology is focussed upon determining the values of the $\Omega_{\mathrm{x},0}$s. The latest results from the Planck satellite's measurement of the CMB (Planck Collaboration et al., 2016) find a matter density $\Omega_{\mathrm{m},0} = 0.308 \pm 0.012$, and limits on the curvature of $|\Omega_{\mathrm{k},0}| < 0.005$. The radiation density is much smaller, $\Omega_{\mathrm{r},0} \approx 9.1 \times 10^{-5}$, and assuming the Universe is spatially flat[3] ($\Omega_{\mathrm{k},0} = 0$) then $\Omega_{\Lambda,0} = 0.692 \pm 0.012$.

It is useful to define the function

$$E(a) = \sqrt{\Omega_{\mathrm{m},0}/a^3 + \Omega_{\mathrm{r},0}/a^4 + \Omega_{\mathrm{k},0}/a^2 + \Omega_{\Lambda,0}}\,, \tag{1.10}$$

such that $H(a) = H_0 E(a)$. We can then solve for the evolution of the scale factor, $a(t)$,

[3]We will discuss in §1.2.3 the inflationary paradigm, which provides strong motivation for $\Omega_{\mathrm{k},0} \approx 0$.

starting from $a = 0$ at $t = 0$ through

$$t = \int_0^t \mathrm{d}t' = \int_0^a \frac{\mathrm{d}a'}{\dot{a}'} = \int_0^a \frac{\mathrm{d}a'}{a'H(a')} = \frac{1}{H_0}\int_0^a \frac{\mathrm{d}a'}{a'E(a')}. \tag{1.11}$$

For general $E(a)$ this can be solved numerically, but for simple universes we can proceed analytically. For example, a matter-dominated critical density universe ($\Omega_{\mathrm{m},0} = 1$, $\Omega_{\mathrm{r},0} = \Omega_{\mathrm{k},0} = \Omega_{\Lambda,0} = 0$), has $E(a) = a^{-3/2}$. Putting this into equation (1.11) we find $a = (t/t_0)^{2/3}$, with $t_0 = \frac{2}{3H_0}$ the age of this universe when $a = 1$. A similar calculation for a radiation-dominated critical density universe gives $a = (t/t_0)^{1/2}$, $t_0 = \frac{1}{2H_0}$; while for a universe containing only a cosmological constant $a \propto \exp(Ht) = \exp\left(\sqrt{\Lambda/3}\,t\right)$. This final case is known as a De Sitter universe, in which the Hubble law is not just a constant in space, but in time as well. The time in this case is not particularly well defined, as this Universe lacks a big-bang singularity ($a = 0$). A pure De Sitter universe is uninteresting because it contains no matter, however, in our current best model for the expansion history of our Universe, there are two phases in which the expansion is approximately De Sitter. The first is inflation, during which a large vacuum energy drives an exponential expansion of the early Universe. The second is at late times, as dark energy (which is currently consistent with a cosmological constant) becomes the dominant energy component in the Universe. We will discuss both of these later in this chapter.

Cosmological redshift

While we are discussing the expansion history of the Universe, it is useful to consider the observational consequences of this expansion. In particular, the stretching of the wavelength of light (or gravitational waves for that matter) that we call 'redshift'. If two photons are emitted at times t_e and $t_\mathrm{e} + \Delta t_\mathrm{e}$, and arrive at an observer at times t_o and $t_\mathrm{o} + \Delta t_\mathrm{o}$, then for $\Delta t_\mathrm{e} \ll t_\mathrm{o} - t_\mathrm{e}$ it can be shown that $\Delta t_\mathrm{o}/\Delta t_\mathrm{e} = a(t_\mathrm{o})/a(t_\mathrm{e})$. If we take $1/\Delta t_\mathrm{e}$ to be the frequency of a photon when it is emitted, then the wavelength when the photon is emitted is $\lambda_\mathrm{e} = c\,\Delta t_\mathrm{e}$. The wavelength of the photon when it is received is $\lambda_\mathrm{o} = c\,\Delta t_\mathrm{o}$. We define the redshift as

$$z = \frac{\lambda_\mathrm{o} - \lambda_\mathrm{e}}{\lambda_\mathrm{e}}, \tag{1.12}$$

i.e. the fractional change in a photon's wavelength from when it was emitted to when it was received. With this definition, $z = (\Delta t_{\rm o} - \Delta t_{\rm e})/\Delta t_{\rm e} = a(t_{\rm o})/a(t_{\rm e}) - 1$. Therefore, if we receive a photon now (when $t = t_0$ and $a = 1$), the scale factor when it was emitted was

$$a(t_{\rm e}) = \frac{1}{1+z}. \tag{1.13}$$

How can we measure z? If the emission results from some physical process for which we know the wavelength (such as an absorption line in a star's spectrum), then we know $\lambda_{\rm e}$. So the wavelength of this emission in a measured spectrum, $\lambda_{\rm o}$, allows us to calculate z.

1.1.3 Structure formation

So far our calculations have assumed the Universe to be homogenous and isotropic. Given that on sufficiently large scales this appears to be true, this approach is thought to be valid for calculating the expansion history of the Universe.[4] However, we know that on small scales there is structure, and in this section we discuss where this comes from. We start with the equations that govern the evolution of the density $\rho(\mathbf{r}, t)$ and velocity $\mathbf{v}(\mathbf{r}, t)$ of an ideal fluid:

$$\text{Continuity} \quad : \quad \frac{\partial \rho}{\partial t} + \nabla.(\rho \mathbf{v}) = 0 \tag{1.14}$$

$$\text{Euler} \quad : \quad \frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v}.\nabla)\mathbf{v} = -\frac{1}{\rho}\nabla p - \nabla\Phi \tag{1.15}$$

$$\text{Poisson} \quad : \quad \nabla^2\Phi = 4\pi G\rho. \tag{1.16}$$

The first two equations describe the conservation of mass and momentum, with the third describing how matter sources the gravitational field.

These equations are solved by the homogeneous solution $\rho(\mathbf{r}, t) = \bar{\rho}(t)$, $\mathbf{v} = H\mathbf{r} \equiv \mathbf{v}_0$ and $\Phi = \frac{2\pi G\bar{\rho}}{3}|\mathbf{r}|^2 \equiv \Phi_0$. The continuity equation then reads

$$\frac{\partial \bar{\rho}}{\partial t} + 3H\bar{\rho} = 0, \tag{1.17}$$

with solution $\bar{\rho} = \rho_0/a^3$. The Euler equation reads

$$\dot{H}\mathbf{r} + H^2\mathbf{r} = -\frac{4\pi G\bar{\rho}}{3}\mathbf{r}, \tag{1.18}$$

[4] Though some authors claim (see for example Buchert & Räsänen, 2012) that small scale inhomogeneities can affect the global expansion history through a process known as 'backreaction'.

which implies $a \propto t^{2/3}$. We therefore see that this homogeneous solution is the matter-dominated critical density universe discussed in §1.1.2. This is to be expected given that our starting equations were Newtonian, such that they are valid when space is flat (a critical density universe) and the fluid flow is non-relativistic (matter, not radiation). This second condition is met for the vast majority of the Universe's history, since the density of matter became greater than that of radiation approximately $50\,000$ years after the big bang.

In equations (1.14)–(1.16), the time derivatives are at fixed physical position, $\mathbf{r}$, with $\nabla \equiv \frac{\partial}{\partial \mathbf{r}}$. Due to the expansion of the Universe, these equations are more easily understood in comoving coordinates, $\mathbf{x}$. These are defined by

$$\mathbf{r} = a(t)\mathbf{x}. \tag{1.19}$$

The velocity can then be split into the velocity coming from the expansion (at fixed comoving position) and the peculiar velocity (corresponding to a changing comoving position)

$$\mathbf{v} = \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} = \dot{a}\mathbf{x} + a\dot{\mathbf{x}} \equiv \mathbf{v}_0 + \mathbf{v}_\mathrm{p}, \tag{1.20}$$

with $\mathbf{v}_\mathrm{p}$ the peculiar velocity.

To transform the Continuity, Euler and Poisson equations into comoving coordinates we note that

$$\frac{\partial}{\partial \mathbf{r}} = \frac{1}{a}\frac{\partial}{\partial \mathbf{x}} \equiv \frac{1}{a}\nabla_\mathbf{x} \tag{1.21}$$

and that time derivates at fixed $\mathbf{r}$ transform to time derivates at fixed $\mathbf{x}$ following

$$\left(\frac{\partial}{\partial t}\right)_\mathbf{r} = \left(\frac{\partial}{\partial t}\right)_\mathbf{x} - H\mathbf{x}.\nabla_\mathbf{x}. \tag{1.22}$$

It is also useful to define the overdensity field $\delta(\mathbf{x}, t)$ by

$$\rho(\mathbf{x}, t) = \bar{\rho}(t)\left[1 + \delta(\mathbf{x}, t)\right]. \tag{1.23}$$

Making these substitutions we arrive at the comoving equations:

$$\text{Continuity} \quad : \quad \left(\frac{\partial \delta}{\partial t}\right)_\mathbf{x} + \frac{1}{a}\nabla_\mathbf{x}.\left[(1+\delta)\mathbf{v}_\mathrm{p}\right] = 0 \tag{1.24}$$

$$\text{Euler} \quad : \quad \left(\frac{\partial \mathbf{v}_\mathrm{p}}{\partial t}\right)_\mathbf{x} + \frac{1}{a}(\mathbf{v}_\mathrm{p}.\nabla_\mathbf{x})\mathbf{v}_\mathrm{p} + H\mathbf{v}_\mathrm{p} = -\frac{1}{a}\nabla_\mathbf{x}\phi - \frac{c_\mathrm{s}^2}{a}\frac{\nabla_\mathbf{x}\rho}{\rho} \tag{1.25}$$

$$\text{Poisson} \quad : \quad \frac{1}{a^2}\nabla_\mathbf{x}^2\phi = 4\pi G\bar{\rho}\delta, \tag{1.26}$$

where $\phi = \Phi - \Phi_0$ is sometimes known as the *peculiar gravitational potential*, and gives the fluctuations in the potential about the homogeneous solution we met earlier. We have also introduced the adiabatic sound speed, $c_\mathrm{s} \equiv \left(\frac{\partial p}{\partial \rho}\right)^{1/2}$.

To make progress with solving these equations, we assume that δ and $\mathbf{v}_\mathrm{p}$ are small, so that we can drop terms that are second order in them. The first-order continuity and Euler equations are

$$\frac{\partial \delta}{\partial t} + \frac{1}{a}\nabla_\mathbf{x}.\mathbf{v}_\mathrm{p} = 0 \tag{1.27}$$

$$\frac{\partial \mathbf{v}_\mathrm{p}}{\partial t} + H\mathbf{v}_\mathrm{p} = -\frac{1}{a}\nabla_\mathbf{x}\phi - \frac{c_\mathrm{s}^2}{a}\nabla_\mathbf{x}\delta, \tag{1.28}$$

where the time derivatives are now implicitly at fixed $\mathbf{x}$.

Taking $\frac{\partial}{\partial t}$(1.27) $- \frac{1}{a}\nabla_\mathbf{x}.$(1.28) and then making substitutions using the comoving continuity and Poisson equations we arrive at

$$\frac{\partial^2 \delta}{\partial t^2} + 2H\frac{\partial \delta}{\partial t} = (4\pi G\rho_0 + \frac{c_\mathrm{s}^2}{a^2}\nabla_\mathbf{x}^2)\delta \tag{1.29}$$

If we take a critical density universe containing only pressureless matter, then $H = 2/3t$, $c_\mathrm{s} = 0$ and $H^2 = \frac{8\pi G\rho_0(t)}{3}$, which implies $4\pi G\rho_0 = 2/3t^2$. Equation (1.29) becomes

$$\frac{\partial^2 \delta}{\partial t^2} + \frac{4}{3t}\frac{\partial \delta}{\partial t} - \frac{2}{3t^2}\delta = 0. \tag{1.30}$$

Then with the *ansatz* $\delta = At^n$, we find that $n = 2/3$ or $n = -1$. The second of these is a decaying mode, and is therefore uninteresting, but $n = 2/3$ leads to density perturbations that increase in magnitude as the Universe evolves. In fact, $\delta \propto t^{2/3}$ means that $\delta \propto a$. So for small perturbations (remember that these results came from linearising the equations, and so are only valid when $\delta \ll 1$) in a matter dominated universe, the fractional density perturbations grow linearly with the scale factor.

This evolution of density perturbations immediately poses a problem when combined with information from the CMB and observations of the local Universe. The temperature fluctuations in the CMB suggest that when $a \approx 0.001$, the amplitude of typical density fluctuations were $\delta \sim 10^{-5}$ (Smoot et al., 1992). Extrapolating this to the present day, we should expect only percent level fluctuations in the density field. However, the Universe we see is rich with gravitationally collapsed structure, $\delta \gg 1$. A good explanation for this arises if there is a second matter component that is not coupled to the photons and

baryons. Its fluctuations can therefore grow independently, and can be larger at the time of recombination than the fluctuations in the baryon-photon fluid. After recombination, when baryons are no longer tightly coupled to the photons, they can fall into the potential wells of overdensities in the other matter component. In fact, such a matter component was already suspected to exist before the fluctuations in the CMB had been measured. We call this *dark matter*.

1.2 Dark matter

Comparing the structure in the local Universe with what we would expect from the CMB in a baryon only Universe was not the first piece of evidence for dark matter (DM). In fact, as early as the 1930's, Fritz Zwicky had discovered evidence for unseen matter in galaxy clusters (Zwicky, 1937). He estimated the mass of the Coma cluster by looking at the line-of-sight velocity distribution of its galaxies. He then compared this value to the total luminosity of the galaxies in Coma, and found that it was around 100 times more massive than one would expect if the mass was just that due to stars. He argued that there must be a dominant mass component in the cluster that was not optically bright, which he called 'dark matter'.

1.2.1 Evidence for dark matter

Since Zwicky's early observations, there have been many more lines of evidence that point towards the existence of DM. The archetypal example is the work on galaxy rotation curves by Vera Rubin in the 1970's. Using a high-resolution spectrograph, she was able to measure the velocity of stars in edge-on spiral galaxies. What she found was that the velocity of stars in the outskirts of spiral galaxies was roughly constant with radius (Rubin et al., 1980), which assuming Newtonian gravity, implies that the enclosed mass grows linearly with radius. This is not true of the visible mass, which is centrally concentrated, and would lead to stellar velocities decreasing towards large radii. If one maintains that Newtonian gravity (equivalent to GR in the weak-field regime relevant for galactic dynamics) is the correct description, then the obvious solution is that there must be non-luminous material

that is more extended than the stellar distribution, i.e. DM.

More recently, the mass of galaxies (Hoekstra & Jain, 2008) and galaxy clusters (Dahle, 2006) has been directly inferred from gravitational lensing. We outline the mathematical details of this technique in §5.3.3, but in short: the path of photons is deflected by gravity.[5] This results in multiple lines of sight leading to the same part of a lensed background object, distorting the image of the lensed object. For simple lenses, such as point masses, lensing typically distorts lensed objects to appear as multiple arcs, or if the lens and source are well aligned along the line of sight, a ring of light known as an *Einstein ring*. The curvature of these arcs or rings tells us about the extent to which light is being bent, and can be used to infer the matter distribution that is doing the lensing (for a review of lensing, see Bartelmann, 2010). For the case of galaxies and galaxy clusters, the results of lensing suggest that the bulk of the lensing mass is otherwise-unseen DM (Massey et al., 2010).

On top of the dynamics and lensing signal of gravitationally collapsed structures, other cosmological probes point towards the existence of DM. These include the *Large scale structure*, essentially the locations of galaxies in the Universe and the way in which they are clustered. When assuming that the matter content of the Universe is dominated by collisionless DM, the structure predicted (Springel et al., 2006) provides a striking match to what is observed (Rodríguez-Torres et al., 2016). This includes detections of the so-called *Baryon Acoustic Oscillation* feature (Cole et al., 2005; Eisenstein et al., 2005), whereby sound waves in the baryons and photons in the early Universe result in a preferred separation between pairs of galaxies – detectable in the local Universe.

All of these pieces of evidence for DM rely on the way DM affects things around it through gravity. This has spurred the development of alternative gravity theories, that would do away with the need for DM, by instead altering the gravitational force law such that the baryonic content is sufficient to explain the observed properties of the Universe. The most popular of these models is *Modified Newtonian Dynamics* (MOND), which posits that for accelerations below some threshold,[6] the force–acceleration relation is modified

[5]The magnitude of the deflections being twice as large in GR than if treating photons as massive particles moving at the speed of light with Newtonian gravity and classical mechanics.

[6]The preferred value for this threshold is typically around $10^{-10}\,\mathrm{m\,s^{-2}}$ (Milgrom, 1983; Brownstein & Moffat, 2006; McGaugh et al., 2016)

such that particles experience larger accelerations than they would according to Newton's second law. While this theory is reasonably successful at explaining galactic rotation curves without needing to invoke DM (Milgrom, 1983; McGaugh & de Blok, 1998; Gentile et al., 2011; McGaugh et al., 2016), this is less true of the other pieces of evidence pointing towards the existence of DM (e.g. Aguirre et al., 2001; Soussa & Woodard, 2004; Pointecouteau & Silk, 2005; Nusser & Pointecouteau, 2006; Takahashi & Chiba, 2007). These failures of MOND are exacerbated by the properties of the theory leading to ambiguities when attempting to apply it on cosmological scales (Scott et al., 2001).

A final challenge for MOND came in the form of the merging galaxy cluster known as the *Bullet Cluster*. This system is discussed extensively in Chapter 5, with an image in Fig. 5.14, but in short: the collision of two massive galaxy clusters has produced a system where the hot X-ray emitting gas has been stripped away from the location of the galaxies (Markevitch, 2006). The mass in gas is roughly ten times more than the mass in the galaxies, but the dominant gravitational lensing signal comes from the location of the galaxies (Bradač et al., 2006). Modifying the lensing in a MONDian sense may be able to boost the strength of the lensing above that from the baryonic material with GR, but as the gas is the dominant baryonic matter component, the lensing signal should come from there (though see Angus et al., 2006). This situation is naturally explained with DM (Clowe et al., 2006), as if the DM is approximately collisionless, then it will not be stripped like the gas, and the dominant mass component would be colocated with the galaxies, as inferred from lensing.

1.2.2 Experimental searches for dark matter

Understanding the nature of DM is now a major goal of not just the astrophysics community, but the particle physics one as well. In fact, the astrophysical evidence for non-baryonic DM is one of the strongest pieces of evidence for physics beyond the Standard Model (SM) of particle physics, driving experimental innovation to uncover its nature. There are three main types of experiment that could hope to shed light on the identity of DM, known as direct detection, indirect detection and production. All three of these rely upon DM being coupled to SM particles through some force other than gravity, and are shown schematically in Fig. 1.1.

Most experiments searching for DM have focussed on so-called weakly-interacting massive particle (WIMP) DM. These are particles with weak interactions, and masses that are similar in mass to the proton or heavier (typically $1 - 100\,\mathrm{GeV}$ Bertone, 2010). A major cause for this as the leading DM paradigm is that a thermally produced WIMP, with a typical weak-force cross-section and mass in the range previously mentioned, has a relic abundance in good agreement with the observed abundance of DM (e.g. Steigman et al., 2012).[7] Such a DM candidate is also popular as it arises naturally in popular extensions to the SM such as supersymmetry (Jungman et al., 1996; Bertone et al., 2005). This being said, there is experimental work being done to find non-WIMP DM. The axion is another popular candidate for DM (Duffy & van Bibber, 2009) because it would simultaneously explain why quantum chromodynamics does not seem to break CP-symmetry (the 'Strong-CP problem' Peccei & Quinn, 1977). There are therefore experiments attempting to detect axionic DM (for a recent review see Graham et al., 2015), although in what follows we will focus on WIMP searches.

Direct detection of DM

Direct detection experiments aim to detect DM particles as they scatter from atomic nuclei in a laboratory. Such a recoil transfers energy to the nucleus, which is then detected in different manners depending on the type of detector. Most common are cryogenically cooled crystal lattices, where a WIMP-nucleus collision produces ionisation and phonon signals, and noble liquid detectors, where interactions with DM lead to scintillation.

The Sun orbits the galaxy at roughly $240\,\mathrm{km\,s^{-1}}$ (Hunt et al., 2016), while the rms velocity of DM particles in the solar neighbourhood is predicted to be $\sim 270\,\mathrm{km\,s^{-1}}$ (Vogelsberger et al., 2009). This means that a DM particle scattering with an Earth-based DM detector will typically have a velocity of order a few hundred $\mathrm{km\,s^{-1}}$. Combining this with an assumed mass of $1 - 100\,\mathrm{GeV}$, the kinetic energy of incident DM particles is $\sim 1 - 100\,\mathrm{keV}$. In this

[7] This result has been dubbed the *WIMP miracle*, though, as pointed out by Subir Sarkar, the same calculation for baryons would predict a number density 10^9 times lower than that actually observed (which we do not call the *baryon disaster*), because for baryons the number density is determined by an initial baryon/anti-baryon asymmetry.

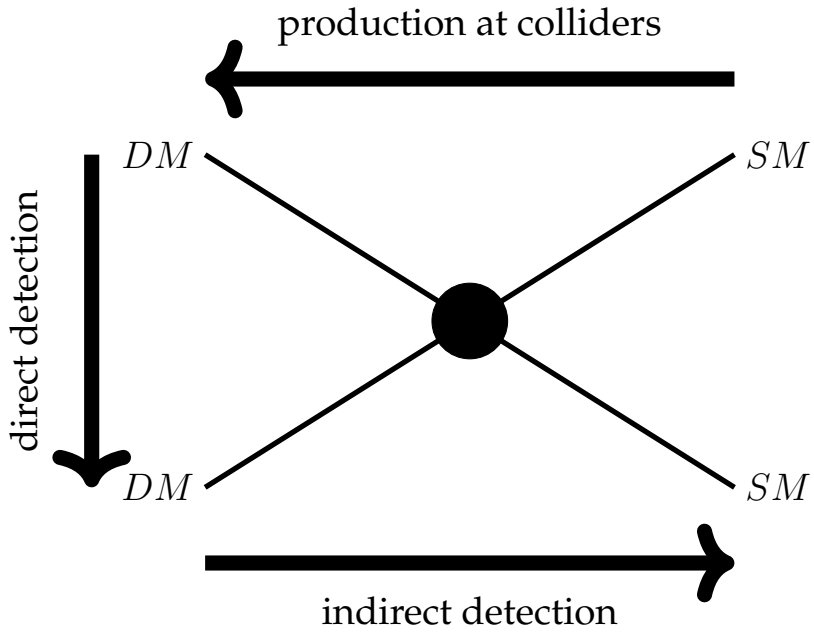


Figure 1.1: An illustration of different methods of detecting DM. In direct detection, a DM particle scatters from a standard model (SM) particle, while indirect detection involves the production of SM particles from DM. Finally, collider searches look for the production of DM particles from high-energy collisions of SM particles.

energy range, there are a significant number of background sources that could be mistaken for DM. In particular: cosmic rays, α-particles, electrons, and photons can have energies in this range, and can scatter from atomic nuclei. The background from cosmic rays can be reduced by placing these experiments deep within mines, as cosmic rays are efficiently stopped by kilometres of earth. To reduce the impact of other backgrounds, a range of materials encase these experiments with the aim of stopping contaminant particles from entering the experiments.[8]

Direct detection experiments have placed limits on the DM-nucleon cross-section for different DM masses, with the tightest constraints coming from LUX (Akerib et al., 2013), XENON100 (Aprile et al., 2013) and SuperCDMS (Agnese et al., 2014). There have been a

[8]The CDMS team shielded their experiment with a range of materials including lead, which is required to block external gamma rays. Terrestrial lead contains significant levels of the unstable isotope ^{210}Pb which would produce unwanted background radiation. CDMS therefore used lead salvaged from a ship that had sunk off the coast of France in the eighteenth century, and had very low ^{210}Pb levels (Ananthaswamy & O'Hara, 2010).

few recent experiments that have not worried about removing non-DM events, and instead looked for an annual modulation in their total rate of events (Bernabei et al., 1998; Aalseth et al., 2011; Angloher et al., 2012). A couple of these have found modulations that would be consistent with the annual modulation in DM events expected due to the motion of the Earth about the sun changing the Earth's velocity through the DM halo on an annual cycle (Bernabei et al., 2003, 2008; Aalseth et al., 2014). However, these detections are inconsistent with other upper limits on DM cross-section from LUX, XENON100 and SuperCDMS, placing severe doubt on these claimed detections.

Indirect detection of DM

Indirect detection is the detection of the products from either the annihilation or decay of DM particles. Assuming a model for the DM particle, the branching fractions to different final state particles can be calculated. The end states that we are interested in are typically γ-rays, cosmic rays and neutrinos, all of which we are able to detect. For example the Fermi Large Area Telescope (Fermi-LAT) can detect γ-rays with energies in the range $20\,\mathrm{MeV}$ to more than $300\,\mathrm{GeV}$ (Atwood et al., 2009). These are the sorts of energies expected from WIMP annihilation, and as the rate of annihilation is proportional to the square of the DM density, looking with Fermi-LAT at regions where we expect a high DM density is a sensitive probe of DM annihilation. In fact, there have been multiple reports of unexplained excesses in the Fermi data (Hooper & Goodenough, 2011; Bringmann et al., 2012), towards the galactic centre where we expect the DM density to be high. However, the interpretation of such excesses as a result of DM remains contentious (e.g. Bartels et al., 2016; Lee et al., 2016; Fornasa et al., 2016), due to other possible sources of γ-rays that would also cluster around the galactic centre.

Production of DM

The prospect of producing DM during the high energy collisions at particle colliders such as the Large Hadron Collider (LHC) provides another window into the nature of DM. Any DM produced would not show up in the particle detectors, but would carry away energy and momentum. Their existence could therefore be inferred from the amount of energy and

momentum unaccounted for after a collision. So far, there are no suggestions of significant DM production at the LHC, and only upper limits have been placed on DM's interaction cross-section with SM particles. These limits are complementary to those from direct detection. For example, current direct detection experiments are most sensitive to DM in the mass range $10 - 1000\,\mathrm{GeV}$, while the LHC is most sensitive for lower mass WIMPs (Rajaraman et al., 2011).

1.2.3 The ΛCDM 'concordance cosmology'

Despite no conclusive detection of DM particles, the overwhelming astrophysical evidence in favour of DM has led to it becoming an established part of cosmology's standard model. The other key ingredient in this standard model is dark energy, which drives an accelerated expansion of the late-time Universe.

Dark energy

This accelerated expansion was first observed by making measurements of the flux-redshift relation of Type Ia supernovae. These supernovae are known as 'standard candles' (or 'standardisable candles'), as their intrinsic luminosity can be inferred from the shape of their light-curve. By combining the measured flux from a supernova with its inferred luminosity, one can calculate the distance[9] to the supernova. By then also measuring the redshift of the supernova, and doing this for many supernovae, we find the distance to different redshifts, from which the expansion history can be calculated. This showed that the rate of expansion is currently accelerating (Riess et al., 1998; Perlmutter et al., 1999), with whatever causes this acceleration dubbed 'dark energy'. The form of the expansion is consistent with the dark energy simply being Einstein's cosmological constant, so that we live in a Universe with a non-zero Ω_Λ. In fact, our current best estimates for the energy content of the Universe suggest that the cosmological constant is the dominant energy component at late times, with $\Omega_{\Lambda,0} \approx 0.7$ and $\Omega_{\mathrm{m},0} \approx 0.3$ (Planck Collaboration et al., 2016).

[9] More precisely, the 'luminosity distance'.

Inflation

We have seen in §1.1.3, how small density perturbations grow in an expanding Universe. However, what gave rise to the initial fluctuations that have evolved into the large scale structure of our Universe today? Historically there were different theories for what seeded this structure. For example, topological defects formed in the early Universe known as *cosmic strings* were thought to be good candidates (Kibble, 1976; Vachaspati, 1986; Brandenberger, 1991). However, precision cosmological measurements, such as those made of the CMB (Smoot et al., 1992; Spergel et al., 2003; Planck Collaboration et al., 2016), suggest that a process known as *inflation* was probably responsible. In the inflationary paradigm, quantum fluctuations are amplified to macroscopic scales by an exponential expansion of the Universe.

That inflation naturally provides a mechanism for generating initial density perturbations is certainly in its favour, but it is not the reason for which an early inflationary period was originally invoked. When Guth (1981) first proposed inflation, it was as a solution to the *Horizon* and *Flatness* problems. The first of these is that the Universe appears to be homogeneous, despite some patches of the Universe that we now observe being causally disconnected. Taking as an example the CMB, different regions of the CMB sky were too far apart (at the time they emitted the CMB photons we observe) to have communicated with each other via signals that travel at the speed of light. Despite this, the temperature of the CMB is uniform to one part in 10^5 across the whole sky. Inflation explains this by modifying the expansion history of the Universe such that the different patches of the CMB sky were once in causal contact (and so could achieve thermal equilibrium with one another), and were only later brought out of causal contact by a rapid expansion of space.

In order to illustrate the flatness problem, we return to the Friedmann equations. Taking equation (1.6) and using $\rho_{\rm crit} = 3H^2/8\pi G$,

$$H^2 = H^2 \frac{\rho}{\rho_{\rm crit}} - \frac{kc^2}{a^2}, \tag{1.31}$$

with ρ the energy density in all forms (matter, radiation and a cosmological constant). Then, using $\Omega = \rho/\rho_{\rm crit}$ and $H = \dot{a}/a$

$$1 - \Omega = -\frac{kc^2}{\dot{a}^2}. \tag{1.32}$$

As the Universe expands, $\dot{a}$ decreases, and Ω is driven away from unity. This means that $\Omega \approx 1$ is unstable. The observed Universe has a total energy density close to the critical density today, which requires that it be extremely close to the critical density in the past. Guth (1981) demonstrated that the same exponential expansion in the early Universe that could solve the the horizon problem would also drive $\Omega \to 1$, with both problems requiring the scale factor to have rapidly increased by a factor $\gtrsim 5 \times 10^{27}$, or 64 'e-folds'.

Before describing the quantitative predictions from inflation, we stop to introduce the power spectrum, $P(k)$. We start by defining the Fourier Transform of the overdensity field

$$\tilde{\delta}(\mathbf{k}, t) = \int \delta(\mathbf{x}, t) \exp(-i\mathbf{k}.\mathbf{x}) \, \mathrm{d}^3\mathbf{x}, \tag{1.33}$$

from which we can define the power spectrum

$$P(k, t) \equiv \left\langle |\tilde{\delta}(\mathbf{k}, t)|^2 \right\rangle . \tag{1.34}$$

The angled brackets denote the expected value, and the power spectrum is a function of $k \equiv |\mathbf{k}|$ because of isotropy. Determining $P(k)$ at different cosmological epochs, and comparing with the predictions of different cosmological models is an important aspect of modern day cosmology.

The power spectrum associated with inflation is $P_{\mathrm{i}}(k) \propto k^{n_{\mathrm{s}}}$, with n_{s} known as the scalar spectral index. The exponential expansion during inflation is driven by the presence of a scalar field, ϕ, known as the inflaton. The value of n_{s} is determined by the potential of the inflaton, $V(\phi)$. A large number of potentials have been studied in the literature, with many of them predicting $n_{\mathrm{s}} \approx 1$ (Tegmark, 2005). In fact, the latest measurements from the Planck satellite suggest $n_{\mathrm{s}} = 0.966 \pm 0.006$, in good agreement with the prediction from a simple quadratic potential $V \propto \phi^2$ with 60 e-folds: $n_{\mathrm{s}} = 0.967$ (Okada et al., 2014).

Transfer function

Given the power spectrum of fluctuations produced by the end of inflation, predicting the large scale structure of the Universe (at least on scales where the density field is linear) is then simply a case of evolving these density perturbations forward in time, following equation (1.29). Since equation (1.29) for the evolution of small perturbations is linear, each

Fourier mode $\tilde{\delta}(\mathbf{k}, t)$ grows independently of the others. We have already seen that in a matter-dominated universe, $\delta \propto a$, but at early enough times the Universe must have been radiation dominated, as radiation is diluted more quickly than matter as the Universe expands. For matter perturbations in a radiation-dominated universe, the solutions to equation (1.29) suggest that perturbations in the matter grow at most logarithmically, as they are stabilised by pressure in the dominant radiation component. This pressure support can only act over scales that are smaller than the particle horizon, which is the maximum distance from which particles could have travelled to the observer in the age of the universe, and so the maximum scale over which causal physics can operate. By considering causally disconnected patches of space as their own mini-universes, and finding how the density in slightly over or under-dense mini-universes evolves as a function of H, it can be shown that super-horizon fluctuations grow as $\delta \propto a^2$ in a radiation-dominated universe, and $\delta \propto a$ in a matter-dominated one. This means that during matter domination, the horizon plays no special role in the growth of fluctuations, but at earlier times the evolution of fluctuations depends strongly on whether they are on scales larger or smaller than the horizon.

The comoving scale of the particle horizon is

$$\chi_{\rm H}(t) = \int_0^t \frac{c\,{\rm d}t'}{a(t')}. \tag{1.35}$$

Before matter-radiation equality at $t = t_{\rm eq}$, the universe is radiation dominated and $a \propto t^{1/2}$. We can write this as $a = (t/t_{\rm eq})^{1/2}\, a_{\rm eq}$, from which the horizon for $t < t_{\rm eq}$ can be found: $\chi_{\rm H} = 2\, c\, t_{\rm eq}\, a/a_{\rm eq}^2$. So in the radiation-dominated era, the comoving horizon grows in proportion to a. If we take two perturbations that enter the horizon in the radiation-dominated era, with comoving wavenumbers k_1 and k_2, and corresponding length scales λ_1 and λ_2, then the ratio of the scale factors when they enter the horizon is $a_1/a_2 = \lambda_1/\lambda_2 = k_2/k_1$. Given that outside the horizon perturbations grow as a^2 and inside the horizon they effectively stop growing, the ratio of the growth of these two perturbations is $(a_1/a_2)^2$. Smaller wavelength perturbations enter the horizon earlier, and have their growth quenched when they do so, resulting in less growth than for a larger wavelength perturbation.

To quantify the effects of causal physics on the growth of structures, we define the trans-

fer function

$$T(k,t) = \frac{\delta(k,t)}{\delta(k,t=0)} \Big/ \frac{\delta(k=0,t)}{\delta(k=0,t=0)} \,. \tag{1.36}$$

In words, this function describes by how much a fluctuation with wavenumber k has grown compared to a fluctuation with $k = 0$. In equation (1.36), this growth is measured relative to the size of these fluctuations at $t = 0$; what is meant by this is the primordial fluctuations, e.g. those present at the end of inflation. Once the transfer function is known, the power spectrum can be written as

$$P(k,t) = P_{\rm i}(k)\, T^2(k,t)\, D^2(t). \tag{1.37}$$

$D(t)$ is the linear growth factor, describing the growth of long wavelength perturbations ($k \to 0$).

If we define the wavenumber that enters the horizon at $t_{\rm eq}$ to be $k_{\rm eq}$, then the above discussion suggests that the transfer function should be constant for $k < k_{\rm eq}$, and $\propto k^{-2}$ for $k > k_{\rm eq}$. Note however, that a more detailed calculation of the transfer function (e.g. Eisenstein & Hu, 1998) shows that the transition from $T \approx \text{constant}$ to $T \propto 1/k^2$ is gradual. The full picture is also complicated by the different constituents of the Universe. DM, baryons and neutrinos all evolve differently in the early Universe, and the evolution of density perturbations must be determined numerically, using codes like the publicly available CMBFAST (Seljak & Zaldarriaga, 1999).

As well as the abundances of the different particle species, the nature of DM can affect the form of $T(k,t)$. Heavy DM particles will be non-relativistic at $t_{\rm eq}$, and so there will be negligible DM pressure during the matter-dominated era – when the small scale fluctuations that lead to galaxies grow. Lighter DM particles can have significant thermal velocities at $t_{\rm eq}$, leading to a suppression of small-scale fluctuations by the free-streaming of DM.

The temperature at $t_{\rm eq}$ is roughly $1\,\mathrm{eV}$, so thermally-produced DM with this sort of mass will lead to significant levels of free-streaming, removing structure on scales below that of galaxy clusters. This sort of DM is known as *hot*, and was shown to be inconsistent with the observed galaxy clustering in the Universe by White et al. (1983). Slightly heavier DM, with a keV mass, would smooth out fluctuations below a comoving scale of around $100\,\mathrm{kpc}$ (Viel

et al., 2005). This DM is known as *warm*, and would suppress the number of low-mass DM structures. Finally, heavier DM leads to negligible free-streaming, and is known as *cold*. For example, a $100\,\mathrm{GeV}$ WIMP (around the mass expected for a SUSY neutralino) would only lead to a suppression of DM structure at masses below an Earth mass (Angulo & White, 2010). This is far below the mass of the DM haloes that host galaxies, and so free-streaming in this case has a negligible impact on the universe we observe.

1.2.4 Challenges facing ΛCDM

The ΛCDM model has been immensely successful at explaining observations of our Universe, and one could worry that this would leave cosmologists with nothing left to uncover. However, there are still theoretical challenges surrounding our ignorance towards the identity of the dark matter and what is driving the late-time accelerated expansion.

Problems with the cosmological constant

The observed late-time acceleration is well described by a cosmological constant, which is unclustered and has an energy density that is constant in time. However, the inferred value of this constant is considered 'unnatural', being many orders of magnitude smaller than the zero-point energy suggested by quantum field theory (Carroll, 2001). This is known as the *cosmological constant problem*. The *coincidence problem* is a second challenge for the naturalness of the measured $\Omega_{\Lambda,0}$; given that matter and a cosmological constant have energy densities that evolve very differently with scale factor, it is surprising that we are viewing the Universe at a time when their energy densities are similar. Also surprising, is the comparable magnitude of the density in baryons to the density in DM, given that most models for DM production are not closely related to baryogenesis.

The missing satellites problem

Aside from these somewhat philosophical questions, there are discrepancies on the scale of individual galaxies between observations and the predictions from simulations with ΛCDM. Kauffmann et al. (1993) used 'Monte Carlo merger trees' (Kauffmann & White,

1993) to demonstrate that Milky Way size haloes with CDM should contain well over 100 subhaloes capable of hosting observable satellite galaxies, a factor of 5–10 higher than the number that had actually been observed. Using N-body simulations, Klypin et al. (1999) made similar predictions for the amount of substructure expected around the Milky Way, with a paper title that asked the question "Where are the Missing Galactic Satellites?", from which the *missing satellites problem* takes its name. Moore et al. (1999) similarly found that a large number of subhaloes should survive within the Milky Way halo, which has now been confirmed by much higher resolution simulations of Milky Way-like haloes with CDM (Diemand et al., 2008; Springel et al., 2008).

The problem of the relatively low number of observed Milky Way satellite galaxies, when compared with the large number of predicted DM substructures, has multiple possible solutions. An obvious solution is that small DM subhaloes are inefficient at forming stars. The photoionisation background suppresses the cooling of primordial gas (Efstathiou, 1992; Bullock et al., 2000; Benson et al., 2002; Somerville, 2002), which could be sufficient to explain the dearth of observed satellite galaxies (Gnedin & Kravtsov, 2006). Another exciting alternative, is that this problem is telling us something about the nature of DM. For instance, warm DM would lead to a smaller number of DM subhaloes around the Milky Way than CDM.

Since this problem was first raised, the number of faint satellite galaxies has increased dramatically (Willman, 2010; McConnachie, 2012; Bechtol et al., 2015; Drlica-Wagner et al., 2015) with over 50 satellite galaxies of the Milky Way now known, and suggestions that a factor of 3–5 more satellites are still to be discovered (Hargis et al., 2014). At the same time, our understanding of reionization and the effects of stellar feedback on further star formation (Ceverino & Klypin, 2009; Ceverino et al., 2014; Hopkins et al., 2014; Trujillo-Gomez et al., 2015) have also improved, and with it, the theoretical prediction for the number of luminous satellites has fallen. In fact, the missing satellites problem no longer appears to be much of a problem, and far from pointing towards a need for non-standard DM, the number of observed satellites can be used to rule out DM models that would erase too much substructure (Macciò & Fontanot, 2010; Lovell et al., 2014; Kennedy et al., 2014; Bose et al., 2017).

The core/cusp problem

Simulations of structure formation in a CDM universe predict a universal density profile, with $\rho \propto 1/r$ in the central regions, irrespective of halo mass or cosmology (Navarro, Frenk & White, 1997). This increasing central density towards the centre of a DM halo is known as a *cusp*. On the other hand, observations of objects ranging in size from dwarf galaxies (Łokas, 2002; de Blok et al., 2003; Oh et al., 2011) to galaxy clusters (Sand et al., 2004, 2008; Newman et al., 2009, 2013) typically favour the presence of a shallower central density slope, or even constant density DM *cores*.

Solutions to this apparent discrepancy within a CDM framework typically invoke baryonic feedback processes that can alter the DM profile in the inner regions (Navarro et al., 1996; Mashchenko et al., 2006; Governato et al., 2010; Zolotov et al., 2012; Governato et al., 2012; Pontzen & Governato, 2012; Teyssier et al., 2013), or observational biases, that can lead to inferring a core when a cuspy halo is present (Swaters et al., 2003; Hayashi et al., 2004; Dalcanton & Stilp, 2010; Pineda et al., 2017). However, not all observed galaxies appear to have cores (Oman et al., 2015), and this problem is now better characterised by the statement that observed galaxies appear to have more diverse rotation curves than simulated ones, with some observed galaxies having large deficits in the inferred amount of mass in their inner regions. We discuss the diversity of dwarf galaxy rotation curves shortly.

Too big to fail

Boylan-Kolchin et al. (2011) made a detailed comparison between local dwarf galaxies and CDM-only simulations. Their most massive DM substructures around simulated Milky Way-like haloes were considerably more massive than estimated dwarf galaxy masses made from line-of-sight velocity measurements (Walker et al., 2009; Wolf et al., 2010). If the results from the N-body simulations are representative of the real Universe, there must be a significant number of massive dark subhaloes around the Milky Way. These subhaloes that do not contain stars despite their large mass have been dubbed 'too big to fail' (TBTF).

Immediately one can see that this problem is related to the previous two: if there were fewer DM substructures then the 'most massive' substructures would extend down to

lower masses and hence lower circular velocities, while if DM haloes have cores this would lower the circular velocities at fixed subhalo mass. The reason that this problem gained particular traction is that the observational measurement (of the mass enclosed within the half-light radius) is thought to be particularly robust, being largely insensitive to the unknown velocity anisotropy (Wolf et al., 2010) that makes mass determinations within both larger and smaller radii uncertain.

Much like the missing satellites and core/cusp problems, TBTF has had numerous proposed solutions within the ΛCDM framework. It has been demonstrated that strong gas outflows (driven by supernovae for example) can lead to rapid fluctuations in the gravitational potential that transfer energy to the DM particles and reduce the central density of DM (Navarro et al., 1996; Pontzen & Governato, 2012; Governato et al., 2012; Oñorbe et al., 2015). Other authors have argued that there is not sufficient energy in the supernovae that have taken place within dwarf galaxies to remove the amount of DM mass from the central regions required to solve TBTF (Peñarrubia et al., 2012; Garrison-Kimmel et al., 2013), although recent work (Sawala et al., 2016; Dutton et al., 2016; Fattahi et al., 2016) suggests that it may be possible to match dwarf galaxy data without extreme outflows or the formation of cored DM profiles. The loss of baryons due to reionisation as well as environmental effects such as ram pressure stripping of gas and tidal stripping of DM, lower the mass of haloes hosting dwarf satellite galaxies, producing results consistent with observations. However, these environmental processes should not be important for field galaxies where it has also been reported that there is a TBTF problem (Garrison-Kimmel et al., 2014; Klypin et al., 2015; Papastergis et al., 2015).

The diversity of dwarf galaxy rotation curves

Dwarf galaxies have a striking diversity of measured rotation curves (Oman et al., 2015). This appears to be at odds with CDM, which predicts a universal density profile for DM haloes (Navarro et al., 1997). While this diversity may derive from incorrectly determined distances to and inclinations of these dwarf galaxies (Oman et al., 2016; Read et al., 2016), if dwarf galaxies really do have diverse rotation curves then the ΛCDM paradigm would need substantial revision. One such revision, would be if DM particles could interact with

one another through forces other than just gravity. DM that behaves in this way is known as self-interacting dark matter (SIDM), and we will now discuss the motivation for invoking such DM.

1.3 Self-interacting dark matter

Motivated by the small scale challenges to ΛCDM, Spergel & Steinhardt (2000) proposed that DM could be self-interacting. In their original proposition, these interactions are isotropic elastic scatterings with an interaction cross-section that is independent of velocity. If the cross-section for these interactions is sufficiently large, then structure formation on small scales is altered from the case of collisionless DM. The redistribution of energy and momentum by DM particle collisions decreases the central density of DM haloes and tends to make the DM velocity distribution isotropic, leading to more spherical haloes (Spergel & Steinhardt, 2000; Burkert, 2000; Yoshida et al., 2000b). The presence of such a dark force would have significant implications for both particle physics and astrophysics. Large self-interactions would rule out some of the most popular DM candidates such as axions (Duffy & van Bibber, 2009) or supersymmetric neutralinos (Bertone et al., 2005), while changing cosmological structure formation on small scales.

1.3.1 Astrophysical motivation and constraints

Dwarf galaxies

The main astrophysical motivation for SIDM comes from the small-scale challenges faced by the ΛCDM model that were previously discussed (for a review see Weinberg et al., 2015). While there is strong evidence that baryonic processes can alter the CDM predictions from those of a DM-only universe (e.g. Schaller et al., 2015a), it is not yet clear whether correctly including the relevant astrophysics will be sufficient to bring simulated CDM dwarf galaxies in line with those that are observed. An exciting alternative is that these small-scale discrepancies are telling us something about the nature of DM. CDM can be modified by removing the assumption that DM is collisionless, and instead allowing DM particles to

scatter with one another. For moderate scattering cross-sections, the success of ΛCDM in describing the observed large-scale structure of the Universe is maintained, while the predictions for the internal-structure of DM haloes (including those that host dwarf galaxies) are altered.

Initial excitement about SIDM was related to its ability to produce constant density cores in dwarf galaxies (Spergel & Steinhardt, 2000; Yoshida et al., 2000b), as well as reduce the amount of substructure in DM haloes – thus addressing the core/cusp and missing satellites problems. The reduction in substructure was proposed to happen by the process of *evaporation*: when a subhalo with a low velocity dispersion travels through a more massive, high velocity dispersion halo, the scattering of subhalo DM particles with main halo DM particles will most often result in neither particle being bound to the subhalo. Through this process, the DM in the subhalo will gradually be lost to the main halo, particularly for subhaloes passing close to the centre of the main halo. The large DM–DM scattering cross-sections required to significantly reduce the expected number of dwarf galaxy-hosting subhaloes around the Milky Way have now been ruled out (D'Onghia & Burkert, 2003), although currently allowed cross-sections can alter the mass function of subhaloes in the inner regions of galaxies and galaxy clusters (Rocha et al., 2013).

The production of constant density cores also naturally solves the TBTF problem. The half-light radii of the observed dwarf galaxies used to infer the TBTF problem (Boylan-Kolchin et al., 2011) are $\lesssim 1\,\mathrm{kpc}$ (Wolf et al., 2010). This is within the radius that is significantly affected by moderate SIDM cross-sections (Vogelsberger et al., 2012; Rocha et al., 2013; Zavala et al., 2013; Vogelsberger et al., 2014b; Fry et al., 2015), with the reduced central densities in SIDM haloes lowering the circular velocity profiles in the inner regions and bringing simulated SIDM-only systems into better agreement with stellar velocity dispersion measurements than CDM-only simulations (Elbert et al., 2015). Getting good agreement between the TBTF observations and SIDM-only simulations does not require a fine-tuned cross-section. The non-linear relationship between the SIDM cross-section and DM halo core sizes means that cross-sections spanning at least two orders of magnitude ($\sigma/m = 0.5$–$50\,\mathrm{cm^2\,g^{-1}}$) can alleviate TBTF (Elbert et al., 2015).

Recently, it has been suggested that SIDM might also explain the striking diversity of

measured galaxy rotation curves (Kaplinghat et al., 2014b; Pace et al., 2016; Elbert et al., 2016; Creasey et al., 2017). The interaction between a baryonic disk and an SIDM halo in which it lives could lead to small changes in the baryonic component of a dwarf galaxy producing dramatic changes in the total density profile. This interaction between baryons and the density profile of SIDM haloes has been studied in the context of analytical models (Kaplinghat et al., 2014b; Pace et al., 2016) as well as N-body simulations imposing an analytical disk potential (Elbert et al., 2016; Creasey et al., 2017). In both cases it has been demonstrated that SIDM can both increase and decrease the central density of DM in the presence of baryons, depending on how centrally concentrated the baryonic component is. For a diffuse baryonic component, SIDM alters the DM density profile in the same way as for the DM-only case, producing a constant density core and reducing the mass in the centre of the halo. However, a concentrated baryonic component can contract the DM halo (Blumenthal et al., 1986; Jesseit et al., 2002; Gnedin et al., 2004). With CDM this leads to a mild increase in the central density compared with the DM-only case, but with SIDM this contraction increases the rate of scattering, which can become so efficient at transporting energy away from the centre of the halo that it induces 'core collapse' (Kochanek & White, 2000; Balberg et al., 2002; Koda & Shapiro, 2011, and see also §4.2.5) and the production of very dense halo cores. SIDM may therefore explain the diverse rotation curves found by Oman et al. (2015), which poses a problem for CDM (though see Read et al., 2016), even with the inclusion of baryonic feedback processes that can alleviate the TBTF problem (Oman et al., 2016).

Galaxy clusters

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 inferred offset between one of the bright central galaxies in Abell 3827 and its DM halo (Massey et al., 2015) is interpreted as resulting from SIDM, then it corresponds to an isotropic scattering cross-section of $\sigma/m \sim 1.5\,\mathrm{cm}^2\,\mathrm{g}^{-1}$ (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., 2015c).

There are also suggestions from combined lensing and stellar kinematics studies that observed galaxy clusters have a deficit of DM in their centres compared with what is predicted (Sand et al., 2004, 2008; Newman et al., 2009, 2013). That being said, other lensing studies find agreement with CDM predictions (e.g. Okabe et al., 2013), and the low central DM densities found by some authors may be explained by the assumption of isotropic stellar orbits and/or a systematic error in the assumed stellar mass-to-light ratio (for a detailed discussion see Schaller et al., 2015b). If taken at face value, then the DM density profiles inferred by Newman et al. (2013) would be consistent with them being formed by SIDM with $\sigma/m \approx 0.1\,\mathrm{cm}^2\,\mathrm{g}^{-1}$ (Kaplinghat et al., 2016).

SIDM constraints

Since SIDM was first proposed as an alternative to collisionless DM, work has been done to constrain the self-interaction cross-section. Due to the high densities and relative velocities of particles, galaxy clusters are the systems in which the scattering rate would be the highest, unless the cross-section is a decreasing function of relative velocity. This means that clusters have been used to place the tightest constraints on velocity-independent SIDM cross-sections. The absence of large offsets between the DM and collisionless galaxies in cluster mergers have been used to limit the cross-section to $\sigma/m \lesssim 0.5\,\mathrm{cm}^2\,\mathrm{g}^{-1}$ (Markevitch et al., 2004; Randall et al., 2008; Dawson et al., 2012; Kahlhoefer et al., 2014; Harvey et al., 2015), with similar constraints being derived from the consistency in mass-to-light ratios of post-collision systems and isolated clusters (Randall et al., 2008). Kim et al. (2017) have recently shown that SIDM in cluster mergers can lead to long-lived offsets between brightest cluster galaxies (BCGs) and their associated cluster-scale DM haloes, potentially allowing cross-sections to be constrained to $\sigma/m \lesssim 0.1\,\mathrm{cm}^2\,\mathrm{g}^{-1}$.

The shapes and radial density profiles of clusters have provided complementary tests of the SIDM paradigm. The core sizes (Yoshida et al., 2000b; Meneghetti et al., 2001) and ellipticities (Miralda-Escudé, 2002) of clusters initially provided very stringent constraints, with claims that the observed triaxiality of clusters implied $\sigma/m \lesssim 0.02\,\mathrm{cm}^2\,\mathrm{g}^{-1}$ (Miralda-

Escudé, 2002). A cross-section of this size would render SIDM essentially irrelevant in the dwarf galaxies where it was originally hoped that it could bring theoretical predictions and observations in to line. However, this tight constraint came from an analytical calculation, which assumed that DM haloes should be close to spherical within a radius where DM particles have interacted on average once. Peter et al. (2013) revisited this constraint, by analysing N-body simulations of galaxy clusters with SIDM. They found that in 3D, haloes could still have significant asphericity inside of the radius at which particles had scattered on average once since the halo formed, and that this was exacerbated by the fact that the observational probes used by Miralda-Escudé (2002) were sensitive to the projected density, which even in the centre of a halo receives a contribution from mass at large 3D radii, which can be significantly triaxial. Peter et al. (2013) therefore revised the constraints placed by Miralda-Escudé (2002), finding that DM with $\sigma/m \lesssim 1\,\mathrm{cm}^2\,\mathrm{g}^{-1}$ was still viable. Rocha et al. (2013) used the same set of N-body simulations, and derived similar constraints from comparing the central densities of their simulated clusters with those of observed clusters.

Away from galaxy clusters, there are fewer constraints on the DM cross-section. On dwarf galaxy scales, it remains an open question whether the cores formed in SIDM bring simulations into closer or more distant agreement with observations. Also, with very large cross-sections, the DM density at the centre of haloes actually increases (Kochanek & White, 2000; Moore et al., 2000; Yoshida et al., 2000a; Koda & Shapiro, 2011), such that there is no SIDM cross-section for which cores in dwarf galaxies become too large to (currently) be allowed. Milky Way-like galaxies have the highest ratios of stellar mass to total mass (Behroozi et al., 2013; Moster et al., 2013; Vogelsberger et al., 2014a; Schaye et al., 2015), and have inner regions that are dominated by baryons. These are the same regions where SIDM effects would show up, and so it is hard to discern SIDM from CDM using such galaxies. That being said, Di Cintio et al. (2017) have recently shown that the decreased dynamical friction caused by the lower DM density in SIDM galaxies compared with CDM, can result in large offsets between supermassive black holes and the centre of their host galaxies, not expected in CDM. The presence or absence of such offsets may provide a test of SIDM on Milky Way-like galaxy scales.

Velocity-dependent SIDM

The tightest constraints on the DM–DM scattering cross-section come from galaxy cluster scales, where the relative velocity between DM particles is high. Meanwhile SIDM's ability to solve the TBTF problem, or produce diverse rotation curves, is on the galaxy or dwarf galaxy scale. This was noted by Fry et al. (2015) who found that cross-sections consistent with cluster scale constraints could not significantly reduce the central density of haloes with peak circular velocities below $30\,\mathrm{km\,s^{-1}}$. For this reason, there has been increased interest in SIDM with a velocity-dependent cross-section (Loeb & Weiner, 2011; Vogelsberger et al., 2012; Zavala et al., 2013; Vogelsberger & Zavala, 2013). As we will see in the next section, a cross-section that decreases with increasing relative velocity is naturally achieved within many particle physics models that give rise to SIDM.

The velocity dependence that arises in particle models of SIDM is often similar to scattering in a Coulomb potential, where the cross-section $\propto 1/v^4$ (Ackerman et al., 2009; Feng et al., 2010; Buckley & Fox, 2010). This strong velocity dependence means that the cross-section in dwarf galaxies can be many orders of magnitude larger than on cluster scales. As such, these models can readily produce effects in dwarf galaxies while evading all current constraints that come from galaxy clusters. For velocity-dependent SIDM models, observations at different velocity scales provide complementary information, so it is important to investigate SIDM effects in a wide range of systems (Kaplinghat et al., 2016).

1.3.2 Particle physics models

The cross-sections for DM interactions with SM particles are experimentally constrained to be very small, suggesting that if DM can interact through the strong, weak or electromagnetic forces that it does so with small cross-sections that are insufficient to affect structure formation. Because of this, particle models of SIDM typically invoke a new force in the dark sector. In most cases these models are built precisely because they would be interesting astrophysically, a potential weakness of the SIDM paradigm, though in Feng et al. (2010); Buckley & Fox (2010); Finkbeiner et al. (2011) there are examples of SIDM behaviour

arising naturally from DM models designed to explain an unrelated anomaly.[10]

There is an ever-growing menagerie of particle physics models of DM that would behave in an SIDM-like manner (Bento et al., 2000; Kusenko & Steinhardt, 2001; Mohapatra et al., 2002; Foot, 2004; Feng et al., 2009; Ackerman et al., 2009; Loeb & Weiner, 2011; Blennow et al., 2012; Tulin et al., 2013b; Petraki & Volkas, 2013; Cline et al., 2014b; Petraki et al., 2014; Boddy et al., 2014; Kouvaris et al., 2015; Soni & Zhang, 2016; Kainulainen et al., 2016; Han & Zheng, 2017), though not all of them lead to the simple behaviour originally invoked by Spergel & Steinhardt (2000). In some models, the DM can undergo inelastic or dissipative processes (Fan et al., 2014; Schutz & Slatyer, 2015; Foot & Vagnozzi, 2016; Boddy et al., 2016), with the potential for DM bound states to form (Boddy et al., 2016; Cirelli et al., 2017). Many of these models lead to DM scattering that is anisotropic and/or has a strong velocity-dependence (Ackerman et al., 2009; Loeb & Weiner, 2011; Blennow et al., 2012; van den Aarssen et al., 2012; Tulin et al., 2013b; Petraki & Volkas, 2013; Petraki et al., 2014), which could be appealing given our discussion above.

Some proposed scenarios are more elaborate still, with only a sub-dominant component of the DM being self-interacting (Fan et al., 2013; Boddy et al., 2014). A complicated dark sector such as this could result in novel phenomenology, potentially explaining the formation of supermassive black hole seeds at high redshift (Pollack et al., 2015) and (even more speculatively) the demise of the dinosaurs (Randall, 2015).

While SIDM may not be unique in offering a solution to these 'small-scale problems' (Pontzen & Governato, 2012; Governato et al., 2012; Sawala et al., 2016), there are numerous DM particle candidates that give rise to scattering between DM particles (Carlson et al., 1992; Bento et al., 2000; Kusenko & Steinhardt, 2001; Mohapatra et al., 2002; Feng et al., 2009; Tulin et al., 2013b; Kaplinghat et al., 2014a; Boddy et al., 2014; Wang et al., 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 then this would have a profound effect on particle physics theories of DM, ruling

[10] In this case an observed excess of cosmic ray positrons in the 1 to 100 GeV energy range (Adriani et al., 2009).

out many of the favoured (and most searched for) DM candidates.

Finally, we note that late-time scattering affecting the distribution of DM in gravitationally collapsed objects is not the only way for DM–DM interactions to alter cosmological structure formation. If the DM scattering is mediated by light particles (dark radiation), then primordial perturbations in the dark sector undergo collisional (Silk) damping on scales below the dark radiation's diffusion length (Buckley et al., 2014). This leads to a power spectrum of fluctuations that is qualitatively similar to those from warm DM (Bringmann et al., 2016), although BAO-like oscillations in the DM-dark radiation fluid can lead to oscillations in the power spectrum within the damping envelope (Buckley et al., 2014; Cyr-Racine et al., 2016). However, there are SIDM models for which the power spectrum is unchanged from that for collisionless CDM, and even for models that lead to a suppression of small scale power, this suppression is not uniquely defined by the late-time scattering properties of the model (Buckley et al., 2014; Vogelsberger et al., 2016). In other words, a particular SIDM cross-section can be achieved by different combinations of particle physics parameters, which would give different length scales below which the power spectrum is suppressed. For this reason, this thesis focusses only on the late time scattering effects of SIDM, assuming a CDM initial power spectrum throughout.

1.4 Thesis outline

The remainder of this thesis is organised as follows. Chapter 2 introduces a method to estimate the rate of SIDM scattering in collapsed structures throughout the history of the Universe, showing at what times and in which objects we expect significant rates of DM scattering. In Chapter 3 we provide a short summary of N-body simulation techniques, before going on to discuss how we implemented DM scattering into the GADGET N-body code. Chapter 4 is then a suite of tests that demonstrate the efficacy of our SIDM implementation, with a discussion of the range of numerical parameters for which we achieve converged results. Chapters 5 and 6 then use this code to simulate the Bullet Cluster with SIDM, first with a simple model for DM scattering and focussing on the importance of the techniques used when comparing the simulation results with observations, and then

looking at how the predictions change when using more complicated models of particle scattering. Finally we summarise our findings in Chapter 7.

Chapter 2

Self-interacting dark matter scattering rates through cosmic time

2.1 Introduction

Assessing the effects of dark matter particle phenomenology on structure formation is usually done using cosmological simulations. However these simulations can only access a finite range of objects due to their limited resolution. An alternative to simulations, originally pioneered by Press & Schechter (1974, hereafter PS) and later extended by use of Excursion Set Theory (Bond et al., 1991) and consideration of ellipsoidal collapse (Sheth et al., 2001), is used to calculate 'analytical' mass functions. This is done using linear theory to evolve the density field, and assuming a simple model for gravitational collapse in which regions denser than some density threshold collapse to form virialised objects. Using the PS formalism is attractive as it allows us to look at all scales and redshifts simultaneously, while we can easily separate the contribution from haloes of different masses to quantities such as the mean scattering rate for SIDM particles through cosmic time.

This chapter follows a similar procedure to Cirelli et al. (2009), who estimated the DM annihilation rate through cosmic time. The rate of interactions in a DM halo can be calculated given a particle model and the density profile of the halo. Then with a mass function (from PS theory or equivalent) it is possible to work out the total rate of scattering in the universe. For the simplest model of particle annihilation, the DM cross-section, σ, mul-

tiplied by the relative velocity of particles, v, is constant. As the rate of interactions is proportional to $\langle \sigma v \rangle$ this simplifies the calculation relative to a case where σ has some other velocity dependence. In this work we use DM models that have interaction cross-sections that differ from $\sigma \propto 1/v$, first using the simplest model for particle scattering in which σ is a constant.

This chapter is aimed at estimating the rate of scattering in DM haloes of different masses through cosmic time. The high redshift Universe is very dense, and were it to turn out that the scattering rate was therefore high, the survival of the first seeds of structure formation could provide a useful constraint on the self-interaction cross-section of DM. This work should also be helpful when assessing the importance of resolution in N-body simulations of SIDM, because they can only resolve objects above a certain mass. While only the resolved objects from simulations are usually of interest, objects build up in a hierarchical fashion, such that resolved objects at some epoch, are made from the merging of smaller (potentially unresolved) objects from an earlier time. It is therefore important to assess whether these small objects that merged should have been affected by DM self interactions.

The chapter is organised as follows. In §2.2 we discuss the calculation of the DM interaction rate through cosmic time for a velocity-independent scattering cross-section, while in §2.3 we show the effects of changing the models and parameters that went into our original calculation. In §2.4 we perform the same calculation with velocity-dependent cross-sections, focussing in particular on two models simulated by Vogelsberger & Zavala (2013). Finally, we give our conclusions in §2.5. Throughout this chapter we assume a Planck 2013 cosmology (Planck Collaboration et al., 2014) unless stated otherwise, and also assume that self-interactions do not affect large scale structure formation.

2.2 Interaction rate over cosmic time

In this section we first discuss the number density of DM haloes of different masses and how this evolves with redshift. By then looking at the scattering rate of DM particles in the haloes that exist at a particular redshift we can calculate the rate of DM particle scattering

at that epoch. This calculation assumes that DM scattering is only between particles within the same DM halo, and neglects the fact that scattering rates would be enhanced during the merging of DM haloes, when the relative velocities between particles can be larger. As haloes only spend a small fraction of time undergoing major mergers, the contribution of mergers to the integrated number of scatterings should not be too significant.

2.2.1 Mass function of collapsed structures

We initially calculate the number of structures of a given mass using Press-Schechter (PS) theory, considering alternative formulations in §2.3.2. The primordial fluctuations $\delta = (\rho - \langle\rho\rangle)/\langle\rho\rangle$ in the Universe's matter density field ρ, are evolved using linear theory as discussed in §1.1.3. The spherical collapse model (e.g. Lacey & Cole, 1993) predicts that volumes of radius R in which the mean overdensity δ_R exceeds a critical threshold $\delta_R > \delta_c = 1.686$ will collapse under their own gravity. We assume gravitational collapse to be immediate leading to a virialised halo with mass $M = \frac{4}{3}\pi R^3\langle\rho\rangle$.

To find these volumes, consider smoothing the density distribution on a scale R. Assuming the density fluctuations form a Gaussian random field, the fraction of the Universe in regions with an overdensity greater than δ_c is

$$F(\delta_R > \delta_c) = \int_{\delta_c}^{\infty} \frac{1}{\sqrt{2\pi\sigma_R^2}} \exp\left(-\frac{\delta_R^2}{2\sigma_R^2}\right) \mathrm{d}\delta_R. \tag{2.1}$$

This depends only on σ_R^2, the variance of δ_R on this scale. Because δ_R has zero mean,

$$\sigma_R^2 = \left\langle\delta_R^2\right\rangle = D^2(z) \int k^2 P(k) \widetilde{W}_R^2(k)\, \mathrm{d}k, \tag{2.2}$$

where the linear growth factor $D(z)$ governs the amplitude of perturbations at redshift z, and $\widetilde{W}_R(k)$ is the Fourier Transform of a real-space spherical top hat filter of radius R.

The power spectrum $P(k)$ is obtained by multiplying the power spectrum of fluctuations generated by inflation by the Transfer Function $T(k)$, as outlined in §1.2.3. For sim-

plicity we use the Eisenstein & Hu (1998) zero-baryon CDM model in which

$$\begin{aligned} T(q) &= \frac{L_0}{L_0 + C_0\, q^2}, \\ L_0(q) &= \ln(2e + 1.8q), \\ C_0(q) &= 14.2 + \frac{731}{1 + 62.5q}, \end{aligned} \tag{2.3}$$

and q is related to k by

$$q = \frac{k}{\Omega_m\, h^2\, \mathrm{Mpc}^{-1}}\, \Theta_{2.7}^2\,, \tag{2.4}$$

where $T_{\rm CMB} = 2.7\,\Theta_{2.7}\,\mathrm{K}$. We look at the effect of changes to the high-k power spectrum by integrating the mass function down to different minimum masses, as explained in §2.2.3.

PS theory then interprets the fraction of the Universe's volume for which $\delta_R > \delta_c$ as the fraction of the Universe's mass that has collapsed to form objects with mass $M \geq \frac{4}{3}\pi R^3\,\langle\rho\rangle$. In this transition from smoothing over volumes to mass scales, it is also convenient to eliminate time dependence from the rms density fluctuations, i.e. we define the rms mass fluctuations on scale M as $\sigma_M \equiv \sigma_R(z)/D(z)$, such that $D(z=0) = 1$. Thus the fraction of the mass in the Universe in collapsed objects with mass greater than M, at redshift z, is

$$F(M, z) = \int_{\delta_c/\sigma_M D(z)}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{\xi^2}{2}\right)\, \mathrm{d}\xi, \tag{2.5}$$

where $\xi = \delta_M/\sigma_M D(z)$. This depends only on the rms density fluctuations (in the lower limit of integration) and the linear growth factor. Differentiating it with respect to mass yields the *multiplicity function*

$$\frac{\mathrm{d}F}{\mathrm{d}\,\ln M}(z) = \sqrt{\frac{2}{\pi}} \left|\frac{\mathrm{d}\,\ln\sigma_M}{\mathrm{d}\,\ln M}\right| \nu \exp\left(-\frac{\nu^2}{2}\right), \tag{2.6}$$

where we have introduced $\nu = \delta_c/\sigma_M D(z)$ and multiplied by a factor of two to account for mass that is initially in under-dense regions.[1] This describes how the mass in the Universe is divided amongst objects of different mass and is plotted in the top panel of Fig. 2.1.

[1] Consider what happens when we take $M \to 0$ in equation (2.5). On small scales the rms fluctuations are very large, and the lower limit in the integration tends to zero. This implies $F(0, z) = \frac{1}{2}$, and only half of the mass in the Universe is in collapsed objects. On small enough scales the density field is always non-linear, and so we would expect all mass in the Universe to be in collapsed objects if we take $M \to 0$. The missing half of the Universe corresponds to regions that are below the collapse threshold when smoothed on a scale M, but would be above the collapse threshold if smoothed on some larger scale. For more information, see the discussion of the 'cloud-in-cloud' problem in Bond et al. (1991).

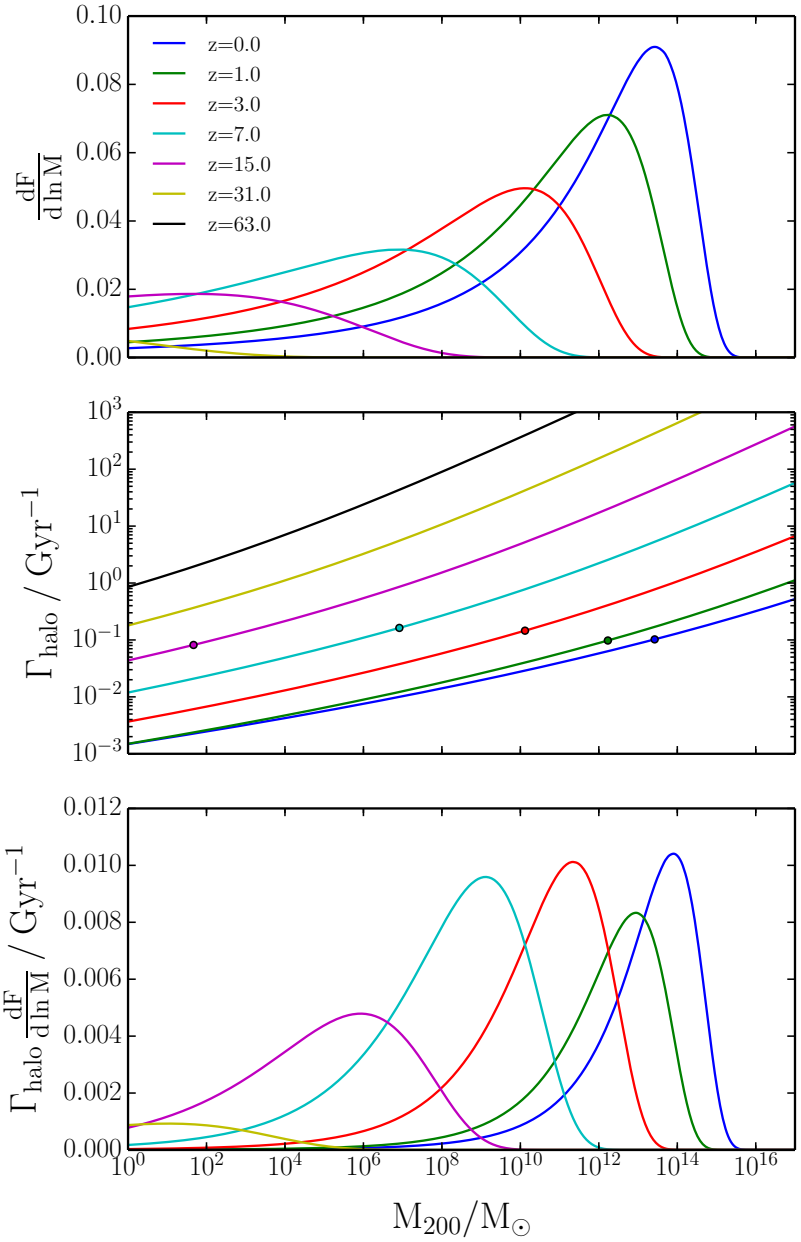


Figure 2.1: *Top panel*: The Multiplicity Function, which shows how mass in the Universe is split between objects of different mass, as described by Press-Schechter theory. Different coloured lines show different redshifts. *Middle panel*: The interaction rate per particle as a function of halo mass, assuming NFW density profiles and the Duffy et al. (2008) concentration-mass relation, with a velocity-independent cross-section of $1\,\mathrm{cm}^2\,\mathrm{g}^{-1}$. Circles highlight the mass at which the Multiplicity Function peaks, illustrating a relatively constant interaction rate per unit mass in the Universe's most typical haloes. *Bottom panel*: The product of curves in the two upper panels, illustrating the relative contribution of haloes in different logarithmic mass bins to the total interaction rate per particle. In this scenario, the main location for scatterings gradually transitions to more and more massive structures.

2.2.2 Interaction rates in collapsed structures

The scattering rate of an individual dark matter (DM) particle i, with velocity $\mathbf{v_i}$, is

$$\Gamma_i = \int f_v(\mathbf{v})\, \rho\, \frac{\sigma}{m}\, |\mathbf{v_i} - \mathbf{v}|\, \mathrm{d}^3\mathbf{v}, \tag{2.7}$$

where f_v is the velocity distribution function,[2] ρ the local density, and (σ/m) the cross-section for DM–DM scattering (which could depend on $|\mathbf{v_i} - \mathbf{v}| \equiv v_{\rm pair}$) divided by the DM particle mass. Integrating over the velocity distribution function gives the scattering rate for a particle at position $\mathbf{r}$,

$$\Gamma_i(\mathbf{r}) = \frac{\langle \sigma\, v_{\rm pair} \rangle(\mathbf{r}) \rho(\mathbf{r})}{m}. \tag{2.8}$$

For a halo of mass M containing N particles, the mean scattering rate per particle is

$$\Gamma_{\rm halo}(M) = \frac{1}{N} \sum_{i=1}^{N} \Gamma_i. \tag{2.9}$$

Integrating over radius r gives

$$\Gamma_{\rm halo}(M) = \frac{1}{M} \int_0^\infty 4\pi r^2 \rho(r) \Gamma_i(r)\, \mathrm{d}r \tag{2.10}$$

$$= \frac{1}{M} \int_0^\infty 4\pi r^2 \rho^2(r) \frac{\langle \sigma\, v_{\rm pair} \rangle(r)}{m}\, \mathrm{d}r. \tag{2.11}$$

We assume that the collapsed haloes from PS theory have spherically symmetric Navarro, Frenk & White (1997, hereafter NFW) density profiles,

$$\frac{\rho(r)}{\rho_{\rm crit}} = \frac{\delta_{\rm NFW}}{(r/r_s)(1 + r/r_s)^2}, \tag{2.12}$$

where r_s is a scale radius, $\delta_{\rm NFW}$ a dimensionless characteristic density, and $\rho_{\rm crit} = 3H^2/8\pi G$ is the critical density. We assume that the mass of a halo fills a spherical region of radius r_{200}, within which the mean density is $200\, \rho_{\rm crit}$ and the total mass is M_{200}. Outside this region we assume the density to be zero. For brevity we will also refer to r_{200} as $r_{\rm v}$ and M_{200} as M. The concentration parameter is defined as $c \equiv r_{200}/r_s$ and can be related to the characteristic density by

$$\delta_{\rm NFW} = \frac{200}{3} \frac{c^3}{\ln(1 + c) - c/(1 + c)}. \tag{2.13}$$

[2] Here f_v is normalised such that $\int f_v(\mathbf{v})\, \mathrm{d}^3\mathbf{v} = 1$.

Note that the NFW profile is obtained from non-interacting dark matter simulations. Dark matter scattering reduces the density in the centre of DM haloes, producing a constant density core (Burkert, 2000; Yoshida et al., 2000b; Spergel & Steinhardt, 2000; Colín et al., 2002; Rocha et al., 2013; Zavala et al., 2013). Assuming an NFW profile, the average radius at which interactions take place (assuming an isotropic velocity dispersion) is $0.32\,r_s$ independent of halo concentration. This is similar to the radius for which the radial density profiles seen in the simulations of Rocha et al. (2013) first drop below the NFW prediction. These simulations used the maximum allowed velocity-independent cross-section, and so cores in other models would likely be smaller. Also, while the density in the centres of haloes decreases, DM scattering increases the velocity dispersion in halo centres, which should cancel some of the effect. Nevertheless we acknowledge that these DM interactions are moderately self-regulating because they form cores that will tend to decrease the interaction rate, but proceed assuming an NFW profile for the DM density. If we relax this assumption then the scattering rates calculated would be lower, but a full treatment of the effect that scattering has on the phase space distribution of haloes, and so the subsequent scattering rates, requires full N-body simulations that are beyond the scope of this chapter.

To calculate the mean pairwise velocity of particles, we integrate over the velocity distribution functions of particle pairs. Assuming that their velocities are isotropic and follow a Maxwell-Boltzmann distribution[3] with one dimensional velocity dispersion $\sigma_{\rm 1D}$, this gives $\langle v_{\rm pair}\rangle = (4/\sqrt{\pi})\,\sigma_{\rm 1D}$. For an NFW halo, the velocity dispersion of particles is (Łokas & Mamon, 2001)

$$\begin{aligned}\sigma^2_{\rm 1D}(s,c) = \frac{1}{2}c^2 g(c) s (1+cs)^2 \frac{GM}{r_{\rm v}} \Bigg[\pi^2 - \ln(cs) - \frac{1}{cs} \\ - \frac{1}{(1+cs)^2} - \frac{6}{1+cs} + \left(1 + \frac{1}{c^2 s^2} - \frac{4}{cs} - \frac{2}{1+cs}\right) \\ \times\ \ln(1+cs) + 3\ln^2(1+cs) + 6{\rm Li}_2(-cs)\Bigg]\,,\end{aligned} \tag{2.14}$$

where $s \equiv r/r_{200}$, $g(c) \equiv [\ln(1+c) - c/(1+c)]^{-1}$, and ${\rm Li}_2(x)$ is the dilogarithm (commonly

[3]High resolution simulations of CDM report departures from Gaussianity for the distribution of velocity components along the principal axes of the velocity dispersion tensor (Vogelsberger et al., 2009), but this approximation is sufficient for our work.

referred to as Spence's function), defined by

$$\mathrm{Li}_2(x) = \int_x^0 \frac{\ln(1-u)}{u}\,\mathrm{d}u. \tag{2.15}$$

Returning to equation (2.11), and changing integration variable from r to s, we find

$$\Gamma_{\rm halo}(M, r_{\rm v}, c) = 16\sqrt{\pi}\frac{r_{\rm v}^3}{M}\frac{\sigma}{m}\int_0^1 s^2\rho^2(s,c)\,\sigma_{\rm 1D}(s,c)\,\mathrm{d}s, \tag{2.16}$$

where we have now assumed that the DM-DM cross-section is velocity-independent (this restriction is relaxed in §2.4). Both $\rho(s,c)$ and $\sigma_{\rm 1D}(s,c)$ depend on the virial mass and radius of a halo, and can be written as dimensionless functions of s and c multiplied by the dimensional quantities $M/r_{\rm v}^3$ and $\sqrt{GM/r_{\rm v}}$ respectively. We can then see that $\Gamma_{\rm halo}$ will be a function of the halo concentration scaled by power-laws in M and $r_{\rm v}$. Specifically, at fixed cross-section and halo concentration, $\Gamma_{\rm halo} \propto M^{3/2}\,r_{\rm v}^{-7/2}$.

At a particular cosmic time, $M \equiv M_{200}$ and $r_{\rm v} \equiv r_{200}$ are not independent, because $M_{200}/r_{200}^3 \propto \rho_{\rm crit}(z)$ by definition. Using this, we find $\Gamma_{\rm halo} \propto M^{1/3}\,\rho_{\rm crit}^{7/6}$, such that

$$\begin{aligned}\Gamma_{\rm halo}(M, z, (\sigma/m), c) = \Gamma_{\rm halo}(M_0, z_0, (\sigma/m)_0, c)\left(\frac{M}{M_0}\right)^{1/3} \\ \times \left(\frac{\rho_{\rm crit}(z)}{\rho_{\rm crit}(z_0)}\right)^{7/6}\left(\frac{(\sigma/m)}{(\sigma/m)_0}\right).\end{aligned} \tag{2.17}$$

We calculate $\Gamma_{\rm halo}(M_0, z_0, (\sigma/m)_0, c)$ with $M_0 = 10^{10}\,\mathrm{M}_\odot$, $z_0 = 0$ and $(\sigma/m)_0 = 1\,\mathrm{cm}^2\,\mathrm{g}^{-1}$, by numerically integrating equation (2.16). We can then calculate $\Gamma_{\rm halo}$ for haloes with different masses and at different redshifts using equation (2.17).

At fixed mass, redshift and cross-section, $\Gamma_{\rm halo}$ is found to increase significantly with increasing halo concentration. The logarithmic slope of the $\Gamma_{\rm halo}(c)$ relation is ~ 1.7 for $c = 5$, and ~ 2.5 for $c = 30$, with $\Gamma_{\rm halo} \propto c^2$ for concentrations around 10. As halo concentrations generally decrease with increasing halo mass, the mass dependence of $\Gamma_{\rm halo}$ is suppressed below the $\Gamma_{\rm halo} \propto M^{1/3}$ seen in equation (2.17). The overall form of $\Gamma_{\rm halo}(M, z)$ depends upon the concentration-mass-redshift relation. Following Duffy et al. (2008, hereafter D08), we shall initially assume

$$c(M, z) = 5.72\left(\frac{M}{10^{14}\,h^{-1}\mathrm{M}_\odot}\right)^{-0.081}(1+z)^{-0.71}. \tag{2.18}$$

Using this $c(M, z)$ relation we show $\Gamma_{\rm halo}(M, z)$ in the middle panel of Fig. 2.1. $\Gamma_{\rm halo}$ increases rapidly with increasing redshift at fixed mass, and increases with mass at fixed redshift. As objects grow in mass through cosmic time, the scattering rate in typical haloes at each redshift evolves slowly. Note that several more recent works show that this simple power law dependence of $c(M)$ should flatten at low masses, which we discuss in §2.3.1.

2.2.3 DM's cosmic scattering rate

Multiplying the multiplicity function from §2.2.1 by the interaction rate in individual haloes from §2.2.2 gives the contribution of haloes of different mass to the total rate of particle scattering in the Universe (see bottom panel of Fig. 2.1). Integrating this quantity over all halo masses at different redshifts yields the mean scattering rate of all particles at that redshift, $\Gamma(z)$, which we refer to as the 'Cosmic Scattering Rate'. This is plotted in Fig. 2.2, where it can be seen that after a gradual rise from the early universe to $z \approx 6$, $\Gamma(z)$ is constant to within a factor of two to the present day.

For this analysis, we assume that haloes form down to masses of $10^{-12}\,{\rm M}_\odot$. In the real Universe, self-interacting dark matter creates a small-scale cut-off in the power spectrum due to collisional damping. For DM composed of weakly interacting massive particles (e.g. neutralinos), the minimum mass of collapsed objects is $\sim 10^{-6}\,{\rm M}_\odot$ (Hofmann et al., 2001). If DM were axions then this minimum mass would be $\sim 10^{-12}\,{\rm M}_\odot$ (Kolb & Tkachev, 1996). For the general class of self-interacting dark matter models that lead to astrophysically interesting scattering rates in the late-time Universe, collisions in the early Universe suppress power on larger scales, or even introduce acoustic oscillations in the dark matter-dark radiation system (Buckley et al., 2014). There is a rich possible phenomenology affecting the high-k power spectrum, which is highly model-dependent.

We investigate the approximate effect of a cutoff in the power spectrum by integrating $\Gamma_{\rm halo}({\rm d}F/{\rm d}\ln M)$ down to different minimum masses, $M_{\rm min}$, shown as the extra lines in Fig. 2.2. Furthermore, in numerical simulations, only haloes above a given mass scale are resolved, and only the DM interactions above those scales can be tracked. We therefore include lines with large $M_{\rm min}$ in Fig. 2.2, to act as predictions for the expected scattering rate in cosmological simulations. Note that the results as $M_{\rm min} \to 0$ converge particularly

slowly for the D08 concentration-mass relation, due to the high concentration of very small haloes. Nevertheless, these results are less sensitive to changing $M_{\rm min}$ than those for a simple annihilation channel where $\sigma v_{\rm pair}$ is constant (Mack, 2014) and low mass haloes make a dominant contribution to the total scattering rate.

In addition to the rate of cosmic scattering, an interesting quantity is the mean cumulative number of interactions that particles have undergone. As each scattering event is a two-body interaction, this is twice the number of interactions per particle. We call this quantity $N_{\rm scatter}$ and plot it as a function of redshift in the bottom panel of Fig. 2.2. While the cosmic scattering rate is markedly different at intermediate and high redshifts when using different minimum masses, the values of $N_{\rm scatter}(z=0)$ are more robust. For $(\sigma/m) = 1\,{\rm cm}^2\,{\rm g}^{-1}$, $N_{\rm scatter}(z=0)$ is 0.87 with $M_{\rm min} = 10^{-12}\,{\rm M}_\odot$ and 0.68 with $M_{\rm min} = 10^{10}\,{\rm M}_\odot$.

A noticeable feature of $\Gamma(z)$ in the upper panel of Fig. 2.2 is the upturn after $z \approx 1$. This is not present when using more recent $c(M,z)$ relations with more complex redshift dependences than the simple $(1+z)^{-0.71}$ in the D08 relation. This upturn is not physical, and arises because the concentration is defined in terms of r_{200} which in turn depends on $\rho_{\rm crit}$. When the Universe is matter-dominated $\rho_{\rm crit} \propto (1+z)^3$, such that at fixed halo mass $r_{200} \propto (1+z)^{-1}$. At late times, when there is a significant dark energy contribution to the Universe, the evolution of $\rho_{\rm crit}$ slows and is no longer given by a simple power law in $(1+z)$. This affects the r_{200} of haloes, and hence halo concentrations, such that a simple power-law cannot accurately capture $c(M,z)$.

2.3 Sensitivity to astrophysical assumptions

In the previous section we considered the redshift dependence of DM scattering rates and showed that with a velocity-independent cross-section, the mean rate of particle scattering in the Universe initially grows and then starts to decrease after $z \approx 6$, dropping by less than a factor of two to the present day. In this section we explore the sensitivity of this result to the assumptions of the model.

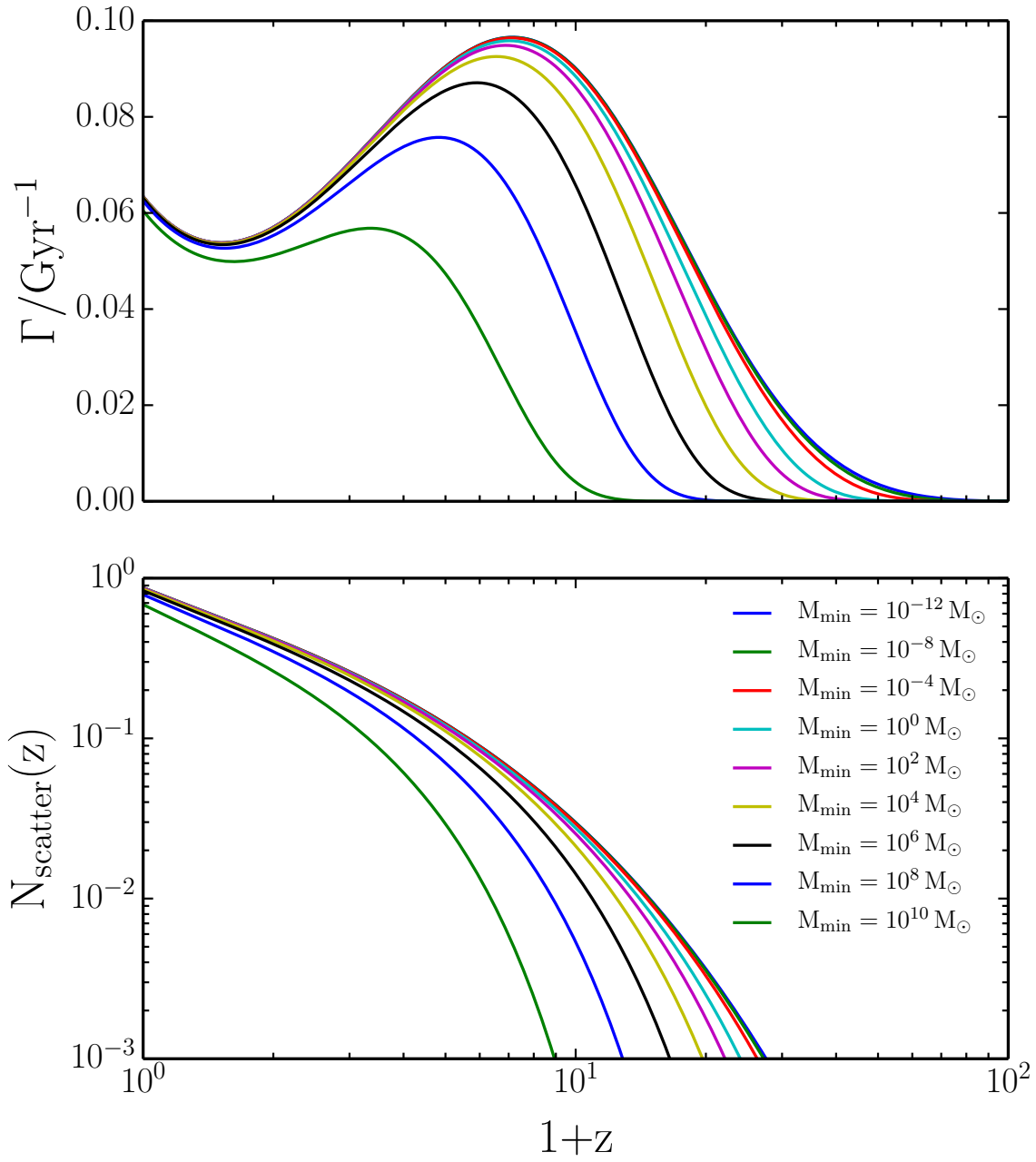


Figure 2.2: *Top panel*: The mean scattering rate of particles in the Universe calculated from Press-Schechter theory, assuming the NFW density profile, the D08 concentration-mass relation and $\sigma/m = 1\,\mathrm{cm}^2\,\mathrm{g}^{-1}$. The different lines count only scatterings in haloes more massive than $10^{10}\mathrm{M}_\odot$ (bottom line) to $10^{-12}\mathrm{M}_\odot$ (top line). The scattering rate varies by less than a factor of two from $z \approx 6$ onwards. *Bottom panel*: The mean cumulative number of interactions that particles have undergone as a function of redshift. The different lines again include only those interactions in haloes more massive than a given threshold. With a velocity-independent cross-section, most scattering is at late redshifts where there is more time. This results in most scattering being in high-mass haloes, so that $N_{\rm scatter}(z = 0)$ varies by less than 25% between $M_{\rm min} = 10^{-12}\mathrm{M}_\odot$ and $M_{\rm min} = 10^{10}\mathrm{M}_\odot$.

2.3.1 Concentration-mass-redshift relations

The concentration-mass-redshift relation, $c(M, z)$, of D08 is attractive for its simplicity and because over a small range of redshifts and halo masses, concentrations can be well fit by simple power laws in M and $(1 + z)$. However, numerical studies that have resolved structures over a wide range of halo masses have found that concentrations are not well fitted by simple power laws. Examining the results of the Millennium Simulation (Springel et al., 2005b) from $z = 3$ to $z = 0$ it is clear that the form of $c(M, z)$ is not separable, with the mass dependence evolving with redshift (Gao et al., 2008). This evolution takes the form of a flattening of the $c(M)$ relation at increasing redshift, such that concentrations of very massive galaxy cluster haloes evolve only weakly with redshift while the concentrations of smaller haloes decrease rapidly with increasing redshift.

The $c(M, z)$ relation is found to be remarkably complex, particularly when considering the dependence on cosmological parameters. Prada et al. (2012, hereafter P12) find that this complex relationship is a result of the 'wrong' physical quantities, M and z, being used. Analogous to studies of the halo mass function, in which a much simpler fitting formula is possible when one considers the mass function as a function of $\ln \sigma_M^{-1}$ rather than a function of M (Jenkins et al., 2001), the $c(\ln \sigma_M)$ relationship is found to be simpler than $c(M)$.

The behaviour of this relationship can be explained by models in which the concentration of a halo is related to its accretion history (Wechsler et al., 2002; Zhao et al., 2003). Ludlow et al. (2014, hereafter L14) found that if the mass of a halo, $M(z)$, was plotted against the critical density, $\rho_{\rm crit}(z)$, then the relationship $M(\rho_{\rm crit})$ was well fit by an NFW profile, with associated concentration $c_{\rm MAH}$. They also found a simple relation between $c_{\rm MAH}$ and the concentration of the halo, allowing the concentration-mass relation to be predicted from the mass-accretion history of haloes. The statistics of the mass-accretion of DM haloes can be found from simulations, or calculated using the conditional probabilities[4] found in extensions of PS theory (Bond et al., 1991; Bower, 1991; Lacey & Cole, 1993; Kauffmann & White, 1993).

Different methods for measuring $c(M, z)$, either from simulations or analytical calcu-

[4]The conditional probability that the material making up an object of mass M_1 at redshift z_1 is in an object of mass M_0 at redshift z_0.

lations, give similar results around the peak of the multiplicity function ($M \approx M^*$), but differ significantly at high and low masses. While the cosmic scattering rate is dominated by haloes around $M^*(z)$, the scattering rate in haloes is highly sensitive to the halo concentration, and so even small differences between $c(M, z)$ relations can lead to significant changes in $\Gamma(z)$. In Fig. 2.3 we show $\Gamma(z)$ calculated as in Fig. 2.2 but for a variety of $c(M, z)$ relations.

Noticeable in Fig. 2.3 is that using $c(M, z)$ from L14 gives a scattering rate at intermediate redshifts a factor of two above that found using other $c(M, z)$ relations. The L14 analytical model was calculated for relaxed haloes, which are generally dynamically older, making them more concentrated than unrelaxed haloes of a similar mass. The cuts made to remove unrelaxed haloes are one of the two main reasons why $c(M, z)$ relations from simulations disagree with each other, the other being the way in which c is calculated from a mass distribution. For example, Prada et al. (2012) calculate c from the ratio $V_{\rm max}/V_{200}$, where $V_{\rm max}$ and V_{200} are the maximum circular velocity and the circular velocity at r_{200} respectively, while Diemer & Kravtsov (2015) find c by directly fitting the radial density with an NFW profile.

2.3.2 Mass function prescription

It is known that the PS formula does not provide an exact fit to the mass function from simulations. Specifically, it underestimates the number of rare objects in the 'high-mass tail', with an overestimate of the amount of mass around the peak of the multiplicity function (see e.g. Jenkins et al., 2001). A better fit to the mass function from simulations was achieved by Sheth & Tormen (1999, hereafter ST), who found that compared to PS, equation (2.6) becomes:

$$\frac{\mathrm{d}F}{\mathrm{d}\,\ln M} = A\sqrt{\frac{2a}{\pi}}\left|\frac{\mathrm{d}\ln\sigma_M}{\mathrm{d}\ln M}\right|\left[1+\left(a\nu^2\right)^{-p}\right]\nu\exp\left(-\frac{a\nu^2}{2}\right), \tag{2.19}$$

with $A = 0.3222$, $a = 0.707$ and $p = 0.3$. We note that our definition of ν is different from that in the ST paper, with $\nu_{ST} = \nu^2$. The original PS formula can also be described by equation (2.19) with $A = 0.5$, $a = 1$ and $p = 0$.

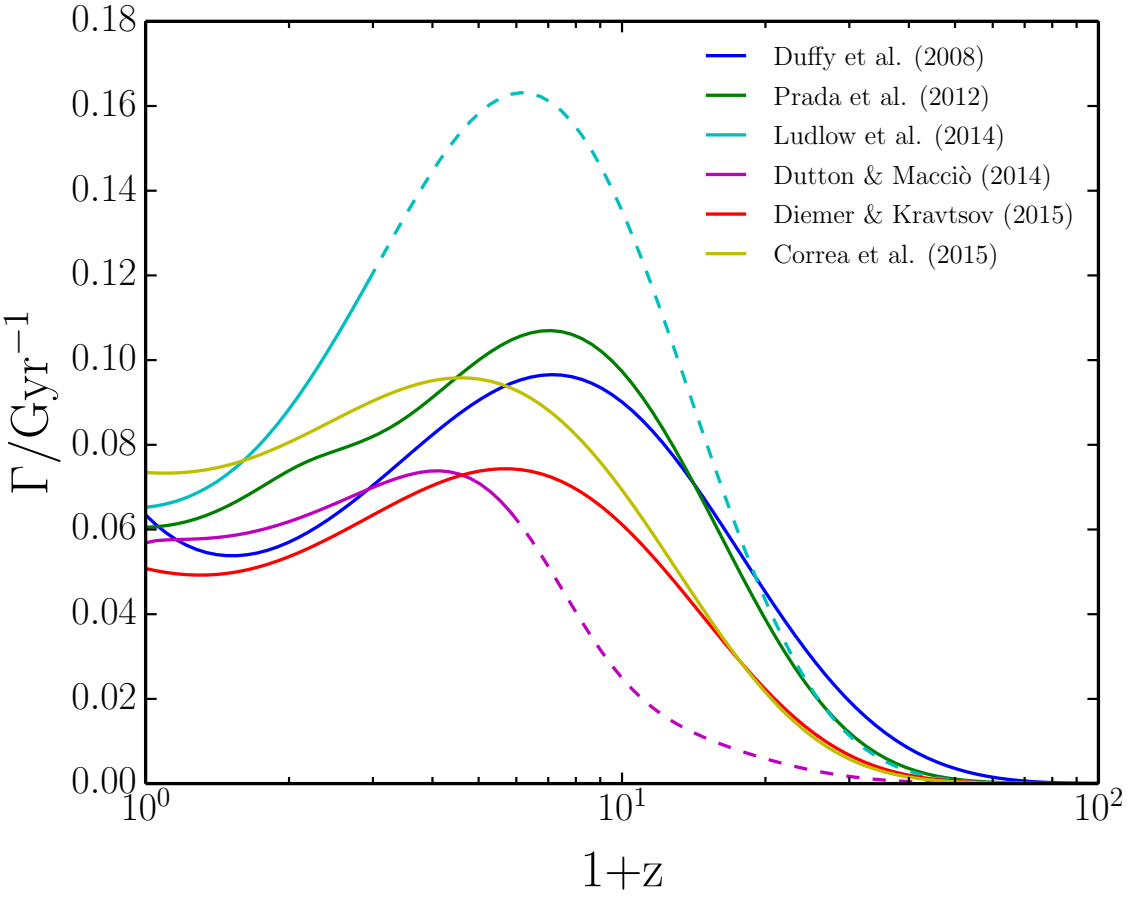


Figure 2.3: The cosmic scattering rate calculated using the concentration-mass-redshift relations of Duffy et al. (2008), Prada et al. (2012), Ludlow et al. (2014), Dutton & Macciò (2014), Diemer & Kravtsov (2015), and Correa et al. (2015). These were calculated assuming a Planck 2013 cosmology (Planck Collaboration et al., 2014), a PS mass-function, and $\sigma/m = 1\,\mathrm{cm}^2\,\mathrm{g}^{-1}$, counting the contribution from all haloes more massive than $10^{-12}\,\mathrm{M}_\odot$. Lines become dashed for redshifts where authors state their relationships may not be valid.

The ST mass function increases the number density of the most massive objects compared to the PS mass function, providing a better fit to simulations (see e.g. Reed et al., 2007). While these differences can be extremely important for some studies (e.g. counting the number density of massive clusters) we find that the different mass functions do not have a large effect on our results. This is because the scattering rate per unit mass in DM haloes increases only gently with increasing halo mass, as can be seen in the middle panel of Fig. 2.1. The shape of $\Gamma(z)$ is similar when either a PS or ST mass function is used, although the normalisation is slightly lower for the latter. By redshift zero there are $\sim 20\%$ fewer DM interactions with an ST mass function.

2.3.3 Varying cosmological parameters

Similar to changing the formalism used to calculate the multiplicity function, small changes to the Cosmological Parameters leave the cosmic scattering rate relatively unchanged because of the weak mass dependence of $\Gamma_{\rm halo}(M)$. Using $c(M,z)$ from D08, we find that changing cosmological parameters from Planck 2013 to WMAP9 decreases the mean number of interactions per particle by redshift zero, $N_{\rm scatter}(z=0)$, by 12%. This is driven by Planck's larger value for $\Omega_{\rm m}$, resulting in larger critical densities at early times. Using earlier WMAP results leads to similar changes, except for WMAP3 for which the anomalously low $\Omega_{\rm m}$ and σ_8 lead to a 33% reduction in $N_{\rm scatter}(z=0)$.

The concentration-mass-redshift relation also depends on cosmological parameters, which is made explicitly clear by relations that relate c to σ_M rather than M directly (e.g. Prada et al., 2012; Diemer & Kravtsov, 2015). This cosmology dependence of $c(M,z)$ makes little difference when moving from Planck 2013 to WMAP9, but further reduces the scattering rate for a WMAP3 cosmology such that $N_{\rm scatter}(z=0)$ is 40% lower than with a Planck 2013 cosmology, using $c(M,z)$ from P12. This increased difference, beyond that seen for a cosmology independent $c(M,z)$, can be understood by noting that haloes of a particular mass form later with smaller σ_8, and are therefore less concentrated.

2.3.4 Scatter in the concentration-mass relation

So far we have assumed that given the mass of a halo we know its concentration through the concentration-mass relation. In practice this relation has some scatter around it, which will impact on the mean scattering rate of haloes. From equation (2.17) the concentration dependence of the scattering rate in haloes is described by $\Gamma_{\rm halo}(M_0, z_0, (\sigma/m)_0, c)$. This is non-linear in c, such that even symmetric scatter in c at fixed mass will alter the mean scattering rate in haloes of that mass.

To discuss how $\Gamma_{\rm halo}$ is affected by scatter in c, it will be useful to introduce c_0, the value of c implied by the $c(M, z)$ relation. Dolag et al. (2004) find that for haloes of fixed mass and redshift, $\ln c$ is normally distributed. If we assume that $\ln c$ is normally distributed with mean $\ln c_0$ and variance $\sigma^2_{\ln c}$, then c follows a log-normal distribution, with probability density function

$$P(c) = \frac{1}{c\,\sigma_{\ln c}\sqrt{2\pi}} \exp\left(-\frac{(\ln c - \ln c_0)^2}{2\,\sigma^2_{\ln c}}\right). \tag{2.20}$$

Including a log-normal distribution of concentrations at fixed mass and redshift leads to an increase in Γ at all concentrations, related to the long tail of the distribution towards high values, as well as a shift in the expectation value of c.[5] If $c(M, z)$ in D08 was a measure of the mean c for a particular mass of halo, then we would have to make the change $\ln c_0 \to \ln c_0 - \sigma^2_{\ln c}/2$ in equation (2.20) to keep $\langle c \rangle = c_0$. However, the $c(M, z)$ relation in D08 was acquired by fitting to the median values of c in each mass bin at each redshift. The median value of c from the probability density function in equation (2.20) is simply $\exp(\ln c_0) = c_0$ as required. Dolag et al. (2004) found that $\sigma_{\ln c} \approx 0.22$, almost independent of the cosmological model. This corresponds to a standard deviation in $\log_{10} c$ of 0.1, or a scatter of 0.1 dex. We find that the shape of $\Gamma(z)$ is effectively unchanged by scatter in $c(M, z)$, but that the normalisation increases with increasing scatter. For a 0.1 dex scatter, the normalisation increases above that of the scatter-free case by less than 15%.

[5] For the distribution in equation (2.20), the expectation value of c is given by $\langle c \rangle = \exp\left(\ln c_0 + \sigma^2_{lc}/2\right) > c_0$.

2.4 Velocity-dependent cross-sections

Having calculated the rate of DM scattering through cosmic time assuming that the cross-section is velocity-independent, we now lift this assumption, and perform the same calculation with velocity-dependent DM-DM cross-sections.

2.4.1 Particle model

For velocity-dependent cross-sections we use the vdSIDMa and vdSIDMb models described in Vogelsberger & Zavala (2013). These are well-motivated by particle physics, and describe the transfer cross-section for elastic scattering mediated by a new gauge boson of mass m_ϕ. This results in an attractive Yukawa potential with coupling strength α_c. These interactions are analogous to the screened Coulomb scattering in a plasma, for which the momentum-transfer cross-section can be approximated by

$$\frac{\sigma_T}{\sigma_T^{\max}} \approx \begin{cases} \frac{4\pi}{22.7}\,\beta^2 \ln\left(1+\beta^{-1}\right), & \beta < 0.1 \\ \frac{8\pi}{22.7}\,\beta^2\,\left(1+1.5\beta^{1.65}\right)^{-1}, & 0.1 < \beta < 10^3 \\ \frac{\pi}{22.7}\,\left(\ln\beta + 1 - \frac{1}{2}(\ln\beta)^{-1}\right)^2, & \beta > 10^3, \end{cases} \tag{2.21}$$

where $\beta = \pi v_{\max}^2/v_{\text{pair}}^2$ and $\sigma_T^{\max} = 22.7/m_\phi^2$ (Feng et al., 2010; Finkbeiner et al., 2011; Loeb & Weiner, 2011). Here $v_{\max}$ is the velocity at which $(\sigma_T\, v_{\text{pair}})$ peaks, with $\sigma_T(v_{\max}) = \sigma_T^{\max}$. We have also introduced the "momentum-transfer cross-section", σ_T, defined as

$$\sigma_T = \int (1-\cos\theta)\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}(\theta)\,\mathrm{d}\Omega \tag{2.22}$$

$$= 2\pi \int_{-1}^{1} (1-\cos\theta)\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}(\theta)\,\mathrm{d}\cos\theta, \tag{2.23}$$

where $\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}$ is the differential cross-section, assumed to be azimuthally symmetric, which describes the probability of particles scattering into a patch of solid angle $\mathrm{d}\Omega$. The transfer cross-section is an effective scattering cross-section that is useful in describing angularly dependent cross-sections (where $\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}$ is not constant). For isotropic scattering ($\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \text{constant}$) the transfer cross-section is simply $\sigma_T = \sigma$, while in general the mean momentum transfer for a scattering process with transfer cross-section σ_T is equal to the mean momentum

transfer for isotropic scattering with $\sigma = \sigma_T$. Throughout the rest of this chapter, when calculating the rate and number of particle scattering events we will use σ_T as if it were the cross-section i.e we will calculate an effective rate of particle scatterings that is the rate of isotropic scattering events that would lead to the same rate of momentum transfer.[6]

The velocity-dependent cross-section in equation (2.21) leads to noticeable changes in $\Gamma_{\rm halo}(M)$. The cross-section diverges as the pairwise velocity tends to zero, such that scattering in low mass haloes (with typical velocities less than $v_{\rm max}$) is enhanced above the constant cross-section case. For $v_{\rm pair} \gg v_{\rm max}$, $\sigma_T \propto v_{\rm pair}^{-4}$, leading to a strong suppression of the scattering rate in DM haloes with velocity dispersions larger than $v_{\rm max}$.

The vdSIDMa and vdSIDMb models have values of $\sigma_T^{\rm max}/m$ and $v_{\rm max}$ chosen to maximise the self-interaction rate at the typical velocity dispersion of Milky Way dwarf spheroidals, while avoiding known astrophysical constraints on the cross-section. Specifically, vdSIDMa and vdSIDMb have $v_{\rm max} = 30\,{\rm km\,s^{-1}}$ and $\sigma_T^{\rm max}/m = 3.5\,{\rm cm^2\,g^{-1}}$, and $v_{\rm max} = 10\,{\rm km\,s^{-1}}$ and $\sigma_T^{\rm max}/m = 35\,{\rm cm^2\,g^{-1}}$ respectively.

2.4.2 vdSIDM cosmic scattering rates

The calculation of the DM scattering rate $\Gamma(z)$ proceeds in a similar manner to §2.2, in that we first find the distribution of haloes of different mass ($\frac{{\rm d}F}{{\rm d}\ln M}$) and then find the scattering rate per unit mass in these haloes, $\Gamma_{\rm halo}(M)$. However, the calculation of $\Gamma_{\rm halo}(M)$ is complicated by the velocity-dependent cross-section, because the cross-section can no longer be taken outside the integral in equation (2.11). Instead, we find $\langle \sigma\, v_{\rm pair}\rangle(r)$ by numerically integrating $\sigma_T(v_{\rm pair})\, v_{\rm pair}$ over the probability distribution of pairwise velocities, again assuming that the velocities of individual particles are drawn from a Maxwell-Boltzmann distribution function with 1D velocity dispersion $\sigma_{\rm 1D}$. This yields

$$\langle \sigma_T v\rangle\,(\sigma_{\rm 1D}) = \frac{1}{2\sigma_{\rm 1D}^3\sqrt{\pi}} \int \sigma_T(v) v^3 e^{-v^2/4\sigma_{\rm 1D}^2}\,{\rm d}v. \tag{2.24}$$

[6] In general, particle orbits within a DM halo are approximately isotropic, so there is no preferred direction for particle scattering. In these cases, the momentum transfer cross-section accurately captures the effects of scattering. However, this may not be the case for systems where there is a preferred direction along which particles approach (Kahlhoefer et al., 2014), such as in a merger. We demonstrate this explicitly in Chapter 6.

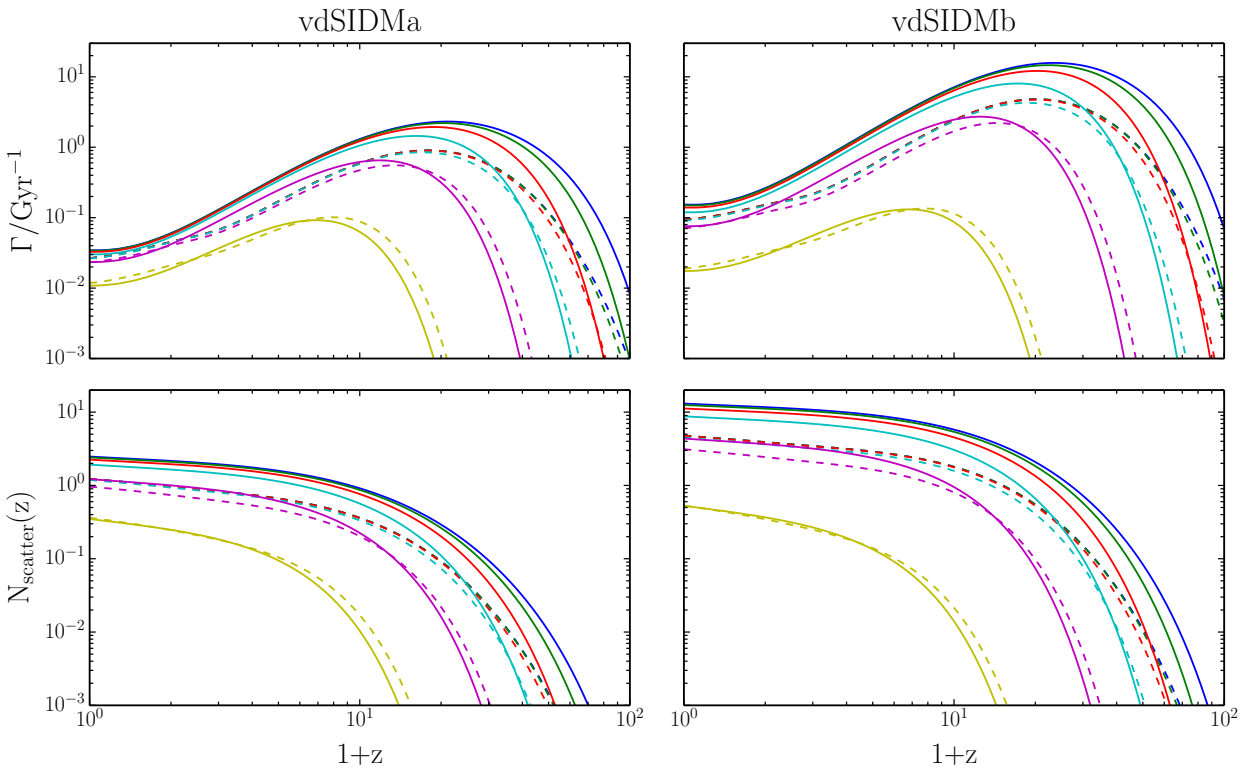


Figure 2.4: Scattering rates (top row) and cumulative number of scatters (bottom row) as a function of redshift, for two different velocity-dependent scattering cross-sections. The left column is for vdSIDMa which has $v_{\mathrm{max}} = 30\,\mathrm{km\,s^{-1}}$ and $\sigma_{\mathrm{max}}/m = 3.5\,\mathrm{cm^2\,g^{-1}}$; while vdSIDMb (right column) has $v_{\mathrm{max}} = 10\,\mathrm{km\,s^{-1}}$ and $\sigma_{\mathrm{max}}/m = 35\,\mathrm{cm^2\,g^{-1}}$. The different line colours correspond to different values for M_{min} of $10^8, 10^4, 1, 10^{-4}, 10^{-8}$, and $10^{-12}\,\mathrm{M_\odot}$, with both Γ and N_{scatter} monotonically increasing with decreasing M_{min}. The solid lines are for the D08 concentration-mass-redshift relation, while the dashed lines use the P12 $c(M, z)$. Unlike the constant cross-section case in Fig. 2.2, $\Gamma(z)$ is now plotted on a logarithmic scale as the scattering rate is larger by around two orders of magnitude at high redshift compared to redshift zero.

Then with $\sigma_{\mathrm{1D}}(r)$ from equation (2.14) we can find $\langle \sigma_T v \rangle\,(r)$, which we use in the numerical evaluation of equation (2.11) to calculate $\Gamma_{\mathrm{halo}}(M)$. Combining $\Gamma_{\mathrm{halo}}(M)$ with the multiplicity function we can calculate the cosmic scattering rate as in §2.2.3.

The scattering rate through cosmic time is plotted for vdSIDMa and vdSIDMb in Fig. 2.4. In contrast to the velocity-independent case in Fig. 2.2, the scattering rate is now displayed on a logarithmic scale. It peaks at redshift $20 - 30$ and falls by two orders of magnitude

before $z = 0$. Most interactions thus occur at early times as can be seen in Fig. 2.5. Half occur before $z = 5.7$ for vdSIDMa and $z = 7.2$ for vdSIDMb (in the latter case, the Universe is $\sim 5\%$ of its present age). This is in stark contrast to the gentler evolution of $\Gamma(z)$ with a constant cross-section (c.f. Fig. 2.2), where half the interactions occur after $z = 0.96$.

With a velocity-dependent cross-section, most scatterings also occur in low mass haloes with typical velocities $v \lesssim v_{\rm max}$. Raising the minimum mass of considered haloes $M_{\rm min}$ from $10^{-12}\,{\rm M}_\odot$ to $10^{8}\,{\rm M}_\odot$ lowers the number of interactions by redshift zero by a factor of six, which can be seen in Fig. 2.6 (for which we introduce $N_0 \equiv N_{\rm scatter}(z = 0)$). For the constant cross-section case, the same change leads to a decrease in $N_{\rm scatter}(z = 0)$ of only 10%.

The choice of concentration-mass-redshift relation becomes more important when the cross-section is velocity-dependent, because different $c(M, z)$ disagree most for low mass haloes and at high redshift. In particular, the simple power law relation from D08 predicts low-mass haloes to be much more concentrated than more recent relations in which $c(M)$ flattens at low mass. This recovers (a less extreme version of) what is seen in estimates of the DM annihilation rate, where $\langle \sigma\, v_{\rm pair} \rangle$ is usually assumed to be constant, resulting in an even larger fraction of interactions occurring in low mass haloes, and hence a cosmic scattering rate with strong dependence on $c(M, z)$ (Mack, 2014; Correa et al., 2015).

As well as the three particle models already discussed (velocity-independent, vdSIDMa and vdSIDMb), we include in Fig.s 2.5 and 2.6 plausible but more extreme velocity-dependent models with lower $v_{\rm max}$. We need not specify the normalisation of $\sigma_T^{\rm max}/m$ for these calculations, but it can be chosen to solve small-scale problems at dwarf galaxy scales, while eluding constraints at cluster scales. As $v_{\rm max}$ is lowered, a larger fraction of interactions happen at high redshift and in low-mass haloes. For the most extreme case considered, with $v_{\rm max} = 10^{-3}\,{\rm km\,s^{-1}}$, half of the interactions have occurred by $z = 19$, and half occur in haloes of mass $< 10^{-6}\,{\rm M}_\odot$. We stress that such models cannot be excluded on particle physics grounds, but it is unclear whether the large number of scatterings in such low mass haloes would leave a detectable signal in the present day universe.

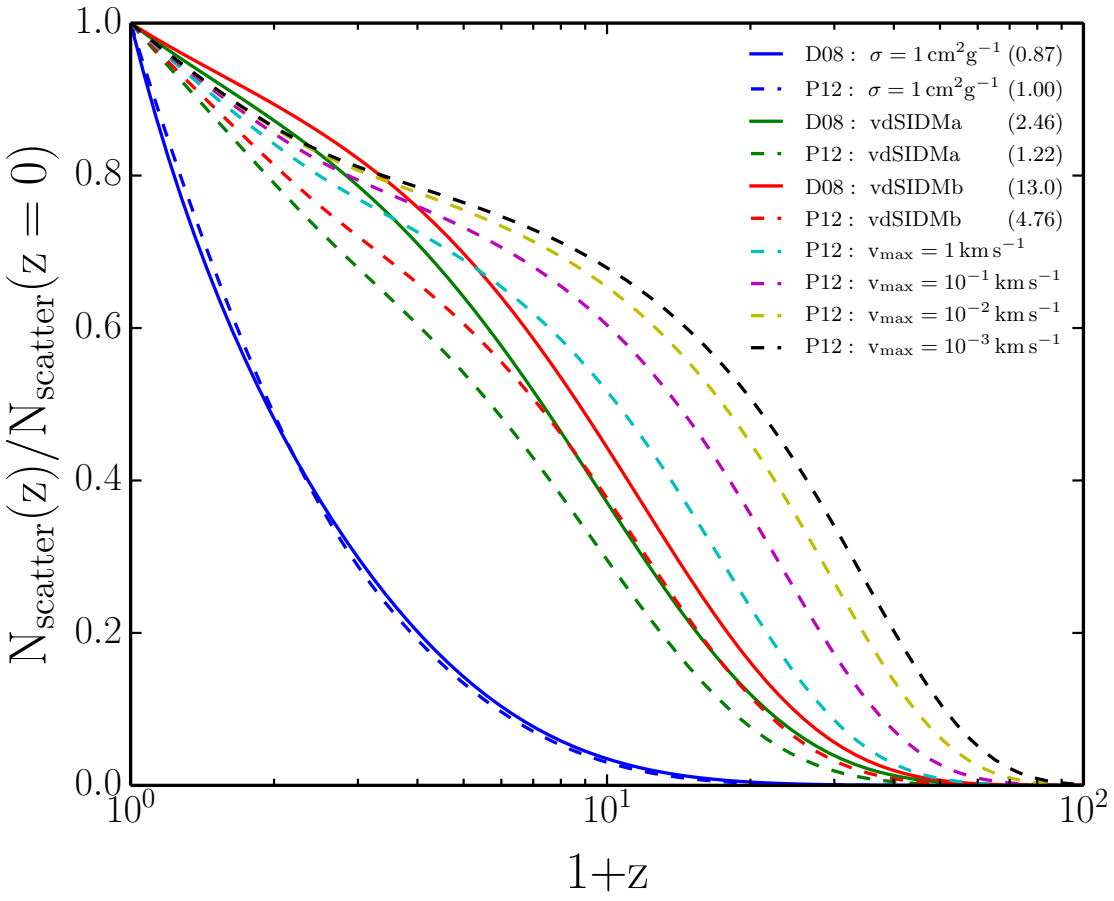


Figure 2.5: *When do scatterings happen?* The cumulative number of interactions as a function of redshift, normalised to unity at redshift zero. The different colours correspond to different particle models for the DM, while the solid and dashed lines are for the D08 and P12 $c(M, z)$ relations respectively. All curves were calculated using $M_{\rm min} = 10^{-12}\,{\rm M}_\odot$. The number in brackets in the legend is $N_{\rm scatter}(z = 0)$ for the relevant model. These are not present for the models with specified $v_{\rm max}$, which represent vdSIDM models with unspecified $\sigma_T^{\rm max}$. Velocity-dependent models with low $v_{\rm max}$ lead to more interactions in haloes with low internal velocities, pushing scattering towards high redshifts where collapsed objects are less massive.

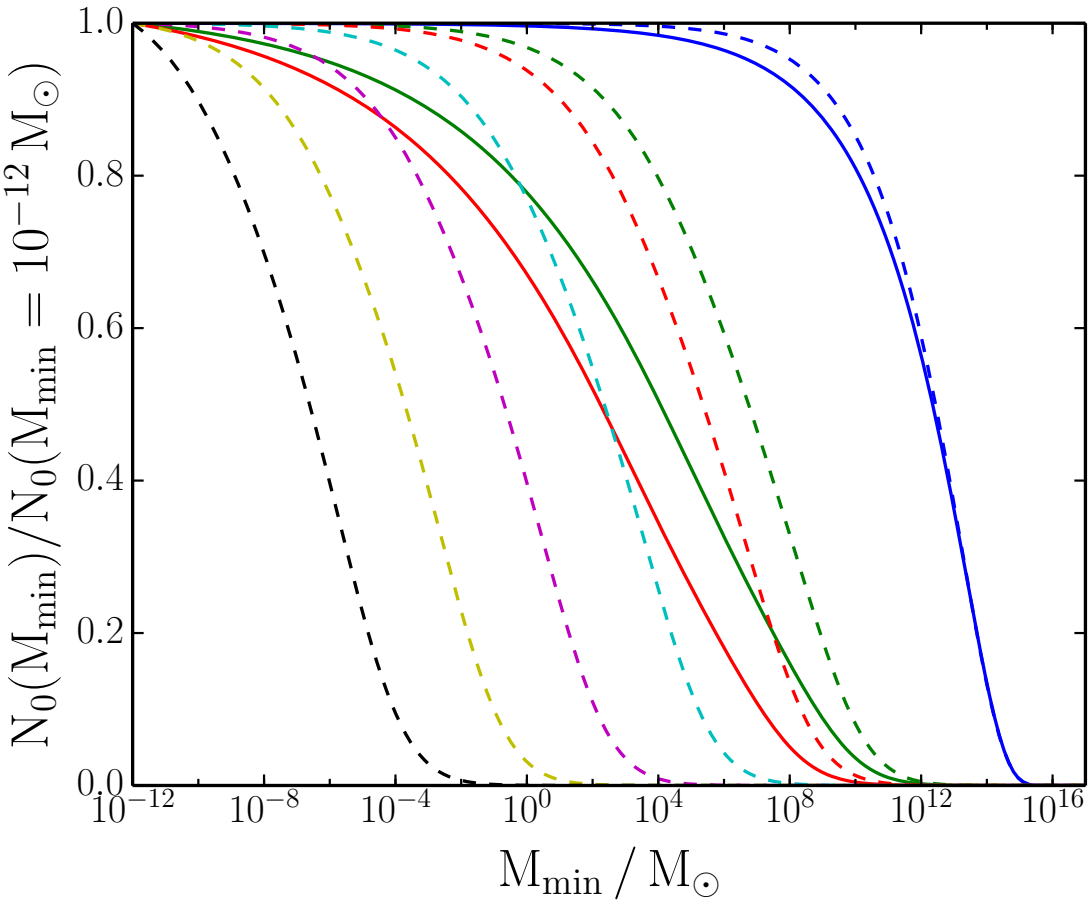


Figure 2.6: *Where do scatterings happen?* The fraction of scatterings by redshift zero that occur in haloes more massive than $M_{\rm min}$, normalised to unity for $M_{\rm min} = 10^{-12}\,{\rm M}_\odot$. Different line styles are as in Fig. 2.5, with colours corresponding to particle models, and solid or dashed lines representing the D08 or P12 concentration-mass-redshift relations respectively. Models with velocity-independent cross-sections have more of their scatterings in high-mass haloes compared to velocity-dependent cases, where the typical halo mass in which most interactions happen is an increasing function of $v_{\rm max}$.

2.5 Conclusions

We have presented an analytical calculation of the mean rate of DM–DM scattering events, for particle physics models with a velocity-independent or velocity-dependent cross-section. In all our calculations, we assume that the self-interactions are a small perturbation to ΛCDM and do not, for example, change the overall growth of structure.

For particle physics models with a velocity-independent interaction cross-section, our results match the canonical picture in which most scatterings occur in massive structures $\gtrsim 10^{12}\,\mathrm{M}_\odot$ at late times $z \lesssim 1$. Our calculations are found to be robust to current uncertainties in cosmological parameters as well as variations in the mass function used. They are also insensitive to the high-k power spectrum (because most scattering events occur in haloes more massive than the cut-off scales due to DM self-interactions in the early universe). The main source of uncertainty in the results is the concentration-mass-redshift relation $c(M, z)$. Its unknown form at high redshift and low mass propagates into a factor of almost three discrepancy in the scattering rate at intermediate redshifts ($z \approx 10$). However, the scattering rate changes by only a factor of two over most of cosmic time, and different $c(M, z)$ relations give similar results after $z \approx 1$, where there is more time. Consequently, the total number of interactions during the entire history of the Universe is uncertain to only a factor of ~ 2.

For particle physics models with a well-motivated velocity dependence, the scattering takes place mainly in low mass objects $\lesssim 10^4\,\mathrm{M}_\odot$ at early times $z \gtrsim 7$. The scattering rate $\Gamma(z)$ peaks at earlier redshifts $z \sim 20$, with a pronounced peak two orders of magnitude higher than the scattering rate at the present day. These numbers are more sensitive to the choice of cosmological and astrophysical parameters, and are dominated by regimes in which the mass function and concentration-mass-redshift relation are least well known. This minimum mass of considered structures, M_{min}, is particularly important, with changes to the small-scale power spectrum induced by DM scattering affecting the cosmic scattering rate.

The dominance of high redshift scatterings in velocity-dependent models profoundly changes their influence on the evolution of structure, and may alter the best strategy to

search for observational signatures. DM particle interactions lead to a transport of particles away from the centres of structures (Kochanek & White, 2000), replacing the cusps found in collisionless CDM simulations with constant density cores.[7] If the SIDM interactions are effectively confined to high redshift, then they may lead to a smearing of small-scale structure more qualitatively reminiscent of warm dark matter. The affected DM structures are also the hosts of the first galaxies, and it is interesting to consider what impact cored haloes could have on early galaxy formation.

High redshift scattering in low-mass objects also has important consequences for attempts to simulate vdSIDM cosmologies. Most scatterings occur in low mass haloes at high redshift that would not be resolved in typical cosmological simulations, but the unresolved interactions could be important for the later dynamics of particles. The large number of self-interactions would create DM cores in high-redshift haloes, and it then becomes an important question – on which there seems little consensus – whether or not the mergers of small cored haloes form cores that persist in large haloes at the present day.

[7] Haloes in which there have been a larger number of interactions presumably have larger cores. However, we caution against qualitative attempts to determine the scattering rate from core sizes: estimates of the core size and ellipticity of a galaxy cluster halo (Miralda-Escudé, 2002) overestimated the effect of SIDM by a factor 50 compared with the results from N-body simulations (Peter et al., 2013).

Chapter 3

Simulating self-interacting dark matter

3.1 An Introduction to N-body techniques

In this chapter we discuss our implementation of DM–DM scattering within the N-body code GADGET (Springel, 2005). Before this, however, we introduce the N-body method as a means to solve for the evolution of a self-gravitating system, assuming initially that DM particles interact only through gravity.

3.1.1 The N-body method

In its simplest incarnation, an N-body code simulates the evolution of N gravitationally interacting bodies (e.g. stars orbiting in a globular cluster) by directly calculating the gravitational force on each body from the $N-1$ other bodies. Such a system can be simulated by representing each body by a particle, described by its mass, position and velocity. The simulation proceeds by making small jumps in time, a 'time-step', during which the velocity of a particle is updated based on the acceleration due to gravitational forces, and its position updated based on its velocity.

For a globular cluster this involves tracking $\sim 10^5$ 'particles', but assuming the DM particle mass to be $100\,\mathrm{GeV}$, a Milky-Way like DM halo contains $\sim 10^{67}$ DM particles. Cosmological simulations therefore cannot come close to simulating individual DM particles, and instead rely on simulating a set of tracer particles sampled from the distribution func-

tion of DM particles. Before going on to discuss the distribution function and the equations that govern its time-evolution, we first discuss two-body interactions and when they are important in self-gravitating systems. Two-body interactions are close encounters between pairs of particles that alter the dynamics of a system compared with particles orbiting in a smooth potential. The importance of these effects depends on the number of particles in a system, so if we are going to use a system with 10^6–10^9 particles to simulate a physical system with 10^{67} particles, we need to be careful not to spuriously introduce these effects.

3.1.2 Two-body interactions

In Appendix B we discuss scattering in a Coulomb potential. The radial dependence of gravitational forces is the same as that for electrostatic forces, and so the equations derived for a Coulomb potential apply equally to two particles that gravitationally scatter. Comparing the gravitational force between two bodies

$$F_{\rm grav} = \frac{Gm^2}{r^2}, \tag{3.1}$$

with the Coulomb force (equation B.1), we can transform equation (B.5) into its gravitational equivalent

$$b = \frac{Gm}{v^2 \tan\frac{\theta}{2}}, \tag{3.2}$$

where b is the impact parameter,[1] m the mass of the particles, and v their initial relative velocity. For small scattering angles, this gives

$$\theta \approx \frac{2Gm}{b\,v^2}. \tag{3.3}$$

If we assume that a particle moves at a speed v through a sea of stationary particles, each with a mass m and number density n, then we expect it to encounter

$$\Gamma_b\,\mathrm{d}b = 2\pi\,b\,\mathrm{d}b\,v\,n \tag{3.4}$$

particles with impact parameters between b and $b + \mathrm{d}b$, per unit time. Over time, these encounters alter the trajectory of the particle under consideration in the manner of a random

[1] For a definition of the impact parameter see Fig. B.1

walk, so that the mean square total deflection angle is found by summing the squares of the individual deflection angles

$$\left\langle \theta_{\rm tot}^2 \right\rangle = \theta_1^2 + \theta_2^2 + \theta_3^2 + ... \tag{3.5}$$

Assuming many small kicks we can write this as an integral

$$\left\langle \theta_{\rm tot}^2 \right\rangle (t) = \int_0^t \int_{b_{\rm min}}^{b_{\rm max}} \Gamma_b \, \theta^2(b) \, {\rm d}b \, {\rm d}t = \frac{8\pi G^2 m^2 n \, t}{v^3} \ln \frac{b_{\rm max}}{b_{\rm min}}, \tag{3.6}$$

where $b_{\rm min}$ and $b_{\rm max}$ are the minimum and maximum impact parameters of encounters, and $\ln \frac{b_{\rm max}}{b_{\rm min}}$ is known as the Coulomb logarithm. In order to calculate the timescale over which 2-body encounters significantly alter the path of our particle, we set $\left\langle \theta_{\rm tot}^2 \right\rangle (t_{\rm relax}) = 1$, from which we find the relaxation time

$$t_{\rm relax} = \frac{v^3}{8\pi G^2 m^2 n \ln \frac{b_{\rm max}}{b_{\rm min}}}. \tag{3.7}$$

The values of $b_{\rm min}$ and $b_{\rm max}$ are not well defined, though thankfully the result is only logarithmically dependent upon them. $b_{\rm max}$ is typically taken to be the characteristic size of the system under study, for example the radius of a globular cluster, while a sensible choice for $b_{\rm min}$ is the 'strong encounter radius'. This is the radius at which the magnitude of the gravitational potential energy of two particles is equal to the initial kinetic energy

$$\frac{G m^2}{r_{\rm strong}} = \frac{1}{2} m \, v^2. \tag{3.8}$$

For pairs of particles with $b \lesssim r_{\rm strong}$ we expect a large change in the velocities of the particles, such that our previous assumption of many small deflections would no longer be valid. We can quantify the effects of these strong encounters by noting that given the definition of $r_{\rm strong}$, we expect the typical time between strong encounters to be

$$t_{\rm strong} = \frac{1}{\pi r_{\rm strong}^2 v \, n} = \frac{v^3}{4\pi G^2 m^2 n} = 2 \ln \frac{b_{\rm max}}{b_{\rm min}} \, t_{\rm relax}. \tag{3.9}$$

Frequent distant encounters are therefore more efficient at altering a particle's trajectory than strong encounters so long as $b_{\rm max} \gg b_{\rm min}$. $t_{\rm relax}$ is therefore the relevant timescale on which two-body effects become important, with strong encounters only contributing as a secondary effect.

Returning to equation (3.7), and considering a uniform density sphere with radius R and total mass M, we can progress by noting that $m\,n$ is simply the density $\rho = M/\frac{4}{3}\pi R^3$. If we also set $b_{\rm max} = R$ and $b_{\rm min} = r_{\rm strong} = 2Gm/v^2$, then the relaxation time is

$$t_{\rm relax} = \frac{v^3 R^3}{6G^2 M m \ln \frac{Rv^2}{2GM}}. \tag{3.10}$$

The virial theorem relating the kinetic and potential energies of a stable system of particles interacting through gravitational forces states that

$$2E_{\rm kin} + E_{\rm pot} = 0. \tag{3.11}$$

For a uniform density sphere, the binding energy is

$$E_{\rm pot} = -\frac{3GM^2}{5R} \tag{3.12}$$

and the kinetic energy is

$$E_{\rm kin} = \frac{1}{2}Mv^2, \tag{3.13}$$

which using the virial theorem implies

$$v^2 = \frac{3GM}{5R}. \tag{3.14}$$

Using this in equation (3.10),

$$t_{\rm relax} = \frac{3}{50}\frac{MR}{mv\ln\left(\frac{3}{10}\frac{M}{m}\right)}. \tag{3.15}$$

Then, noting that M/m is just the number of particles in the system, N, and dropping the numerical pre-factors which are only approximate,

$$t_{\rm relax} \sim \frac{N}{\ln N}\frac{R}{v}. \tag{3.16}$$

R/v is often referred to as the crossing time, and is comparable to the time it takes a particle to orbit within a system. Two-body relaxation is therefore only important over timescales $\sim N/\ln N$ crossing times. For stars within globular clusters, or galaxies within a galaxy cluster, $t_{\rm relax}$ may be less than the age of the Universe, and two-body effects can play an important role. For DM within a galaxy on the other hand, $t_{\rm relax}$ is many orders of magnitude larger than the age of the Universe. The DM therefore behaves as a collisionless fluid, evolving under the influence of a smooth gravitational potential that is generated by its mass distribution.

3.1.3 The Collisionless Boltzmann Equation

We have just seen that for DM particles, two-body gravitational interactions are negligible. DM particles therefore move smoothly through phase space, with their distribution in physical-space and velocity-space defined by the distribution function $f(\mathbf{r}, \mathbf{v}, t)$, with

$$\mathrm{d}M = f(\mathbf{r}, \mathbf{v}, t)\, \mathrm{d}^3\mathbf{r}\, \mathrm{d}^3\mathbf{v} \tag{3.17}$$

the mass of DM in the volume element $\mathrm{d}^3\mathbf{r}$ centred on $\mathbf{r}$, with velocity in the velocity-space element $\mathrm{d}^3\mathbf{v}$ centred on $\mathbf{v}$.

In the absence of collisions, Liouville's theorem states that

$$\frac{\mathrm{d}f}{\mathrm{d}t} = 0. \tag{3.18}$$

The derivative here is the Lagrangian derivative, and Liouville's theorem is the statement that the six dimensional phase-space density about a given system point (i.e. a DM particle) traveling through phase-space is constant with time. The Lagrangian derivative of f can also be written as

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \frac{\partial f}{\partial t} + \frac{\partial f}{\partial \mathbf{r}} \cdot \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} + \frac{\partial f}{\partial \mathbf{v}} \cdot \frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t}, \tag{3.19}$$

which combined with equation (3.18) and using the fact that $\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t}$ is the velocity, $\mathbf{v}$, and that changes in velocity result from the gravitational force per unit mass, $-\nabla\Phi$, gives

$$\frac{\partial f}{\partial t} = -\mathbf{v} \cdot \nabla f + \nabla\Phi \cdot \frac{\partial f}{\partial \mathbf{v}}. \tag{3.20}$$

This is known as the Collisionless Boltzmann equation, or Vlasov equation, which coupled with the Poisson equation

$$\nabla^2\Phi = 4\pi G\rho = 4\pi G \int f\, \mathrm{d}^3\mathbf{v} \tag{3.21}$$

describes the evolution of a collisionless self-gravitating system.

Equations (3.20) and (3.21) completely describe the evolution of the DM fluid[2] and we could calculate the evolution of f on a six-dimensional phase-space grid. However, DM that is 'cold' lives on a thin sheet in phase-space, such that this method would be highly

[2]For now we are assuming that the DM is the only species, and thus makes the only contribution to the gravitational potential.

inefficient. Instead, we turn to N-body simulations, a Monte Carlo technique where we replace the smooth distribution function $f(\mathbf{r}, \mathbf{v}, t)$ with N particles, each described by a mass, position and velocity, such that

$$f(\mathbf{r}, \mathbf{v}, t) \approx \sum_{i=1}^{N} m_i \delta(\mathbf{r} - \mathbf{r_i}(t)) \delta(\mathbf{v} - \mathbf{v_i}(t)). \tag{3.22}$$

More precisely, over any phase-space volume V

$$\int_V f(\mathbf{r}, \mathbf{v}, t)\, \mathrm{d}^3\mathbf{r}\, \mathrm{d}^3\mathbf{v} = \left\langle \sum_{(\mathbf{r_i}, \mathbf{v_i}) \in V} m_i \right\rangle. \tag{3.23}$$

This requires that the initial $\mathbf{r_i}$ and $\mathbf{v_i}$ are selected with probability proportional to $f(\mathbf{r}, \mathbf{v}, t_0)$,[3] where t_0 is the starting time of our simulation.

With an initial set of particles with phase-space coordinates that evenly sample $f(\mathbf{r}, \mathbf{v}, t_0)$, the phase-space density at a later time, $f(\mathbf{r}, \mathbf{v}, t)$, can be approximated by the distribution of particles at this time, if we move the particles through phase space according to

$$\frac{\mathrm{d}\mathbf{r_i}}{\mathrm{d}t} = \mathbf{v_i} \quad \text{and} \quad \frac{\mathrm{d}\mathbf{v_i}}{\mathrm{d}t} = -\nabla\Phi(\mathbf{r_i}). \tag{3.24}$$

We estimate the potential from the masses and positions of the N particles through the Poisson equation, noting that the linearity of this equation means that the potential at $\mathbf{r_i}$ is simply the sum of the potential at $\mathbf{r_i}$ from all other particles in the system

$$\Phi(\mathbf{r_i}) = -\sum_{j \neq i} \frac{G m_j}{|\mathbf{r_i} - \mathbf{r_j}|}. \tag{3.25}$$

This potential (assuming δ-function point mass particles) diverges as particles approach one another, leading to large gradients in Φ, hence large accelerations. These require careful integration with small time-steps to produce the correct evolution of $\mathbf{v_i}(t)$. As discussed in §3.1.2, close encounters should not be important for the evolution of collisionless DM, and so these large accelerations in close encounters are entirely spurious, a result of the small number of particles used in simulations compared to the number of DM particles in the underlying physical systems that is being simulated. It is therefore common practice to 'soften' the gravitational potential, such that the force between two particles falls off as the separation tends to zero.

[3] Assuming that all simulation particles have the same mass.

3.1.4 Gravitational softening

A common example of this softening procedure is so-called Plummer softening, where the potential becomes

$$\Phi(\mathbf{r_i}) = -\sum_{j \neq i} \frac{Gm_j}{\left(|\mathbf{r_i} - \mathbf{r_j}|^2 + \epsilon^2\right)^{1/2}}, \tag{3.26}$$

with ϵ the 'gravitational softening length'. Equation (3.26) corresponds to the potential if each particle had a Plummer density profile (Plummer, 1911)

$$\rho_P(r) = \left(\frac{3m}{4\pi\epsilon^3}\right)\left(1 + \frac{r^2}{\epsilon^2}\right)^{-\frac{5}{2}}. \tag{3.27}$$

However, softening and treating each particle as an extended mass distribution are not entirely equivalent as the force acting on a particle at $\mathbf{r_i}$ is calculated from the gradient of the potential there, not the force that would be exerted on an extended density distribution centred there.

Softening reduces the forces in close encounters, allowing for the use of longer time-steps than for an unsoftened potential. It also provides a better estimate of the forces that would be present if there were a smooth matter distribution rather than N δ-function particles. Another reason often cited for including softening is to reduce two body relaxation. As discussed in Dehnen (2001) and found empirically by (Hernquist & Barnes, 1990), relaxation effects are not suppressed much by softening. This follows from the Coulomb logarithm term in the relaxation time (equation 3.7); relaxation results from both close and near encounters, with each decade in impact parameter contributing equally. Reducing the effects of close encounters therefore cannot prevent relaxation from occurring, which depends predominantly upon the number of particles used in the simulation.

3.1.5 The GADGET N-body code

Having introduced the theoretical background to N-body techniques, we now remark upon how the particle distribution is evolved in practice in the collisionless simulations used in cosmology, in particular in the GADGET code used throughout this thesis.

Gravitational softening in GADGET

We start by discussing gravitational softening, for which Plummer softening is no longer popular. This is because the force due to a Plummer sphere is everywhere less than that due to a point particle, which means that the expected value of the forces in a Plummer softened N-body system systematically deviate from the forces in the underlying system. This force bias can be reduced by using a compact kernel, i.e. one for which the force returns to that for a point particle outside of some radius. The kernel used in GADGET is the spline kernel introduced by Monaghan & Lattanzio (1985) in the context of smoothed-particle hydrodynamics (SPH). This corresponds to a density distribution for each particle

$$\rho_{\rm s}(r,\tilde{\epsilon}) = \frac{8m}{\pi\tilde{\epsilon}^3}\begin{cases} 1-6\left(\frac{r}{\tilde{\epsilon}}\right)^2+6\left(\frac{r}{\tilde{\epsilon}}\right)^3, & 0\leq\frac{r}{\tilde{\epsilon}}\leq\frac{1}{2}, \\ 2\left(1-\frac{r}{\tilde{\epsilon}}\right)^3, & \frac{1}{2}<\frac{r}{\tilde{\epsilon}}\leq 1, \\ 0, & \frac{r}{\tilde{\epsilon}}>1. \end{cases} \tag{3.28}$$

For this kernel, the forces between particles separated by more than $\tilde{\epsilon}$ are the same as those for point particles. In GADGET, the gravitational softening length ϵ, is defined by $\tilde{\epsilon}=2.8\epsilon$. This definition stems from Plummer softening described in §3.1.4. With $\tilde{\epsilon}=2.8\epsilon$, the gravitational potential at zero radius is $-Gm/\epsilon$, the same as for a Plummer sphere (equation 3.27).

Efficient gravitational force calculations

A direct summation of the forces acting on all particle pairs in a simulation would involve $\frac{1}{2}N(N-1)$ individual force calculations. For simulations that can have upwards of a billion particles, this would be completely infeasible in a sensible amount of time and so alternatives to direct summation are used. Two popular solutions to this problem are to use a *tree* or a *particle mesh*. GADGET uses both, so we describe them both here.

Tree: The hierarchical tree algorithm (Barnes & Hut, 1986) works by dividing the simulation box up into sub-regions and dividing these sub-regions into further sub-regions until each region contains either one or no particles. In GADGET this division is done using an octree, where each cubic region is divided into eight equally sized octants. When calculating the forces on a particle, nearby particles can have their forces calculated directly, while

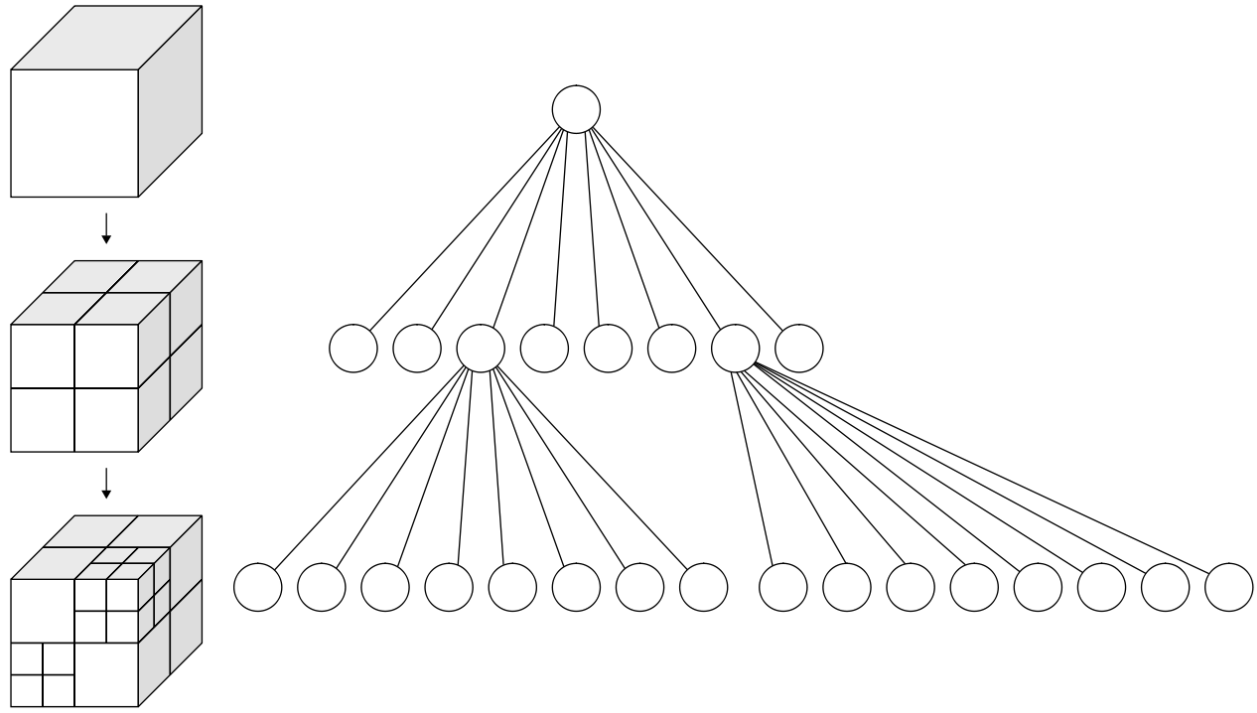


Figure 3.1: A schematic of an octree, showing how the physical simulation volume is divided into octants on the left, and how this can be stored in a tree structure on the right (Source: Wikimedia Commons).

for sets of more distant particles contained in some sub-region, the mass within the sub-region at the centre-of-mass of the sub-region can be used instead. This is because the small scale distribution of matter has only a small influence on the force felt far away due to that matter. This method of calculating forces lends itself naturally to a tree structure, as shown in Fig. 3.1. The root node of the tree contains the mass within the simulation box and the centre of mass of the box, and then eight links to its daughter nodes. These daughter nodes will contain the centre of mass of the particles within the corresponding cubic cell, as well as links to their eight daughter nodes. This continues until we reach leaf nodes, which are those containing one or zero particles.

In order to calculate the force on a particle we start at the root node, and traverse the tree. If the centre of mass of a node is sufficiently far away from the particle for which we are calculating the forces, then all particles within the corresponding cell are treated as a single particle with mass equal to the mass within the cell and at the position of the centre

of mass of the cell. If the node is not sufficiently far away then we continue to walk the tree, opening up its daughter nodes. For a node that is a distance r away, with a cell of side-length l, 'sufficiently far away' usually means $l/r < \theta_{\rm op}$, where $\theta_{\rm op}$ is known as the opening angle. In GADGET this criteria is altered somewhat, so that a node containing mass M is opened if:

$$\frac{GM}{r^2}\left(\frac{l}{r}\right)^2 > \alpha\,|\mathbf{a}|, \tag{3.29}$$

where $\mathbf{a}$ is the acceleration on the particle in question[4] and α is the $\mathrm{ErrTolForceAcc}$ parameter that determines the accuracy of the tree-force calculation.

Particle mesh: The particle mesh algorithm starts by calculating a gridded density field, by assigning the mass of the particles on to a grid. For this purpose GADGET uses a *cloud-in-cell* assignment (Hockney & Eastwood, 1981). The Poisson equation (3.21), can then be efficiently solved, by noting that in Fourier space it becomes

$$-|\mathbf{k}|^2\tilde{\Phi}(\mathbf{k}) = 4\pi G\tilde{\rho}(\mathbf{k}), \tag{3.30}$$

with $\tilde{\Phi}$ and $\tilde{\rho}$ the Fourier transforms of the potential and density respectively.

Using a Fast Fourier Transform, it is efficient to convert between a quantity and its Fourier decomposition. The gravitational forces are found by Fourier transforming the gridded density and using equation (3.30) to calculate $\tilde{\Phi}$ from $\tilde{\rho}$, before performing an inverse Fourier transform to get the potential Φ. The gravitational forces can then be calculated from the gradient of Φ. This method is fast, and has the added bonus of implementing periodic boundary conditions, as usually required for a cosmological simulation. The downside is that the resolution is limited by the size of the grid used for calculating the density and potential, and so if we want high force resolution in high density regions, we would require a fine mesh everywhere. This can be rectified by using an adaptively refined mesh, as done in ART (Kravtsov, 1999), FLASH (Fryxell et al., 2000), RAMSES (Teyssier, 2002) and ENZO (Bryan et al., 2014). In GADGET the solution is to split the force calculation into a long and short-range component. The short-range forces are calculated using a tree, and the long-range forces using a mesh with a fixed cell size.[5]

[4] In GADGET this is estimated from the acceleration in the previous time-step.

[5] This force splitting is done in Fourier space, with $\tilde{\Phi}$ split into two components, $\tilde{\Phi}^{\rm long} = \tilde{\Phi}\exp\left(-|\mathbf{k}|^2 r_{\rm s}^2\right)$

Particle time-steps

The final aspect of GADGET that we will discuss before introducing our SIDM implementation is the time-stepping procedure. This uses *leapfrog integration*, which updates the position and velocities of particles according to

$$\text{Drift} \quad : \quad \mathbf{r}_i = \mathbf{r}_{i-1} + \mathbf{v}_{i-1/2}\,\Delta t \tag{3.31}$$

$$\text{Kick} \quad : \quad \mathbf{v}_{i+1/2} = \mathbf{v}_{i-1/2} + \mathbf{a}_i\,\Delta t, \tag{3.32}$$

where $\mathbf{r}_i$ and $\mathbf{v}_i$ correspond to the positions and velocities at integer time-step i, and $\mathbf{a}_i$ is the acceleration at time-step i, which for gravity is just a function of the $\mathbf{r}_i$. Note that the positions and velocities are not updated at the same integer time-step, but a half-integer time-step apart. The positions are updated from time-step i to time-step $i+1$, by the velocity at a time half-way between i and $i+1$, and vice-versa. It is from this behaviour that the technique gets its name.

Equations (3.31) and (3.32) are complicated somewhat when dealing with time integration in an expanding Universe, as detailed in Quinn et al. (1997). Another complication arises when we want particles to have time-steps that adapt to the environment. For example, particles in the centre of dense collapsed structure need much shorter time-steps to accurately integrate their motion than particles moving through a low density region, and using the shortest required time-step for all particles in the simulation would be wasteful. The Δt are therefore allowed to vary according to

$$\Delta t_{\text{grav}} = \min\left[\Delta t_{\max}, \left(\frac{2\eta\epsilon}{|\mathbf{a}|}\right)^{1/2}\right], \tag{3.33}$$

where ϵ is still the gravitational softening length, $\mathbf{a}$ is the acceleration of the particle for which we are calculating the time-step, and η a numerical parameter that controls the time-integration accuracy.[6] To allow for particles to be synchronised, the time-steps from

and $\tilde{\Phi}^{\text{short}} = \tilde{\Phi} - \tilde{\Phi}^{\text{long}}$, with r_{s} the spatial scale at which the force is split. The long range forces are calculated using the particle mesh method with the potential Φ^{long}, while the force between pairs of particles (or particles and distant cells) can be appropriately modified from the softened Newtonian force to implement the forces coming from Φ^{short}.

[6] η is the GADGET parameter file option ErrTolIntAccuracy.

equation (3.33) are mapped onto a power of two hierarchy, where all time-steps are some maximum allowed time-step divided by a power of two.

Because of the alternating updates of position and velocity, there are two different ways to implement variable time-steps within a leapfrog scheme. These are known as 'drift-kick-drift' (DKD) and 'kick-drift-kick' (KDK). In KDK the time-step effectively changes at an intermediate point in the kick, while in DKD it changes during the drift. Because a particle's time-step is calculated from the acceleration on it, which in turn depends on particle positions, it is determined at the start or end of a drift step. For this reason, the time-asymmetry of KDK is smaller than DKD. The decreased time-asymmetry of KDK leads to better behaviour, as demonstrated for the case of orbits in a Keplerian potential by Springel (2005). To be explicit about the steps involved in KDK integration, each time-step is composed of the following:

$$\mathrm{K}(\Delta t/2) \quad : \quad \mathbf{v}_{i+1/2} = \mathbf{v}_i + \mathbf{a}(\mathbf{r}_i)\,\frac{\Delta t}{2}, \tag{3.34}$$

$$\mathrm{D}(\Delta t) \quad : \quad \mathbf{r}_{i+1} = \mathbf{r}_i + \mathbf{v}_{i+1/2}\,\Delta t, \tag{3.35}$$

$$\mathrm{K}(\Delta t/2) \quad : \quad \mathbf{v}_{i+1} = \mathbf{v}_{i+1/2} + \mathbf{a}(\mathbf{r}_{i+1})\,\frac{\Delta t}{2}, \tag{3.36}$$

where Δt for this time-step is determined by using $\mathbf{a}(\mathbf{r}_i)$ in equation (3.33).

3.2 Collisional dynamics

In §3.1.3 we introduced the collisionless Boltzmann equation (CBE), which describes the evolution of a collisionless fluid such as CDM. When considering SIDM we want to solve a similar equation, with the addition of a collision term that scatters particles into and out of patches of phase space. A question one is often asked when discussing SIDM simulations is how it can be reasonable for a simulation particle – that 'represents' a large number of DM particles – to scatter in one direction. *Is this not akin to saying that a very large number of DM particles all happened to receive the same momentum kick, and if it is the same, why is it a reasonable thing to do?*

A key thing to note about a collision term in the Boltzmann equation is that it scatters particles away from the DM phase space sheet. The CBE combined with the assumption

that DM is cold, tells us that the fine grained distribution function evolves as a 3D sheet in phase space, and that gravity acts to distort this sheet. These distortions do not tear the sheet, such that neighbouring points in phase space remain that way indefinitely. With a collision term this is no longer the case, as neighbouring points in phase space can be scattered away from one another. Thinking of simulation particles as patches of phase space is inconsistent with this picture, because parts of the patch will be scattered while others will not. We therefore find it more satisfactory to think of SIDM simulation particles as phase-space tracers. This distinction is largely philosophical, but goes some way to answering the question of why it is reasonable to scatter simulation particles as if they were individual DM particles.

3.2.1 Calculating the interaction rate

The Boltzmann equation with a collision term can be written as

$$\frac{\mathrm{d}f(\mathbf{r},\mathbf{v},t)}{\mathrm{d}t} = \Gamma_{\mathrm{in}} - \Gamma_{\mathrm{out}}. \tag{3.37}$$

The 'out' term accounts for collisions in which a particle at position $\mathbf{r}$ and with velocity $\mathbf{v}$ scatters from another particle and thus no longer has a velocity $\mathbf{v}$, while the 'in' term accounts for collisions that result in a particle with velocity $\mathbf{v}$ when it previously had some other velocity.

If the interacting particles can be treated as hard spheres with radius r, then particles will interact if they come within a distance $2r$ of each other. The cross-section for particle interactions is therefore $\sigma = \pi(2r)^2$. A particle moving at speed v through a stationary background of such particles, with a particle number density n, will scatter with background particles at a rate

$$\Gamma = n\sigma v. \tag{3.38}$$

This can be generalised for the case where background particles have a distribution of velocities described by $f_v(\mathbf{r},\mathbf{v}) = f(\mathbf{r},\mathbf{v})/\rho(\mathbf{r})$,[7] then a particle at $\mathbf{r}_i$ with velocity $\mathbf{v}_i$ will scatter at a rate

$$\Gamma_i(\mathbf{r}_i,\mathbf{v}_i) = \int f_v(\mathbf{r}_i,\mathbf{v})\rho(\mathbf{r}_i)\,\frac{\sigma}{m}|\mathbf{v}_i - \mathbf{v}|\,\mathrm{d}^3\mathbf{v}, \tag{3.39}$$

[7] Here f_v is normalised such that $\int f_v(\mathbf{r},\mathbf{v})\,\mathrm{d}^3\mathbf{v} = 1$.

from which we recover equation (3.38) using the relationship between mass density and number density ($\rho = mn$) for the case of stationary background particles ($f_v(\mathbf{r}_i, \mathbf{v}) = \delta(\mathbf{v})$, with δ the Dirac delta function).

Γ_i gives the rate of scattering for individual particles, from which we can easily calculate the out term in equation (3.37) as

$$\Gamma_{\text{out}}(\mathbf{r}_i, \mathbf{v}_i) = f(\mathbf{r}_i, \mathbf{v}_i)\Gamma_i(\mathbf{r}_i, \mathbf{v}_i). \tag{3.40}$$

However, we want to implement scattering into an N-body code, for which the scattering rate per particle was already the useful quantity. We therefore wish to implement equation (3.39) at each time-step in the code, for each particle in our simulation. Doing this accounts for Γ_{out} in equation (3.37), while Γ_{in} is accounted for by appropriately updating the velocities of those particles that scatter, such that they scatter in to a new patch of phase space with the correct probability.

3.2.2 Calculating the post-scatter velocities

We assume that the SIDM particle interactions are fully described by an azimuthally symmetric differential cross-section, which could have some velocity dependence, $\frac{\mathrm{d}^2\sigma}{\mathrm{d}\Omega\,\mathrm{d}v}$. From the differential cross-section at a particular velocity, $\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}$, we can calculate the total cross-section at that velocity as

$$\sigma = 2\pi \int_0^\pi \sin\theta \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\,\mathrm{d}\theta. \tag{3.41}$$

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

$$p(\theta)\,\mathrm{d}\theta = \frac{2\pi \sin\theta}{\sigma}\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\,\mathrm{d}\theta. \tag{3.42}$$

3.3 Methods of simulating SIDM

Before going on to explain the method we used to implement DM self-interactions in the GADGET code, we briefly discuss two potential methods of simulating SIDM that were not chosen, and the reasons for doing so. SIDM simulations in the literature are dominated by N-body methods that add velocity-kicks to the gravitational dynamics, but one could

imagine simulating SIDM using a fluid like description, or altering the method for calculating which particles receive kicks from that usually employed.

3.3.1 Fluid-like description of SIDM

For a system of collisional particles, we can define the mean free path, $\lambda = 1/\sigma n$, where σ is the cross-section for particle interactions and n the number density of particles. If λ is much smaller than the characteristic length scale of the system, then the *continuum assumption* applies and we can treat the collection of particles as a continuous fluid. This fluid can be described by macroscopic quantities: density, pressure, temperature and bulk velocity, that are single valued at each location, and are well-defined over 'infinitesimal' volume elements. To see that this is not true for a system of collisionless particles, consider two bulk streams of particles that pass through each other. In the collisionless case there will be two bulk velocities at one location, while with a short λ, particles from the two streams will interact with each other. These interactions turn the two bulk velocities into random motions, which are macroscopically manifest as an increased temperature and pressure.

Can SIDM be treated in such a way? This depends on the cross-section for particle scattering, as well as the typical densities of dark matter. If we take an SIDM cross-section of $\sigma/m = 1\,\mathrm{cm}^2\,\mathrm{g}^{-1}$, and assume a density equal to the average density inside virialised objects at $z = 0$, then we find

$$\lambda = \frac{1}{\sigma n} = \frac{1}{(\sigma/m)\rho} = \frac{1}{1\,\mathrm{cm}^2\,\mathrm{g}^{-1}\,200\rho_{\mathrm{crit},0}} \approx 150\,\mathrm{Mpc}. \tag{3.43}$$

This is much larger than the typical size of haloes, and so SIDM with such a cross-section is very far from behaving like a fluid. The mean free path in equation (3.43) considered a density that is the average within a halo's virial radius. The density inside a halo varies by orders of magnitude, so in the central high-density regions of a halo the DM may behave more like a fluid, however the fact remains that SIDM with a cross-section of the order usually considered is in an intermediate regime where collisions can neither be ignored nor treated as so frequent that they lead to fluid behaviour. This should not be surprising given that in Chapter 2 we found that for $\sigma/m = 1\,\mathrm{cm}^2\,\mathrm{g}^{-1}$ the average number of interactions that a particle undergoes by $z = 0$ is of order unity.

Yoshida et al. (2000a) ran simulations that treated SIDM as a fluid, and found that rather than reducing the central densities of haloes – one of the main astrophysical motivations for SIDM – the central cusp became steeper than the $\rho \propto 1/r$ seen with CDM. The cluster-scale halo they simulated also became significantly more spherical, at odds with observed clusters that can have large ellipticities in projected mass. The sort of cross-sections that would lead to fluid-like SIDM are also ruled out by the Bullet Cluster, where the separation between the hot gas and lensing mass (primarily DM) demonstrates that DM's behaviour cannot be the same as that of the gas. In summary, simulating DM as a fluid is valid for highly collisional SIDM, but not for the cross-sections that are currently allowed by astrophysical observations.

3.3.2 Individually resolved DM interactions

Particularly for the case of isotropic scattering, a natural way to implement scattering would be to treat the DM simulation particles as hard spheres, that scatter with each other when they touch. Particles at positions $\mathbf{x_1}$ and $\mathbf{x_2}$, with velocities $\mathbf{v_1}$ and $\mathbf{v_2}$, collide during the next time-step if

$$|\mathbf{x_1} - \mathbf{x_2} + (\mathbf{v_1} - \mathbf{v_2})\, t| = 2\, r_\mathrm{p} \tag{3.44}$$

has a solution for t in the range 0 to Δt. Here, r_p is the effective radius of a simulation particle for the purposes of scattering, defined by $\pi(2\, r_\mathrm{p})^2 = \sigma_\mathrm{p} = (\sigma/m) m_\mathrm{p}$. This setup is shown in Fig. 3.2.

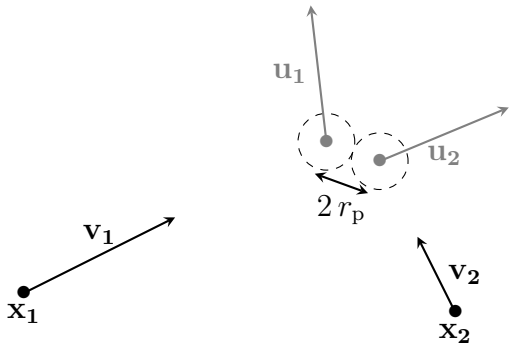


Figure 3.2: Two particles whose positions and velocities mean that they will scatter within the next time-step. For isotropic scattering,

An appealing aspect of such a scheme is that it keeps scattering local, and that the post-scatter velocities can be worked out from the pre-scatter velocities and the axis connecting the particle centres at the point of collision – which is parallel to the exchanged momentum between the two particles. However, this scheme is computationally expensive, as a quadratic equation must be solved for each pair of particles to see if they will scatter within the next time-step. Only neighbours up to some maximum separation, $\max(v_{\rm pair})\,\Delta t$, would need to be checked for potential scattering, but the maximum pairwise velocities of particles, $\max(v_{\rm pair})$, can be large. High-velocity scattering events transfer the greatest amount of energy and momentum and so should not be missed, which would require a neighbour search over a large volume.

Aside from the poor speed performance expected of this method, there is also the worry that it would be particularly sensitive to small-scale details of the simulation, such as the initial conditions. For a cross-section of $1\,{\rm cm}^2\,{\rm g}^{-1}$, the scattering radius is

$$r_{\rm p} = \left(\frac{m_{\rm p}}{10^8\,{\rm M}_\odot}\right)^{1/2} 0.041\,{\rm kpc}, \tag{3.45}$$

while for a Planck 2015 cosmology (Planck Collaboration et al., 2016), the comoving mean inter-particle separation ($l = (m_{\rm p}/\langle\rho\rangle)^{1/3}$) is

$$l = \left(\frac{m_{\rm p}}{10^8\,{\rm M}_\odot}\right)^{1/3} 130\,{\rm kpc}. \tag{3.46}$$

At low redshift the scattering radius is over three orders of magnitude smaller than the mean inter-particle separation, and over an order of magnitude smaller than a typical gravitational softening length, with these ratios increasing as we move to higher resolution. This means that whether or not particles scatter is sensitive to the trajectories of particles on scales much smaller than any other scales in the simulation. As an example of where this might cause a problem, consider a cosmic sheet or filament, collapsing from grid-like initial conditions. If this collapse proceeds along one of the grid axes, then at the point of shell-crossing, the particles will all scatter. Contrast this with the case of collapse at some significant angle to each of the grid axes, where no neighbouring particles will pass within $2\,r_{\rm p}$ of each other during shell-crossing. For glass-like initial conditions this effect would be less severe, but still one should be cautious about having a simulation that is sensitive to

particle positions on scales well below those that would usually be considered 'resolved'. We finally note that this method extends less naturally to the case of velocity-dependent or anisotropic DM–DM scattering cross-sections. For this and the other reasons above we chose not to use this method for our SIDM simulations.

3.4 Implementation

We implemented DM scattering on top of the GADGET code described earlier in this chapter. The scattering is done stochastically, using a similar algorithm to Rocha et al. (2013) that they derive from the Boltzmann equation. 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 (including Kochanek & White, 2000; Yoshida et al., 2000b; Davé et al., 2001; Koda & Shapiro, 2011; Vogelsberger et al., 2012; Fry et al., 2015; Elbert et al., 2016; Kim et al., 2017), with these algorithms primarily differing in the number of neighbours (or search volume) used to find potential scattering pairs.

For ease of discussion, we focus here on the case of isotropic scattering with a velocity-independent cross-section. In Chapter 6 we discuss how to extend this method to the more general case of an arbitrary differential cross-section.

3.4.1 Scattering Rate

For a particle moving with velocity $\mathbf{v}_i$, its scattering rate is given by equation (3.39), which we repeat here

$$\Gamma_i(\mathbf{r}_i, \mathbf{v}_i) = \int f_v(\mathbf{r}_i, \mathbf{v}) \rho(\mathbf{r}_i) \frac{\sigma}{m} |\mathbf{v}_i - \mathbf{v}| \, \mathrm{d}^3\mathbf{v}. \tag{3.47}$$

To calculate scattering probabilities within an N-body simulation, f_v and ρ are estimated from the volume within a distance h_{SI} of a particle's position. We call h_{SI} the 'scatter search radius', and note that we will sometimes simply call it h. To avoid confusion with the Hubble parameter ($h = H_0/100\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$), we will use h_{SI} in this chapter, while in future chapters the meaning of h should be clear from the context.

In equation (3.47), it is the cross-section per unit mass, rather than the cross-section of individual particles, that enters the scattering rate per particle. It is for this reason that it is σ/m that SIDM studies can constrain. Throughout this work, σ and m will refer to the cross-section and mass of individual DM particles, while σ_p and m_p will be used for 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 particles – whether they be individual particles or simulation particles. This means that the cross-section of our simulation particles must scale with the simulation particle mass, following

$$\sigma_\mathrm{p} = \left(\frac{\sigma}{m}\right) m_\mathrm{p}. \tag{3.48}$$

The simplest way to estimate the scattering rate from the particles enclosed in the search region is for all neighbour particles to contribute equally to the estimate of ρ and f_v, independent of their location within the search region. For a local number density n_p of simulation particles, the expected number of particles within the search region is $\langle N_\mathrm{p}\rangle = \frac{4}{3}\pi h_\mathrm{SI}^3\, n_\mathrm{p}$. So using $\rho\,\sigma/m = \rho\,\sigma_\mathrm{p}/m_\mathrm{p} = n_\mathrm{p}\sigma_\mathrm{p}$, we can estimate the integral in equation (3.47) by a sum over the N_p neighbour particles

$$\Gamma_i = \sum_{j=1}^{N_\mathrm{p}} \frac{\sigma_\mathrm{p}|\mathbf{v}_i - \mathbf{v}_j|}{\frac{4}{3}\pi h_\mathrm{SI}^3}. \tag{3.49}$$

From this, the probability of two particles, i and j, separated by a distance less than h_SI, scattering within the next time step, Δt, is given by

$$P_{ij} = \frac{\sigma_\mathrm{p}|\mathbf{v}_i - \mathbf{v}_j|\,\Delta t}{\frac{4}{3}\pi h_\mathrm{SI}^3}, \tag{3.50}$$

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

In this scattering procedure, h_SI is a numerical parameter that has to be chosen. In §4.2.2 we investigate the effects of changing h_SI, using both a fixed h_SI for all particles, as well as a variable h_SI that adapts to the local density. We find that using a fixed value of h_SI, similar in size to the gravitational softening length, provides correct results. This is in contrast to smoothed particle hydrodynamics (Gingold & Monaghan, 1977) (for which adaptive smoothing lengths are necessary) as the scattering is inherently stochastic, and so it is not important that the scattering probability varies smoothly.

Neighbour search

The oct-tree employed by GADGET for the gravity calculation is used to find neighbours from which a particle could scatter in the next time-step. We find particles within a spherical search region with radius $h_{\rm SI}$, by walking the tree, opening a node if it has some overlap with the search region. When using a fixed $h_{\rm SI}$ for all particles, the search is symmetric in the sense that if particle i finds particle j as a potential scattering partner, then j will also find i. For this case, we can ensure that each pair of particles separated by a distance less than $h_{\rm SI}$ is only found once, and scatters with the probability given in equation (3.50). However, for variable $h_{\rm SI}$ this symmetry in neighbour searching is no longer guaranteed, and so instead we allow particle pairs to find each other in both directions, and scatter with a probability equal to half of that in equation (3.50) each time – though in this case they will not necessarily 'meet' twice.

3.4.2 Scattering kinematics

If two particles with identical mass, and velocities $\mathbf{v}_i$ and $\mathbf{v}_j$ are to scatter, then we first move into the centre of momentum frame, in which the velocities are $\mathbf{v}'_i$ and $\mathbf{v}'_j = -\mathbf{v}'_i$. We use the direction of $\mathbf{v}'_i$ to define the z-axis, from which the polar scattering angle θ is measured. We draw a random θ from the probability density function $p(\theta)$ (equation 3.42) to determine the polar angle at which the particles scatter, as well as drawing a random azimuthal angle from the range $[0, 2\pi]$. With these two angles, the scattering kinematics in the centre of mass frame are 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 chapter, the scattering kinematics are particularly simple.[8] The post-scatter velocities are

$$\begin{aligned} \mathbf{u}_i &= \mathbf{V} + w\,\hat{\mathbf{e}} \\ \mathbf{u}_j &= \mathbf{V} - w\,\hat{\mathbf{e}} \end{aligned} \tag{3.51}$$

[8]The fully general case is discussed in Chapter 6.

where the $\mathbf{u}$ are the post-scatter velocities, $\mathbf{V} = (\mathbf{v}_i + \mathbf{v}_j)/2$, $w = |\mathbf{v}_i - \mathbf{v}_j|/2$, and $\hat{\mathbf{e}}$ is a randomly oriented unit vector. When scattering is anisotropic, the post-scatter velocities take a similar form, but the direction of $\hat{\mathbf{e}}$ is drawn from a probability density function that depends on $p(\theta)$.

3.4.3 Multiple scatters within a time-step

As particle scattering is implemented on a particle-pair by particle-pair basis, it is possible for a particle to scatter more than once in a single time-step. While the low rate of particle scattering[9] 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.

Scattering across processors

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. We call this occurrence 'bad

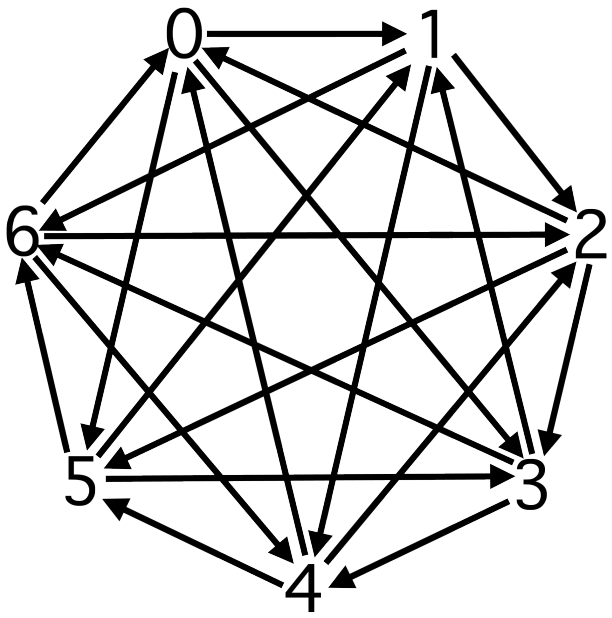
[9] For $\sigma/m = 1\,\mathrm{cm}^2\,\mathrm{g}^{-1}$ we showed in Chapter 2 that the average number of scattering events per particle is $\mathcal{O}(1)$ by redshift zero, and so the frequency of particles scattering twice within a single simulation time-step is very low.

scattering'. 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. The problem is that the collisions have not been time-ordered. With proper time-ordering, the incoming velocity for the second scattering would have been altered by the first scattering event, changing the momentum transfer in the second scattering event.

There are various ways to overcome this problem. A popular method in the literature is the use of shorter time-steps for SIDM than would be used for collisionless DM (e.g. Vogelsberger et al., 2012), to keep the probability of a particle scattering twice in one time-step negligibly small. Unfortunately, this leads to SIDM simulations that take much longer than their CDM counterparts, which is clearly undesirable. For this reason, we sought an approach that would work when time-steps are set by the same gravitational dynamics constraints as are used for CDM simulations.

Due to the domain decomposition in GADGET, most particles that reside in each other's search radii, will be stored on the same processor. Collisions between these particles are easily time ordered, as the particles' velocities can be updated as soon as they scatter. This is also true if a particle exported to another processor scatters twice there. Again, all the necessary velocities can be updated after the first scattering event, such that the second event is correctly modelled. However, a particle scattering on more than one processor, within a single time-step, is problematic.

Given the small size of $h_{\rm SI}$ compared to the size of the domains, the majority of problematic scatterings would involve particles from only two domains. For example, particle i from processor I being exported to processor J and scattering there, whilst a particle from processor J is exported to processor I and scatters off particle i there. Our solution for the case of just two domains is to assign a direction between each pair of processors, and only allow particles to be exported in this direction. To help with load-balancing, the difference between the number of processor pairs in which a processor exports particles and the number in which it imports particles, should not be larger than one. To do this, a directional communications pattern as shown schematically in Fig. 3.3 is used. Assigning a directionality to the particle send/receive process does not completely prevent particles



if $I + J$ is even:
    if $I > J$:
        I exports particles to J
    else:
        J exports particles to I
else:
    if $I > J$:
        J exports particles to I
    else:
        I exports particles to J

Figure 3.3: Pseudocode for the directional communications pattern employed in the scattering algorithm, with an example of the direction in which particles are exported for the case of 7 processors. In the pseudocode, I and J are the ranks of two processors, stored as the GADGET variable ThisTask. These processor ranks run from 0 to $N_{\text{processors}} - 1$, and are the numbers at the heptagon vertices in the schematic diagram on the left.

scattering simultaneously on different processors, as it is still possible where particles on three or more domains are within a distance h_{SI} 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_{SI} is large, such that these events are highly unlikely.

In order to quantify the level at which bad scattering takes place, 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 $\langle \Delta E \rangle = \frac{1}{2} \langle \text{KE} \rangle$, where $\langle \text{KE} \rangle$ is the mean kinetic energy of individual particles. This rises to $\langle \Delta E \rangle \approx 0.87 \langle \text{KE} \rangle$ 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 3.50).

For practical purposes, bad scattering was non-existent once we implemented directional communication. As an example, across all the simulations used in Chapter 5 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 ~ 1 part in $N_{\rm DM}$, where $N_{\rm DM}$ is the number of DM particles in the simulation. In Chapter 5 this corresponds to 1 part in 10^7, making it inconsequential compared to the non-conservation of energy due to gravitational forces. With variable time steps, manifest energy conservation is lost (Dehnen & Read, 2011), and our Chapter 5 simulations had a typical level of energy conservation over the course of a simulation of ~ 1 part in 10^4.

It is possible for bad scattering to have a noticeable impact on the results of a simulation. For example, in §4.2.5 we show the evolution of the density profile of an SIDM halo undergoing core collapse. During this process, the central regions of the halo become exceptionally dense, and the scattering rates very high. We found in this case, that if we ran the simulation on a large number of cores, then there could be appreciable rates of bad scattering, that injected energy into the system and altered the halo's evolution. For this reason, the simulations shown in §4.2.5 were run on a single processor.

3.4.4 Scattering within leapfrog integration

As discussed in §3.1.5, GADGET uses a KDK leapfrog time-stepping scheme. Our SIDM scattering implementation depends on both the particle positions (to find neighbours) and velocities (for scattering probabilities and scattering kinematics), and so here we describe exactly which positions and velocities we use. In our implementation, a time-step for an SIDM particle involves the following:

$$\mathrm{K}(\Delta t/2) \quad : \quad \mathbf{v}_{i+1/2} = \mathbf{v}_i + \mathbf{a}(\mathbf{r}_i)\,\frac{\Delta t}{2}, \tag{3.52}$$

$$\mathrm{S}(\Delta t) \quad : \quad \mathbf{u}_{i+1/2} = \mathrm{Scatter}(\mathbf{v}_{i+1/2}, \mathbf{r}_i, \Delta t), \tag{3.53}$$

$$\mathrm{D}(\Delta t) \quad : \quad \mathbf{r}_{i+1} = \mathbf{r}_i + \mathbf{u}_{i+1/2}\,\Delta t, \tag{3.54}$$

$$\mathrm{K}(\Delta t/2) \quad : \quad \mathbf{v}_{i+1} = \mathbf{u}_{i+1/2} + \mathbf{a}(\mathbf{r}_{i+1})\,\frac{\Delta t}{2}, \tag{3.55}$$

where the scatter step is the statement that the post-scatter velocities $\mathbf{u}_{i+1/2}$ are calculated from the pre-scatter velocities $\mathbf{v}_{i+1/2}$, using the positions from the start of the time-step. It is not clear that this is the best way to implement scattering, and it may be that putting half

of the drift step ($\mathbf{r}_{i+1/2} = \mathbf{r}_i + \mathbf{v}_{i+1/2}\,\Delta t/2$) before the scatter step would be preferable. We did not investigate the effect of making this change, because, as demonstrated in Chapter 4, our code performed as expected in various test cases.

3.4.5 Scattering in cosmological simulations

The scattering implementation as discussed so far has assumed that we are working in non-expanding space. For cosmological simulations, GADGET uses comoving positions, as well as velocities and time intervals that differ from their definitions in non-cosmological simulations. In this section we describe the system of units used internally by GADGET when simulating an expanding patch of space, and explain how our scattering algorithm is implemented in a manner that is consistent with these.

GADGET's system of units for cosmological simulations

We start by re-stating the definition of comoving position, $\mathbf{x}$, related to the physical position, $\mathbf{r}$, by

$$\mathbf{r} = a(t)\mathbf{x}, \tag{3.56}$$

where $a(t)$ is the cosmological scale factor.[10] From this, we can define the physical velocity, $\mathbf{v}$, and peculiar velocity, $\mathbf{v}_\mathrm{p}$, as

$$\mathbf{v} \equiv \dot{\mathbf{r}} = \dot{a}\mathbf{x} + a\dot{\mathbf{x}} \equiv \dot{a}\mathbf{x} + \mathbf{v}_\mathrm{p}\,. \tag{3.57}$$

The internal positions in GADGET are comoving, while the internal velocities are neither physical nor peculiar, instead being the 'canonical momentum'[11]

$$\mathrm{Pos} \equiv \mathbf{x} \tag{3.58}$$

$$\mathrm{Vel} \equiv a^2\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t}. \tag{3.59}$$

[10] We would like to thank Tom Theuns, as the notation in this section draws heavily from his document describing the inner workings of the EAGLE code (Schaye et al., 2015).

[11] This is different from the snapshots, where the velocity is $\mathbf{v}_\mathrm{snap} = \sqrt{a}\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t}$.

The units used in GADGET are set in the parameter file, with defaults of

$$U_M = 10^{10}\ M_\odot = 1.989 \times 10^{33}\,\mathrm{g} = \hat{U}_M\,\mathrm{g} \tag{3.60}$$

$$U_L = 1\,\mathrm{kpc} = 3.085 \times 10^{21}\,\mathrm{cm} = \hat{U}_L\,\mathrm{cm} \tag{3.61}$$

$$U_V = 1\,\mathrm{km\,s^{-1}} = 10^5\,\mathrm{cm\,s^{-1}} = \hat{U}_V\,\mathrm{cm\,s^{-1}} \tag{3.62}$$

which defines the GADGET parameter file options UnitMass_in_g $\equiv \hat{U}_M$, UnitLength_in_cm $\equiv \hat{U}_L$, and UnitVelocity_in_cm_per_s $\equiv \hat{U}_V$. The unit of time is thus a derived unit,

$$U_T = U_L\,U_V^{-1} = 3.085 \times 10^{16}\,\mathrm{s} = 0.979\,\mathrm{Gyr}. \tag{3.63}$$

We will use a hat to indicate a (dimensionless) program value, and note the following relationship between physical masses, velocities and lengths and their respective program values:

$$m = U_M\,h^{-1}\hat{m} \tag{3.64}$$

$$v = U_V\,a^{-1}\hat{v} \tag{3.65}$$

$$l = U_L\,a\,h^{-1}\hat{l}. \tag{3.66}$$

GADGET uses the expansion factor a as the time variable and $\mathrm{d}\log a$ for time-steps. The logarithm of the simulation time, which runs from scale factor a_{begin} to a_{max}, is linearly mapped onto an integer timeline with TIMEBASE integer steps. This defines the difference in $\log a$ between neighbouring points on the integer timeline:

$$\mathrm{Timebase\_interval} \equiv \frac{\log(a_{\mathrm{begin}}/a_{\mathrm{max}})}{\mathrm{TIMEBASE}}. \tag{3.67}$$

The time between two events (such as the start and end time of a time-step – a_i and a_{i+1}) is therefore stored in GADGET as an integer,

$$\mathrm{ti\_step} = \frac{\log a_{i+1} - \log a_i}{\mathrm{Timebase\_interval}}. \tag{3.68}$$

To convert a small interval in $\log a$ to an interval in time, we can apply the chain rule

$$\mathrm{d}t = \frac{a}{\dot{a}}\frac{\mathrm{d}a}{a} = \frac{\mathrm{d}\log a}{H(a)}. \tag{3.69}$$

GADGET uses the following definitions relevant to calculating the Hubble constant, $H(a)$:

$$\text{Hubble} \equiv 100\,\text{km}\,\text{s}^{-1}\,\text{Mpc}^{-1}\,U_T \tag{3.70}$$

$$E(a) \equiv \sqrt{\Omega_\text{m} a^{-3} + \Omega_\Lambda + (1 - \Omega_\text{m} - \Omega_\Lambda) a^{-2}} \tag{3.71}$$

and therefore

$$H(a) = h\,U_T^{-1}\,\text{Hubble} \times E(a). \tag{3.72}$$

Combining this with equations (3.68) and (3.69), we can relate a small time interval, Δt, to an interval in $\log a$, and in turn to GADGET's integer time-step

$$\Delta t = \frac{\Delta \log a}{h\,U_T^{-1}\,\text{Hubble} \times E(a)} = \frac{\text{Timebase\_interval} \times \text{ti\_step}}{h\,U_T^{-1}\,\text{Hubble} \times E(a)}. \tag{3.73}$$

Scattering probabilities in cosmological SIDM simulations with GADGET

Using the above relationships, we can relate the probability for a pair of particles to scatter (equation 3.50)

$$P_{ij} = \frac{(\sigma/m) m_\text{p} |\mathbf{v_i} - \mathbf{v_j}|\,\Delta t}{\frac{4\pi}{3} h_\text{SI}^3}, \tag{3.74}$$

to an expression in dimensionless program values that we can implement in GADGET

$$P_{ij} = \frac{(\sigma/m)}{(\text{cm}^2/\text{g})} \frac{\hat{U}_M}{\hat{U}_L^2} \frac{\hat{m}_\text{p} |\hat{\mathbf{v}}_\mathbf{i} - \hat{\mathbf{v}}_\mathbf{j}|}{\frac{4\pi}{3} \hat{h}_\text{SI}^3} \frac{h}{a^4} \frac{\text{Timebase\_interval} \times \text{ti\_step}}{\text{Hubble} \times E(a)}. \tag{3.75}$$

To do this, we define the following:

$$\hat{\sigma} \equiv \frac{(\sigma/m)}{(\text{cm}^2/\text{g})} \frac{\hat{U}_M}{\hat{U}_L^2} \tag{3.76}$$

$$\hat{f}_\sigma \equiv \frac{h^2}{a^4} \tag{3.77}$$

$$\hat{f}_\text{dt} \equiv \frac{1}{h \times \text{Hubble} \times E(a)} \tag{3.78}$$

$$\hat{\text{d}t} \equiv \text{Timebase\_interval} \times \text{ti\_step}, \tag{3.79}$$

such that the scattering probability can be calculated in terms of program values as

$$P_{ij} = \frac{\hat{\sigma}\,\hat{m}_\text{p} |\hat{\mathbf{v}}_\mathbf{i} - \hat{\mathbf{v}}_\mathbf{j}|}{\frac{4\pi}{3} \hat{h}_\text{SI}^3}\,\hat{f}_\sigma\,\hat{f}_\text{dt}\,\hat{\text{d}t}. \tag{3.80}$$

Chapter 4

Testing the self-interacting dark matter implementation

4.1 Scattering from a uniform background

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 from either the cube or background to scatter more than once. All of the particles in the cube had initial velocities v_0 along the

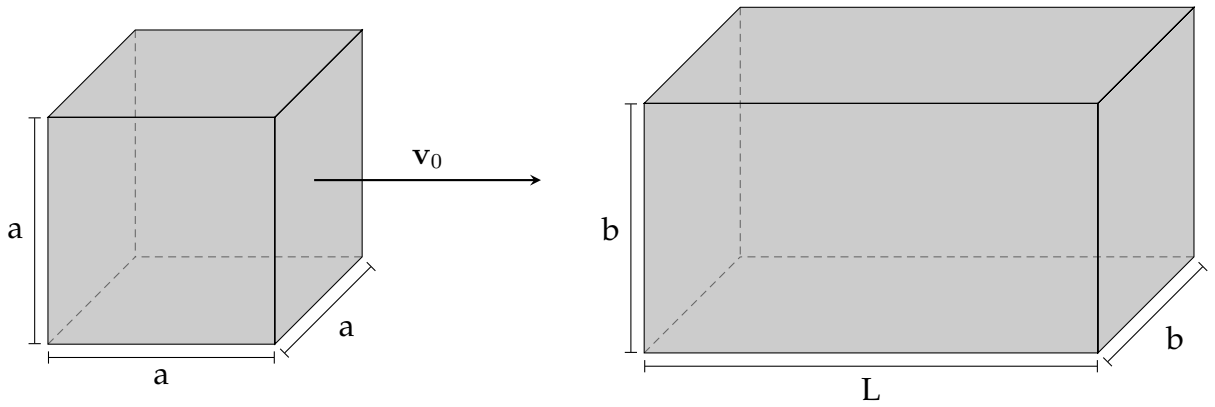


Figure 4.1: The setup of our test case, where a cube containing N_c particles moves through a cuboid with a number density of stationary particles n_b. The cube and background particle positions are drawn randomly within their respective volumes. Our test case used $b = \frac{11}{10}a$ and $L = 2a$, with $N_\mathrm{c} = 10^5$ particles in the cube. The density of points in the background cuboid was $n_\mathrm{b} = 2.5 \times 10^5\, a^{-3}$.

z-axis, and gravity was turned off.

4.1.1 Scattering rates

Predicted scattering rates

The expected rate of scattering for each particle in the cube is $\Gamma = n_\mathrm{b}\, v_0\, \sigma_\mathrm{p}$, where $n_\mathrm{b} = (\rho_\mathrm{b}/m_\mathrm{p})$ is the number density of background particles. This leads to the expected number of interactions after a time t being approximately

$$N_\mathrm{exp} = N_\mathrm{c}\, \Gamma\, t = N_\mathrm{c}\, n_\mathrm{b}\, v_0\, \sigma_\mathrm{p}\, t, \tag{4.1}$$

with N_c the number of particles in the cube.

The expected number of scattering events is complicated by the fact that scattered particles can no longer interact.[1] This effect can be captured analytically by noting that the number of particles in the cube changes in time according to

$$\frac{\mathrm{d}N_\mathrm{c}}{\mathrm{d}t} = -\Gamma\, N_\mathrm{c}, \tag{4.2}$$

with the solution

$$N_\mathrm{c}(t) = N_\mathrm{c}(0) \exp\left(-n_\mathrm{b}\, v_0\, \sigma_\mathrm{p}\, t\right). \tag{4.3}$$

The expected number of scattering events before a time t is then simply $N_\mathrm{c}(0) - N_\mathrm{c}(t)$, such that if we let the cube travel through a uniform background of length L

$$N_\mathrm{exp} = N_\mathrm{c}(0)\left[1 - \exp\left(-n_\mathrm{b}\sigma_\mathrm{p} L\right)\right]. \tag{4.4}$$

This is still a slight simplification, as when particles at the front of the cube (the right hand side of the cube in Fig. 4.1) scatter from background particles, they lower the background density of particles seen by particles towards the back of the cube. This effect can still be modelled analytically, for which we point the interested reader to Appendix A.

[1] Allowing scattered particles to scatter again would also complicate things, as scattered particles no longer have velocities v_0 or 0.

Measured scattering rates

Our test simulations used the parameters given in the caption to Fig. 4.1, with the particle cross-section chosen such that $n_{\rm b}\sigma_{\rm p}L = 0.1$. The number of scattering events measured in this test is plotted relative to the prediction from equation (4.4) in Fig. 4.2 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% of the mean background inter-particle separation. By running the test case with different time steps, we find that this 20% 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, cross-section and the relative velocity of particles.

Probability saturation

In general, the scattering rate is insensitive to h, as the number of neighbour 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 (3.50) ($\propto h^{-3}$). The product of the number of neighbour 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 neighbour 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 neighbour 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 panel of Fig. 4.2. 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. Given that the time-steps in these test simulations were set arbitrarily, there is no significance to the value of h at which probability saturation occurred (counter to the conclusions drawn by Rocha et al., 2013). In §4.2 we discuss scattering in a DM halo, where the time-steps are limited by dynamical considerations, following equation (3.33). We therefore leave the discussion of how to choose h to avoid probability saturation for

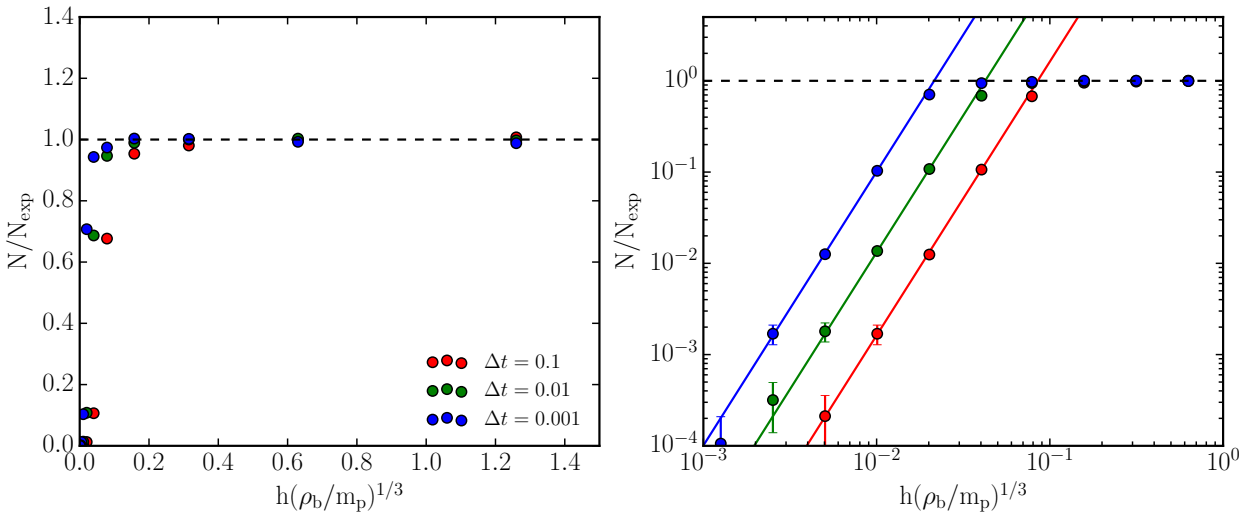


Figure 4.2: The number of scattering events in our test simulations as a function of neighbour-search radius, h. The left panel is similar to Fig. 1 in Rocha et al. (2013), and we find that we also see a decrease in the rate of scattering, below that expected, when using small h. While this happens for $h \lesssim 0.2(m_{\rm p}/\rho_{\rm b})^{1/3}$ 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, here measured in units of a/v_0. As discussed in the text, we can now explain this effect as 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 panel we show the same data as in the left 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, $N_{\rm exp} \approx 10^4$, and the error bars show the 1σ uncertainty, assuming that N is Poisson distributed.

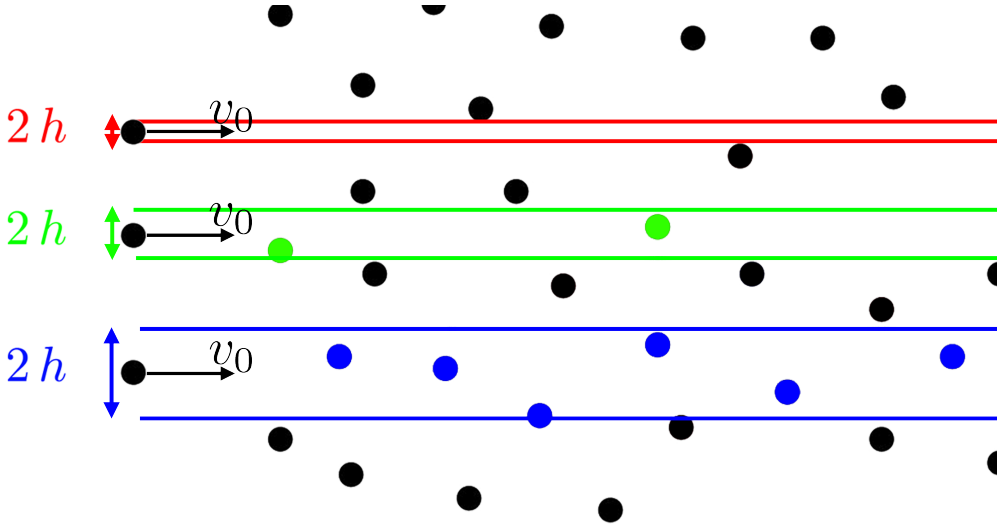


Figure 4.3: Diagram of the repeated sampling problem. The blue, green and red lines mark the regions sampled by particles with different search radii. As h is decreased, the number of potential scattering partners that a particle finds deceases. For sufficiently small h, the expected number of potential scattering partners drops below the expected number of scatters and the scattering rate drops below what it should be.

§4.2.3, where it can be discussed in a context relevant for astrophysical simulations.

Repeated sampling

The problem of probability saturation can be avoided, so long as sufficiently small time-steps are used. However, another problem arises when h is made too small. The problem is easy to see when one considers the simple test case used here. As h is made smaller, the number of background particles that any one particle from the moving cube will sample during the course of the simulation decreases. In the extreme case that fewer than $N_{\rm exp}$ of the $N_{\rm c}$ cube particles will find another particle in their search region, the number of particles that scatter must decrease below that expected. This problem is demonstrated pictorially in Fig. 4.3.

To see when this becomes a problem, we can consider simulating the particles as hard spheres that scatter whenever they are separated by a distance $r_{\rm p} = \sqrt{\frac{\sigma_p}{\pi}}$. This was the method of simulating scattering described in §3.3.2. From considering simulating the interactions in this way, we can see that our method will clearly have broken down for $h \leq r_{\rm p}$.

In fact this problem exists even for larger h than this, though as discussed in §3.3.2, for astrophysically allowed cross-sections, $r_{\rm p}$ is orders of magnitude smaller than any other scale in the simulation, and so this problem can be avoided with a sensible choice for h.

It is interesting to note how breaking this constraint can lead to a decrease in the number of scattered particles, even though the probability of scattering in any one time-step (and hence the scattering rate) is being correctly (albeit stochastically) sampled. As h is decreased, the time-steps must be decreased according to $\Delta t \propto h^3$ to avoid probability-saturation. This in turn means that the distance travelled in a time-step is $\propto h^3$, so that with small h, particles sample the same region of space in many successive time-steps. Pairs of particles that end up close together can scatter off each other several times, keeping the total scattering rate correct, but not distributed between the particles in the correct way. We therefore call this second problem associated with a small neighbour-search radius, *repeated sampling*.

4.1.2 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. 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. 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. These results are shown in Fig. 4.4, with results that match expectations.

4.2 Scattering in an isolated halo

Astrophysically allowed SIDM cross-sections result in DM interactions that only happen at an appreciable rate in the centre of gravitationally collapsed objects. An important test of our code, is therefore how well it captures the rate of scattering inside a DM halo. To

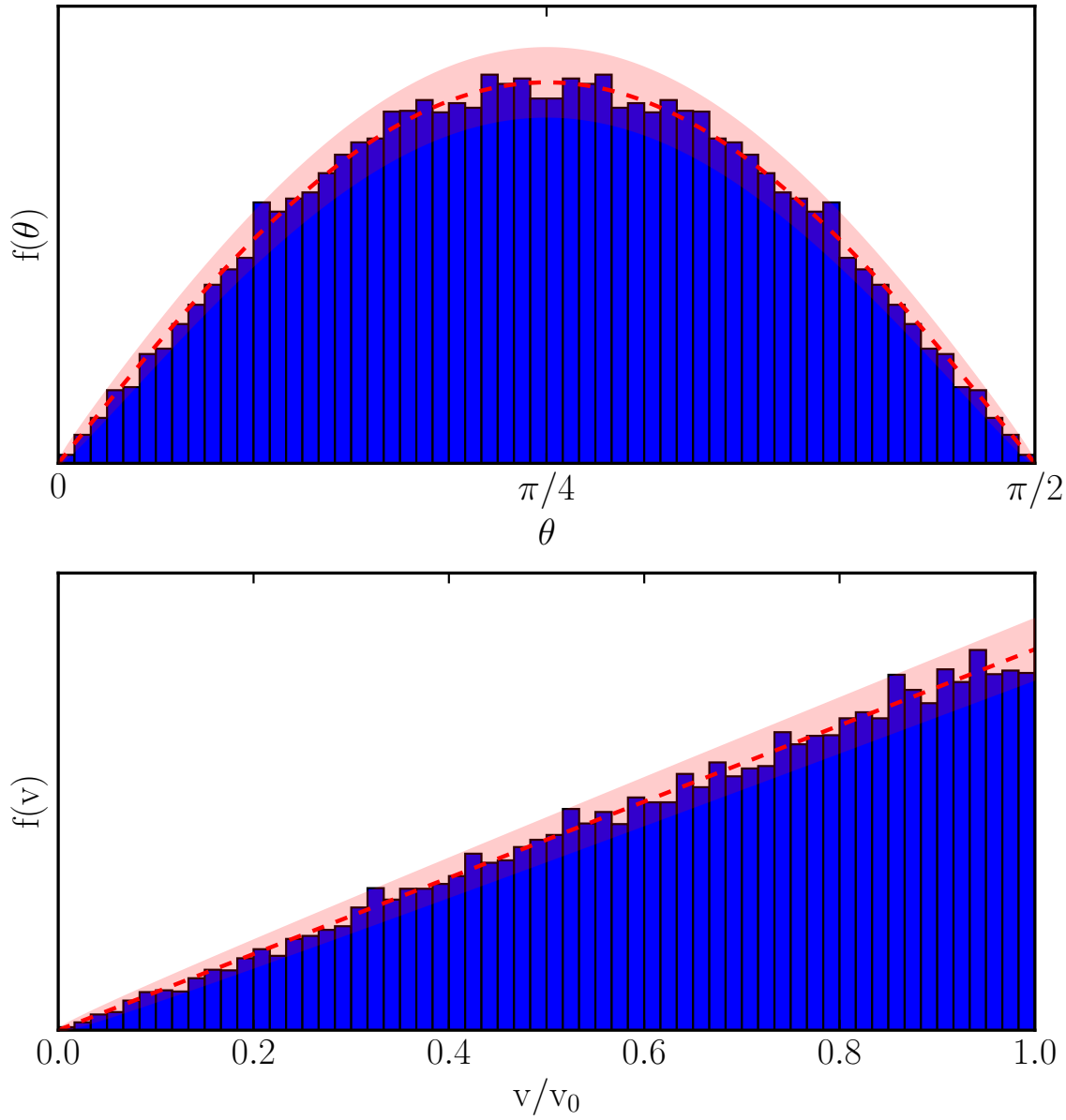


Figure 4.4: The distribution of polar angles and velocity magnitudes of scattered particles in one of our test simulations. The red dashed line shows the expected distribution, with the red shaded region showing the 2σ variation about this expectation, assuming the number of particles in each bin is Poisson distributed. The top panel is symmetric about $\theta = \pi/4$ because (in the frame of the stationary background) particles that scatter leave at right angles to one another, so if a cube particle is scattered by an angle θ_1, then the background particle with which it scattered moves at an angle $\pi/2 - \theta_1$ with respect to the z-axis.

test this, we simulate isolated Hernquist profiles (Hernquist, 1990), which have a density distribution

$$\rho(r) = \frac{M}{2\pi}\frac{a}{r}\frac{1}{(r+a)^3}, \tag{4.5}$$

with M the total mass of the halo, and a the scale radius. These are used because unlike NFW profiles (equation 2.12), 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 generate haloes that closely resemble those found in cosmological simulations, we use a matching procedure between NFW profiles and Hernquist profiles. Given the Hernquist profile's two free parameters, we require two matching criteria. The first of these we take to be matching the normalisation 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,

$$M(<r) = M\frac{r^2}{(r+a)^2}. \tag{4.6}$$

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

$$M = M_{200}\frac{(r_{200}+a)^2}{r_{200}^2} \tag{4.7}$$

$$a = \frac{r_{200}}{\sqrt{\frac{c^2}{2\left[\ln(1+c)-\frac{c}{1+c}\right]}}-1}. \tag{4.8}$$

We note that a similar matching procedure is described in the text of Springel et al. (2005a), 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.

4.2.1 Generating Hernquist profile initial conditions

In order to generate a particle distribution with positions and velocities corresponding to a Hernquist profile, we need the phase space distribution function. For a Hernquist profile

this is only a function of the specific energy,[2] E, given by Hernquist (1990)

$$f(E) = \frac{M}{8\sqrt{2}\pi a^3 v_g^3} \frac{1}{(1-q^2)^{5/2}} \Big[3\sin^{-1} q + q(1-q^2)^{1/2}(1-2q^2)(8q^4-8q^2-3)\Big], \tag{4.9}$$

where

$$v_g = \sqrt{\frac{GM}{a}} \tag{4.10}$$

and

$$q = \sqrt{\frac{-aE}{GM}}. \tag{4.11}$$

First, the positions of particles are generated by drawing uniform random numbers in the interval 0 to 1 and mapping them onto radii. For a uniform random variable $X \sim \mathrm{U}(0,1)$, this is done by finding r such that

$$X = \frac{M(<r)}{M} = \frac{r^2}{(r+a)^2}. \tag{4.12}$$

The angular positions of the particles are simply found by drawing polar and azimuthal angles corresponding to a uniform distribution over the unit sphere.

The velocities are more complicated. Given a particle at radius r with velocity $\mathbf{v}$, its specific energy is

$$E = \Phi(r) + \frac{1}{2}|\mathbf{v}|^2, \tag{4.13}$$

where

$$\Phi(r) = -\frac{GM}{r+a} \tag{4.14}$$

is the gravitational potential due to a Hernquist mass distribution. All particles must be bound to the halo, such that there is a maximum velocity for particles at radius r

$$v_{\mathrm{esc}}(r) = \sqrt{\frac{2GM}{r+a}}. \tag{4.15}$$

To draw a velocity for a particle at radius r according to the distribution function, we can use rejection sampling. We first draw a $\mathbf{v}$ uniformly from within the sphere defined by $|\mathbf{v}| < v_{\mathrm{esc}}(r)$. This velocity is then accepted with a probability proportional to $f(E)$. This can be done by finding the maximum $f(E)$ at that radius, which is $f\Big(\Phi(r)\Big)$, and drawing a uniform random variable $Y \sim \mathrm{U}(0,1)$. We then accept the velocity $\mathbf{v}$ if

$$Y f\Big(\Phi(r)\Big) \leq f\Big(\Phi(r) + \frac{1}{2}|\mathbf{v}|^2\Big), \tag{4.16}$$

[2] This means that the velocities of particles are isotropically distributed at all radii.

otherwise we draw a new random velocity and repeat the process.

4.2.2 Scattering rates in an isolated halo

Having generated Hernquist-profile initial conditions, we now use them to explore SIDM scattering within a DM halo. We checked that our profiles are stable with collisionless DM, finding that with a suitable choice of time-integration and tree-force accuracy parameters, the density profile and velocity distribution remained unchanged except for the formation of a small constant density core – with radius similar to the gravitational softening length. This stability, and the numerical choices necessary to achieve it are discussed further in §4.2.4.

In Fig. 4.5 we plot the scattering rate per particle in an isolated DM halo with a Hernquist density profile. The halo has a total mass of $M = 10^{15}\,\mathrm{M}_\odot$ and a scale radius $a = 1000\,\mathrm{kpc}$. The simulation was run for 2.5 Gyr with 10^6 particles, each with a mass $m_\mathrm{p} = 10^9\,\mathrm{M}_\odot$, and a Plummer-equivalent gravitational softening length $\epsilon = 12\,\mathrm{kpc}$.

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.

Predicted 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 (5.1), 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 $\langle v_\mathrm{pair}\rangle = (4/\sqrt{\pi})\sigma_\mathrm{1D}$, where σ_1D is the one-dimensional velocity dispersion. This can be calculated from the density profile and the Jeans equation, which (assuming an isotropic velocity distribution) gives

$$\begin{aligned}\sigma_\mathrm{1D}^2 &= \frac{GM}{12a}\left\{\frac{12r(r+a)^3}{a^4}\ln\left(\frac{r+a}{r}\right) - \frac{r}{r+a}\left[25 + 52\frac{r}{a} + 42\left(\frac{r}{a}\right)^2 + 12\left(\frac{r}{a}\right)^3\right]\right\} \\ &\equiv \frac{GM}{12a} f_\mathrm{H}(\tilde{r}),\end{aligned} \tag{4.17}$$

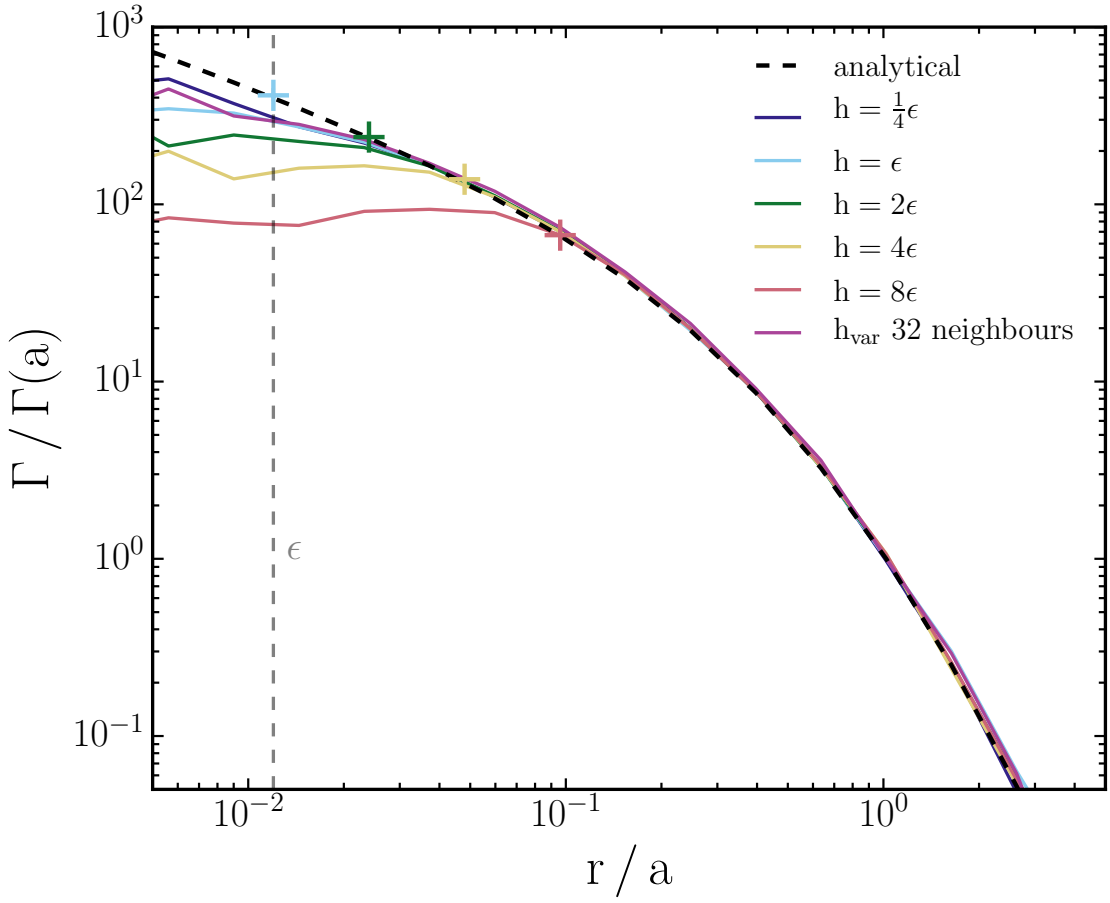


Figure 4.5: The scattering rate per particle in a Hernquist profile DM halo, plotted as a function of radius. Increasing the size of the search radius used for DM scattering leads to a decrease of the scattering rate in the inner regions of the halo. The results converge for $h < \epsilon$ as the density 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 $h_{\rm var}$ 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 $r \lesssim \max(h, 2\epsilon)$ the scattering rate falls below the analytical result.

for a Hernquist profile, where $\tilde{r} = r/a$.

Integrating over the velocity distribution function in equation (2.7) gives the average scattering rate for particles at position $\boldsymbol{r}$,

$$\Gamma(\boldsymbol{r}) = \frac{\langle \sigma\, v_{\rm pair} \rangle(\boldsymbol{r}) \rho(\boldsymbol{r})}{m}. \tag{4.18}$$

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 (5.1) and (4.17). This is shown as the dashed line in Fig. 4.5.

Measured 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 with 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. 4.5 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 analytical 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 with the true unsmoothed rate. The scattering rate in the simulations drops significantly below the analytical 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 discussed in §4.1.1, Rocha et al. (2013) found that their scattering algorithm under-predicted 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% of the mean inter-particle separation. For $h = 0.1\,{\rm kpc}$ in Fig. 4.5, h is $\sim 4\%$ of the mean inter-particle separation at $r = 2a$, but the scattering rate still matches the analytical prediction. This supports our findings in §4.1.1, namely that the Rocha et al. (2013) result stems from probability saturation, and so depends on the time-steps used.

4.2.3 How to choose the neighbour-search radius

Fig. 4.5 demonstrates that using smaller values of $h_{\rm SI}$, or an adaptive $h_{\rm SI}$ that depends on the local density, allows the scattering rate to be more accurately captured. At the same time, using too small an $h_{\rm SI}$ can lead to the problems of probability saturation or repeated sampling. In this section we use simple analytical arguments for the case of a Hernquist profile, to find an appropriate size for the SIDM neighbour search radius.

As previously discussed, the scattering rates for particles should not depend sensitively on $h_{\rm SI}$. With a large $h_{\rm SI}$, particles find many neighbours, but then have a low probability of scattering from each of them. For small $h_{\rm SI}$, the stochasticity is switched from the drawing of an unlikely random number to the unlikely event of finding neighbouring particles with separation $< h_{\rm SI}$. There are however certain advantages to both, and we start this section by briefly outlining the pros and cons of large and small $h_{\rm SI}$, before going on to explain how we choose $h_{\rm SI}$ in our simulations.

The case for small $h_{\rm SI}$

In general a smaller search radius is better because

(i) Scattering is more local, which leads to better conservation of angular momentum, which is not implicitly conserved by non-local scattering.

(ii) The scattering rate resolves local density peaks and troughs. A particle's probability of scattering depends on the local density measured over a scale $\sim h_{\rm SI}$, which if too large could smooth away density peaks that should lead to a locally high rate of

scattering.

(iii) Fewer neighbours are found, making scattering less computationally expensive. This is particularly relevant for simulations run on multiple cores, where a small $h_{\rm SI}$ significantly reduces the fraction of particles with neighbours that reside in different domains. Not only is communication between MPI tasks computationally expensive, but in the case of particle scattering it introduces additional complications discussed in §3.4.3.

The problem with small $h_{\rm SI}$

There is only one obvious drawback to using a small $h_{\rm SI}$, namely that P_{ij} (equation 3.50) increases as $1/h_{\rm SI}^3$, such that for too small an $h_{\rm SI}$, $P_{ij} > 1$. To avoid probability saturation, small time-steps need to be used. Using smaller time-steps than those needed to adequately implement gravitational dynamics would lead to simulations that take much longer than their CDM counterparts, which is clearly undesirable.

Choosing a suitable $h_{\rm SI}$

From the previous discussion it is clear that we want to use as small a search radius as possible, while ensuring that we keep $P_{ij} < 1$. In Fig. 4.6, we show how the expected value of P_{ij} varies with radius in a Hernquist profile, with both a fixed $h_{\rm SI}$ and one that varies to keep a fixed number of neighbours enclosed within the search radius. The scattering rate for a particle depends on the local density and velocity distribution as described by equation (4.18). The density profile and one-dimensional velocity dispersion for a Hernquist halo are given by equations (5.1) and (4.17) respectively.

The gravitational time-step for a particle in GADGET is given by equation (3.33), repeated here for convenience

$$\Delta t_{\rm grav} = \min\left[\Delta t_{\rm max}, \left(\frac{2\eta\epsilon}{|\mathbf{a}|}\right)^{1/2}\right], \tag{4.19}$$

with ϵ the gravitational softening length, $\mathbf{a}$ the acceleration of the particle, and η a numerical parameter that controls the time-integration accuracy. Given the Hernquist density profile,

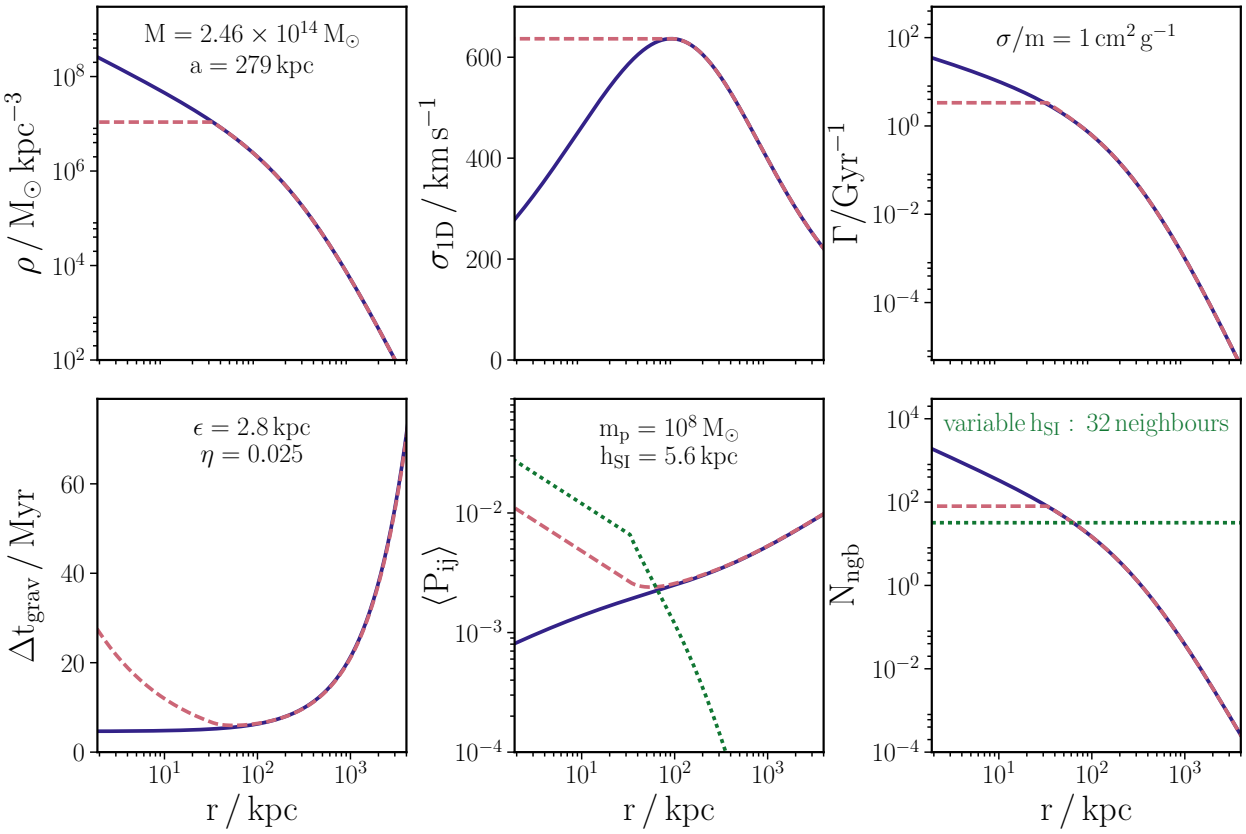


Figure 4.6: The blue solid line in the top-left and top-centre panels show the density profile and velocity dispersion profile in a Hernquist halo, while the top-right panel shows the expected rate of scattering at different radii assuming that $\sigma/m = 1\,\mathrm{cm}^2\,\mathrm{g}^{-1}$. The red dashed line shows the same quantities for a particle distribution that approximately resembles the cored density profiles found in SIDM simulations. The bottom-left panel then shows the time-step calculated for particles, assuming $\epsilon = 2.8\,\mathrm{kpc}$ and $\eta = 0.025$. From the time-step, velocity dispersion and cross-section, the expected pairwise scattering probability can be found, which is plotted in the bottom-centre panel assuming a particle mass of $10^8\,\mathrm{M}_\odot$ and a fixed h_{SI} of $5.6\,\mathrm{kpc}$. The green dotted line shows how this changes for the cored Hernquist profile if h_{SI} varies with density to keep 32 neighbour particles enclosed in each particles' search region. The bottom-right panel shows the expected number of neighbour particles for particles at different radii, the product of N_{ngb} and $\langle P_{ij} \rangle$ being equal to $\Gamma \Delta t_{\mathrm{grav}}$.

we can calculate the gravitational acceleration as a function of radius, and therefore calculate the gravitational time-step. This is plotted in the bottom-left panel of Fig. 4.6, where we ignore that GADGET also imposes a maximum time-step on all particles in a simulation.

For a given particle mass and search radius, we can calculate the mean scattering probability for pairs of particles with separation $< h_{\rm SI}$ from $\langle v_{\rm pair} \rangle$ and Δt. This can be found analytically,

$$\langle P_{ij} \rangle \, (\tilde{r}) = \frac{(\sigma/m) m_{\rm p}}{h_{\rm SI}^3} \left(\frac{3\pi \, \eta \, \epsilon \, a \, f_{\rm H}(\tilde{r})}{2} \right)^{1/2} (1 + \tilde{r}), \tag{4.20}$$

and is plotted in the bottom-middle panel of Fig. 4.6. For a Hernquist profile we find that $\langle P_{ij} \rangle$ decreases towards the centre of a halo, despite the fact that the scattering rate per particle is increasing. This is in part because the time-steps decrease towards the halo centre, decreasing the probability of scattering within a single time-step, and also because the increase in density towards the centre of the halo does not increase $\langle P_{ij} \rangle$, instead increasing the expected number of neighbours within a particle's search region, $N_{\rm ngb}$ (bottom-right panel).

The haloes formed in SIDM do not have the high central densities, and low central velocity dispersions seen in simulated CDM haloes. Instead they form a cored central region with constant density, and an isothermal velocity dispersion. Over a wide range of cross-sections, Hernquist profiles form cores that grow to a size $r_{\rm core} \approx 0.12\, a$, at which point they stabilise (Kochanek & White, 2000, and see Fig. 4.10), with a central velocity dispersion close to the peak velocity dispersion of the corresponding halo with collisionless DM (Rocha et al., 2013). To mimic these effects, we also include in Fig. 4.6 the results for a Hernquist halo altered to have $\rho(r < 0.12a) = \rho(r = 0.12a)$ and a velocity dispersion that, inside of the radius at which it peaks for a standard Hernquist profile, is equal to this peak velocity dispersion.

The radial distribution of P_{ij} in this altered Hernquist profile rises towards low radii, driven by the low accelerations (and subsequently large time-steps) in the centre of a cored halo. However, even with this rise towards small radii, $P_{ij} \lesssim 0.01$ over a large range of radii with our assumed numerical parameters. This is particularly true in practice because

(i) The acceleration in the inner region of a cored halo is increased by departures from spherical symmetry, and by substructures or other nearby haloes.

(ii) GADGET employs a time-step limit, which for the simulations discussed in Chapter 5 (that provided the inspiration for the numerical values in Fig. 4.6) was $\Delta t_{\rm max} = 10\,{\rm Myr}$. As large $\Delta t_{\rm grav}$ at both small and large radii drive up P_{ij}, the time-step limit controls P_{ij} in these regions.

The results presented in Fig. 4.6 demonstrate that for a cluster mass halo simulated with $\mathcal{O}(10^6)$ particles, $h_{\rm SI} \sim \epsilon$ results in $P_{ij} \ll 1$. However, to inform our choice of $h_{\rm SI}$ in other systems, or when simulating systems at higher resolution, it is informative to consider how P_{ij} depends upon the mass of the system being simulated, and the mass of the simulation particles. We do this by assuming that as we change the particle mass, we change the gravitational softening length and SIDM search radius accordingly, with the scaling $m_{\rm p} \propto \epsilon^3 \propto h_{\rm SI}^3$. Also, we assume that halo concentration is independent of halo mass, such that for a Hernquist halo, $M \propto a^3$. Making these assumptions and using equation (4.20) we can see that: when changing the resolution while simulating a fixed physical system $\langle P_{ij} \rangle \, (\tilde{r}) \propto m_{\rm p}^{1/6}$, when simulating systems of different mass with a fixed resolution $\langle P_{ij} \rangle \, (\tilde{r}) \propto M^{1/6}$, and when simulating different systems keeping the number of particles fixed $\langle P_{ij} \rangle \, (\tilde{r}) \propto M^{1/3} \propto m_{\rm p}^{1/3}$.

This result means that as we increase the resolution of our simulations, or simulate systems smaller than galaxy clusters, the pairwise scattering probabilities decrease. As the example case in Fig. 4.6 was a massive system at low resolution, setting $h_{\rm SI} \sim \epsilon$ should correctly implement scattering in most systems we would want to simulate. We also note that if instead of assuming halo concentration to be independent of mass, we had set some concentration-mass relation, this would have changed how $\langle P_{ij} \rangle$ scales with halo mass. For example, if we use the simple relationship from Duffy et al. (2008)

$$c_{200} = 3.93 \left(\frac{M_{200}}{2 \times 10^{12}\, h^{-1}\, {\rm M}_\odot} \right)^{-0.097}, \tag{4.21}$$

and match NFW profiles to Hernquist profiles following the discussion earlier in this chapter, then we find $\langle P_{ij} \rangle \, (\tilde{r}) \propto M^{0.20}$ provides a good fit over a wide halo mass range ($M_{200} = 10^6 - 10^{16}$ ${\rm M}_\odot$).

4.2.4 The evolution of density profiles with SIDM

One of the main astrophysical motivations for SIDM is to alter the central density profiles of DM haloes. In this section we investigate the formation of constant density cores at the centre of SIDM haloes, by allowing Hernquist profiles to evolve under the influence of DM scattering.

Core formation with CDM

To investigate core sizes in SIDM, we need to first ensure that we do not form cores as numerical artefacts. Even for collisionless DM, simulated haloes will form small cores due to gravitational softening. We therefore evolved Hernquist profiles with collisionless DM, for a range of different simulation parameters, and found the range of parameters that formed cores that were stable and acceptably small. The halo used for these tests had $M = 10^{14}\,\mathrm{M}_\odot$ and $a = 225\,\mathrm{kpc}$. Our fiducial set of simulation parameters were a particle number of $N = 128^3$, and a time-integration and force accuracy of $\eta = 0.005$ and $\alpha = 0.0012$ respectively. We used a fiducial gravitational softening length of $\epsilon = 4.4\,\mathrm{kpc}$, and varied the softening with the simulation particle mass according to $\epsilon \propto m_\mathrm{p}^{1/3}$.

From these fiducial parameters, we varied each parameter individually. The resulting radial density profiles after 10 Gyr for each of these test simulations, are shown in Fig. 4.7. In order to reduce the noise in our density profile measurements, the density in each radial bin was calculated as the average of eleven snapshots uniformly distributed in time from $9.5\,\mathrm{Gyr}$ to $10.5\,\mathrm{Gyr}$ after the simulations began. We find that the density profiles are converged with respect to η and α for our fiducial parameters, and that with sufficiently accurate force resolution and time integration accuracy, cores do not grow much beyond a couple of gravitational softening lengths.

When the force accuracy used was not sufficient, cores formed and continued to grow throughout the simulation. These cores had a distinctive shape, with a sharp transition to a flat central density. The required force accuracy parameter (α) to avoid formation of these runaway cores was dependent on the number of particles used, with more particles requiring the use of a smaller α. For this reason, the simulation run with 256^3 particles used

a smaller α than the fiducial value. The opening criterion used in GADGET is more complicated than a simple opening angle (see equation 3.29), making the analysis of the situation challenging. We do not fully understand the behaviour of our simulations when running with large α values, although it is clear that the gravitational forces acting on a particle calculated with tree gravity can be biased by not opening up sufficiently nearby cells. In any case, Fig. 4.7 allows us to empirically test for what range of numerical parameters we have converged results.

CDM run-times

While investigating the importance of different accuracy parameters for the results of our simulations, we also recorded how these parameters affected the run-time of the simulations. The results are plotted in the bottom-right panel of Fig. 4.7. The CPU time (wall clock time multiplied by number of cores) was measured over 0.1 Gyr of the simulation, starting soon after the beginning of the simulation. The measurement was made near the beginning of the simulation, as at this time the numerical parameters had not had much effect on the system being simulated. We did not measure from the beginning of the simulation to avoid including the time associated with various startup procedures.

By fitting power laws to the computation time's dependence on the different accuracy parameters, we can see the computational cost of using different parameter values. The scaling with particle number was super-linear ($t_{\rm CPU} \propto N^{1.42}$), which is expected as the computation per particle increases with increasing particle number. This is because the number of other particles with which a particle interacts increases with increasing N (although with tree gravity the number of force evaluations per particle only scales as $\log N$), and also because the softening and hence time-steps of particles was decreased with increasing particle number. The dependence on η ($t_{\rm CPU} \propto \eta^{-0.44}$) can also be understood, by noting that in equation (3.33) $\Delta t \propto \eta^{1/2}$. As such, the number of time-steps (and hence the CPU time) scales $\propto \eta^{-1/2}$, in rough agreement with our empirical findings. The α dependence is somewhat harder to predict, given that it depends sensitively on the distribution of matter in the simulation. Perhaps surprisingly, given that it is an important parameter in terms of the stability of our haloes, we found that run times depended only weakly upon

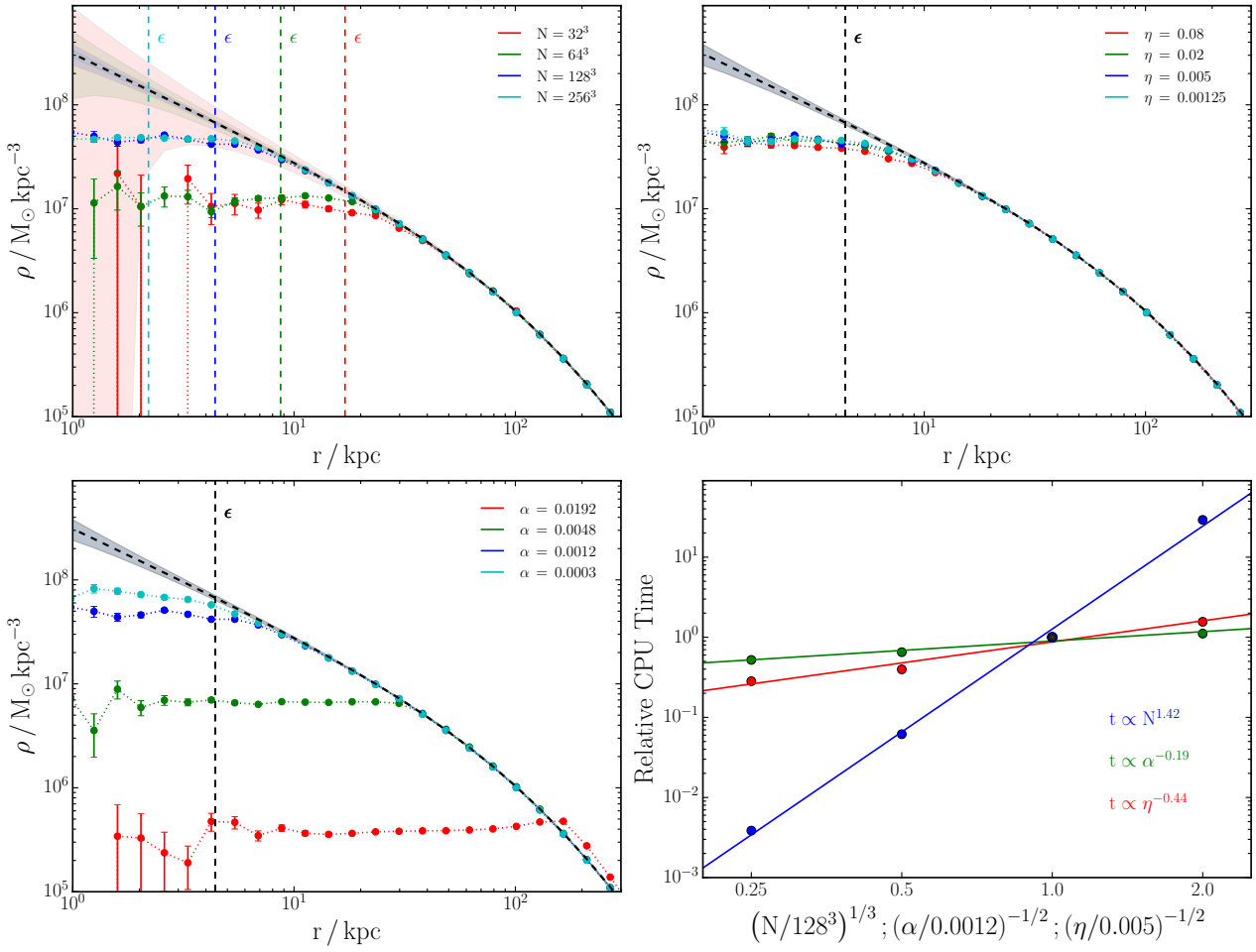


Figure 4.7: Radial density profiles after 10 Gyr, plotted for simulations run with different particle numbers and with different force and time-integration accuracy parameters. The black dashed lines show the Hernquist density profile used for the initial conditions, with the shaded regions showing the expected 1σ variations in the initial density profile, assuming the number of particles in each radial bin is Poisson distributed. The vertical dashed lines show the gravitational softening length. Our fiducial parameters were $N = 128^3$, $\eta = 0.005$ and $\alpha = 0.0012$, from which we varied the parameters individually. The $N = 256^3$ simulation used $\alpha = 0.0003$ rather than the fiducial value because the α at which the simulations converged was seen to be dependent on the number of particles used. The bottom-right panel shows how the CPU time taken for the simulations depends on the chosen numerical parameters.

α, with $t_{\rm CPU} \propto \alpha^{-0.19}$.

Core formation with SIDM

Now with a knowledge of the expected numerical core-sizes, we can investigate the cores that form due to self-interactions. We turn on scattering between DM particles, using the method described in Chapter 3. Similar to our CDM convergence tests, we evolved our Hernquist halo with SIDM with different particle numbers and with different force and time-integration accuracy parameters. The results for $\sigma/m = 1\,{\rm cm}^2\,{\rm g}^{-1}$ are shown in Fig. 4.8. When a large α value is used, the density profile is still unstable, but for the other parameters, the profiles appear well converged. In particular, the simulations that were run with a small number of particles do a reasonable job of capturing the cored structure, though presumably this would stop if the softening became comparable to the physical core size set by the SIDM cross-section.

We find that the resulting density profiles are well fit by a cored-Hernquist profile, defined as

$$\rho(r) = \frac{M}{2\pi}\frac{a}{(r^\beta + r_{\rm c}^\beta)^{1/\beta}}\frac{1}{(r+a)^3}, \tag{4.22}$$

where $r_{\rm c}$ is the core-radius, and β controls how sharply the density profile transitions from $\rho \propto 1/r$ to the constant density core. Examples of cored-Hernquist fits to radial density profiles are given in Fig. 4.9.

By finding the best fit cored-Hernquist profile for a number of time-steps in our simulations, we can see how the core size evolves over time. We fixed $\beta = 4$ when fitting the cored profiles. The fits achieved when fixing β in this way are not as good as when β is allowed to vary (see Fig. 4.9), but this process enables us to describe the evolution of the core in terms of only a single number. This was done for a variety of cross-sections, and is plotted in Fig. 4.10. Following Kochanek & White (2000, hereafter KW00) the plot is made in terms of the dimensionless quantities $r_{\rm c}/a$, $t/t_{\rm dyn}$, and $\hat{\sigma}$, where the dynamical time for a Hernquist profile is defined as $t_{\rm dyn} = 4\pi(a^3/GM)^{1/2}$, and the dimensionless cross-section as $\hat{\sigma} = M(\sigma/m)/a^2$. Plotting the core sizes in this dimensionless way allows the results to be applied to systems with different masses and concentrations.

We ran the same dimensionless cross-sections as KW00, with results that agreed except

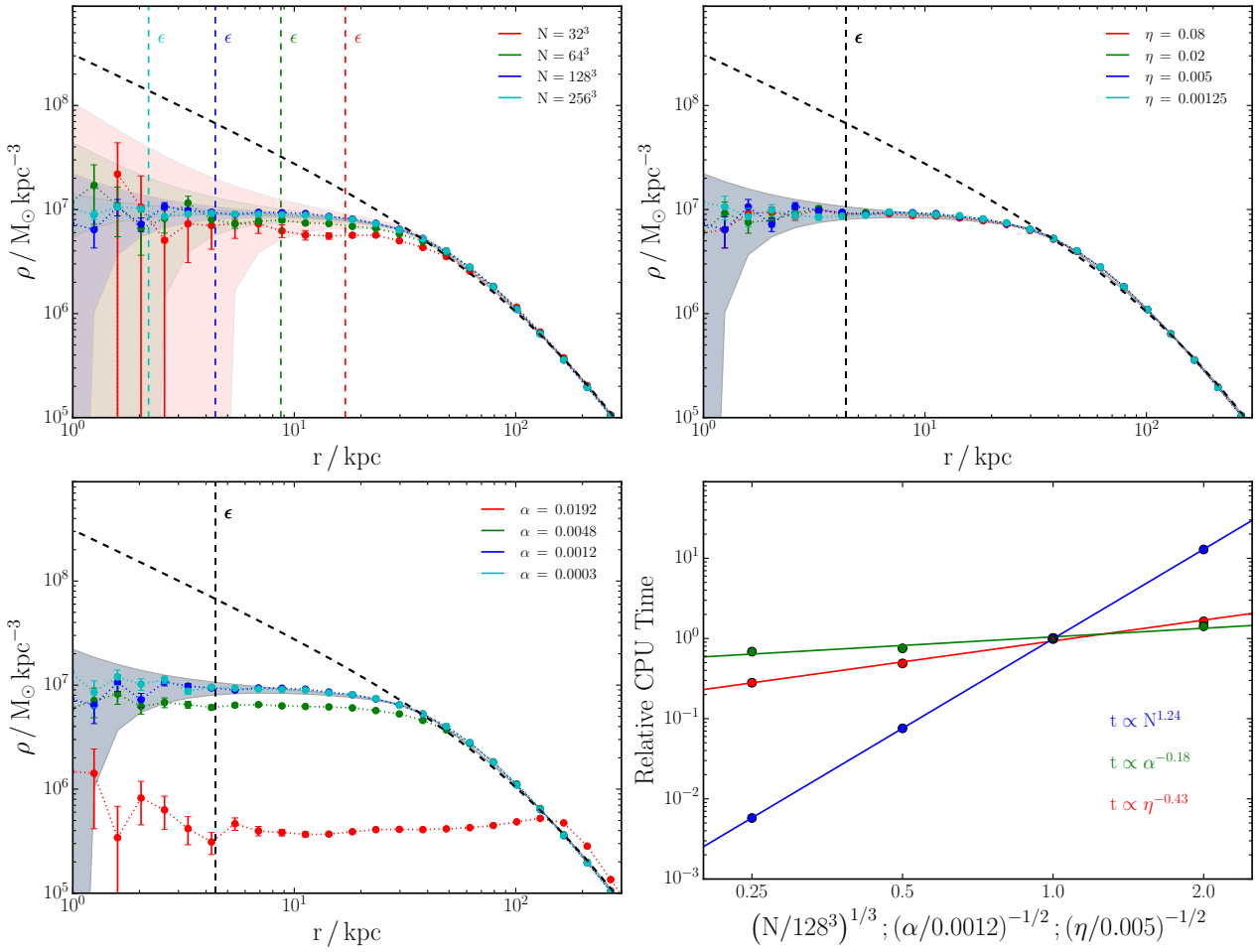


Figure 4.8: Radial density profiles after 10 Gyr from simulations with $\sigma/m = 1\,\mathrm{cm}^2\,\mathrm{g}^{-1}$. As in Fig. 4.7, the simulations were run with different numbers of particles, and different force and time-integration accuracy parameters, varying from a fiducial simulation with $N = 128^3$, $\eta = 0.005$ and $\alpha = 0.0012$. Shaded regions show the expected 1σ variations in the density of the best fit cored-Hernquist profile to our fiducial run.

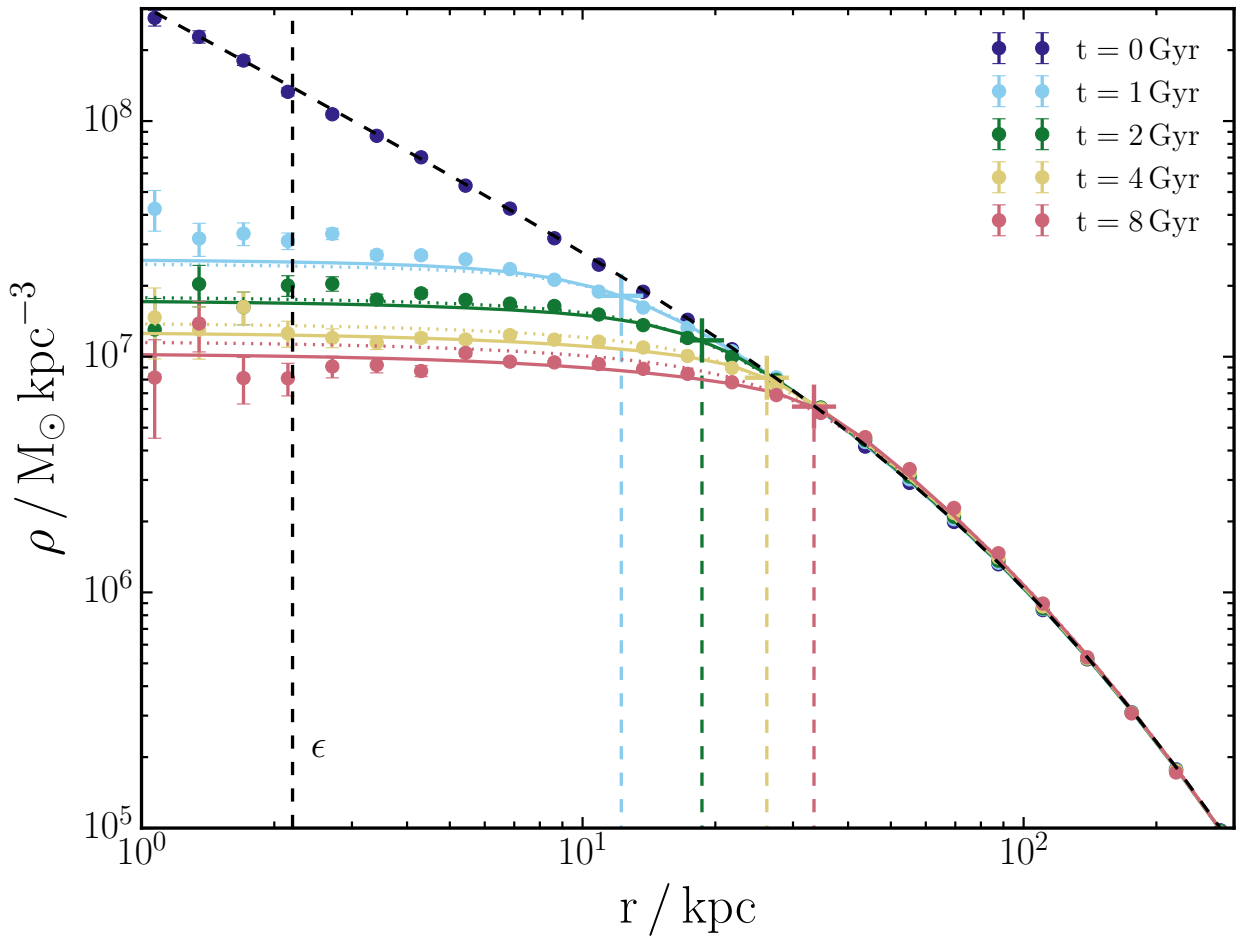


Figure 4.9: The radial density profile for a $10^{14}\,\mathrm{M}_\odot$ Hernquist halo, with $a = 225\,\mathrm{kpc}$. The halo was evolved with $\sigma/m = 1\,\mathrm{cm}^2\,\mathrm{g}^{-1}$. The density profile at different times after the beginning of the simulation is shown by the points, with error bars assuming that the number of particles in each radial bin is Poisson distributed. The best-fit cored Hernquist profile (equation 4.22) at each time has been plotted with a solid line, while the dashed lines are the best-fit cored Hernquist profiles when we fix $\beta = 4$. The coloured vertical dashed lines show the best fit core-radius (r_c) with a variable β.

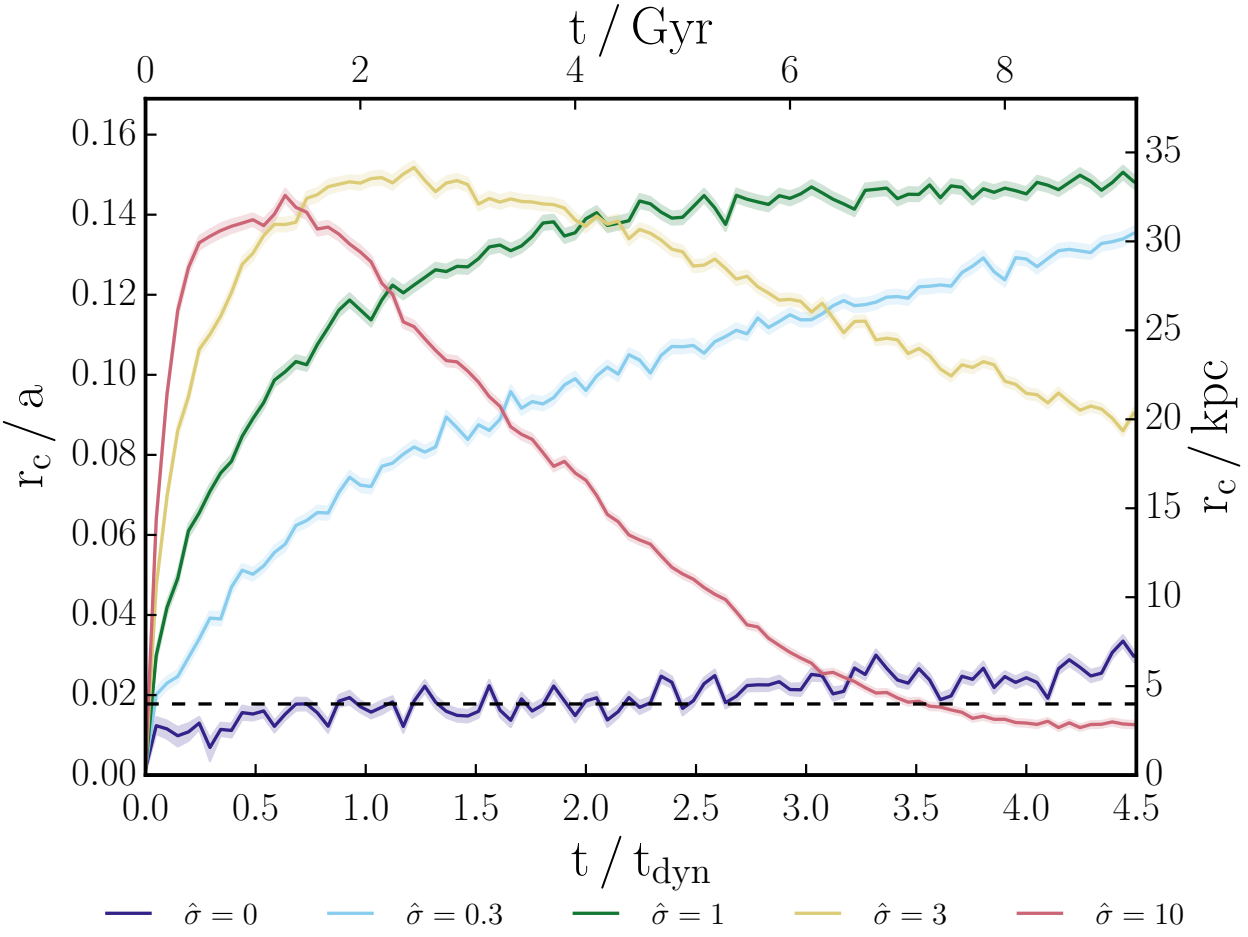


Figure 4.10: The evolution of core sizes starting with a Hernquist profile as the initial conditions. For the collisionless case (dark blue line), a core forms on the scale of the gravitational softening length. For the collisional cases, the core size grows to be $\sim 15\%$ of the scale radius. For the largest cross-sections, the cores then undergo core-collapse within a few dynamical times. For a $10^{14}\,\mathrm{M}_\odot$ halo with $a = 225\,\mathrm{kpc}$, the dynamical time is $2.05\,\mathrm{Gyr}$ and $\hat{\sigma} = 1$ corresponds to $\sigma/m = 2.42\,\mathrm{cm}^2\,\mathrm{g}^{-1}$.

for the normalisation of r_c/a. The core radii found in KW00 were considerably larger than those in Fig. 4.9 (by a factor of 3–4). This is due to their definition of the core size, as they estimate the core size as the radius at which the density drops to a quarter of the central density. As the core forms within a region where $\rho \propto 1/r$, the radius at which ρ has dropped to a quarter of the central density is four times the radius at which the density of the initial profile is equal to the central density of the cored profile, and hence roughly four times the r_c returned by our fitting procedure.

SIDM run-times

The CPU time for our $\sigma/m = 1\,\mathrm{cm}^2\,\mathrm{g}^{-1}$ simulations were measured in the same way as the collisionless ones, and plotted in the bottom-right of Fig. 4.8. The scaling with both η and α was very similar to the collisionless case, while the run-time increased more slowly with increasing particle number than for the collisionless simulations ($t_{\mathrm{CPU}} \propto N^{1.24}$ for SIDM versus $t_{\mathrm{CPU}} \propto N^{1.42}$ for CDM). We found that when running the simulation with the fiducial set of parameters, the SIDM simulation ran 30% slower than the collisionless version. As expected, the neighbour search and scattering procedure increase the CPU time required. However, the formation of large cores in the collisional simulations led to faster simulations over the course of the $\sim 10\,\mathrm{Gyr}$ for which they were run. This is because the dynamical times for particles in a cored profile are longer than for those in a cuspy profile, meaning that once a core has formed due to self-interactions, the time-steps used in the collisional case are longer than in the collisionless case (see the bottom-left panel of Fig. 4.6).

4.2.5 Core collapse

An interesting feature of the evolution of systems with large cross-sections is that after the initial formation of a core, the central density rises again. This is known as gravothermal collapse, and is a phenomenon seen in other gravitationally bound collisional systems, such as globular clusters. Gravothermal collapse in the context of globular clusters has been studied theoretically (Lynden-Bell & Eggleton, 1980) using a model that describes thermal

conduction. More recently, Balberg et al. (2002) applied a similar model to isolated SIDM haloes.

The differences between a system of stars and a system of collisional dark matter, primarily stem from the fact that stars undergo Rutherford scattering, which has strong angular and velocity dependence.[3] This leads to many interactions, each with a small amount of momentum transfer. If the DM cross-section is isotropic, then these systems are dominated by a small number of high momentum transfer events. These two scenarios lead to differences in the way energy is transported through the system. Another difference is that core collapse in globular clusters can be halted by energy injection from three-body encounters (Goodman & Hut, 1989). When a binary star system has a close encounter with a third star, it typically 'hardens', giving out energy to the third star. There is no equivalent of this with SIDM, such that collapse should proceed unchecked.

The density and velocity dispersion for a halo undergoing core collapse are plotted as a function of radius in Fig. 4.11. The velocity dispersion inside the scale radius becomes isothermal within a dynamical time due to the redistribution of energy by the collisions. This isothermal distribution increases in temperature, as core collapse leads to the system becoming more gravitationally bound.

The results for gravothermal collapse of a Hernquist profile from Monte Carlo N-body simulations and a conducting fluid model were compared by Koda & Shapiro (2011). They found that by tuning the two free parameters in their conducting fluid model,[4] they get good agreement between the two methods for both the density profiles and velocity dispersions. The increase in velocity dispersion at large radii seen in Fig. 4.11 does not appear to be present in their results, although they only plot the velocity dispersion out to ten scale radii (~ 2 Mpc in Fig. 4.11), where we just start to see an effect.

It is important to stress that for the $10^{14}\,\mathrm{M}_\odot$ haloes simulated here, core collapse within a Hubble time was only seen for the two largest cross-sections, corresponding to $7.3\,\mathrm{cm}^2\,\mathrm{g}^{-1}$ and $24.2\,\mathrm{cm}^2\,\mathrm{g}^{-1}$. These cross-sections are an order of magnitude above current upper-limits, and Ahn & Shapiro (2005) demonstrated (using the conducting fluid model) that

[3] The 'collisions' in globular clusters are close gravitational encounters, for which the differential cross-section is $\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} \propto (v \sin\frac{\theta}{2})^{-4}$.

[4] These two free parameters relate to the efficiency of heat conduction on large and small scales.

the accretion of new material onto a halo acts to stabilise haloes against core collapse. This suggests that currently viable SIDM models will not result in collapsed haloes within a Hubble time.

4.3 Scattering in a cosmological simulation

As a final test of our code implementation, we ran a small cosmological periodic box, with a Planck 2013 cosmology (Planck Collaboration et al., 2014) and a box-size of $25\,h^{-1}\,\mathrm{Mpc}$. The initial conditions were produced using the IC_GEN code described most recently in Jenkins (2013). The displacements and velocities were computed using second-order Lagrangian perturbation theory, using the method outlined in Crocce et al. (2006). The simulation was run at three different resolutions, corresponding to particle numbers of 64^3, 128^3 and 256^3, and with $\sigma/m = 1\,\mathrm{cm}^2\,\mathrm{g}^{-1}$.

In §3.4.5 we discussed how scattering is implemented within cosmological simulations. Given that GADGET uses comoving positions and canonical momenta as the internal variables for position and velocity, and that the scale factor is used to keep track of time, there are various conversion factors (primarily the scale factor a and Hubble parameter h) that enter into the calculation of scattering probabilities. To test that these factors have been correctly included, we measured the average rate of scattering per particle in the simulations, and compared the results with the analytical predictions for the scattering rate found in Chapter 2.

In Fig. 4.12 we show that there is reasonable agreement between the rates of scattering measured in our simulations and an analytical prediction. For this prediction we used an ST mass function, and the Diemer & Kravtsov (2015) concentration-mass relation. As demonstrated in Fig. 2.3, different concentration-mass relations predict rates of scattering that differ by around a factor of two. Some of the agreement in Fig. 4.12 is therefore due to our choice of concentration-mass relation, but the rough agreement between the simulation results and predictions act as a sanity check of both our SIDM implementation and the framework for analytically predicting $\Gamma(z)$.

The measured $\Gamma(z)$ from our three different-resolution simulations show that as the

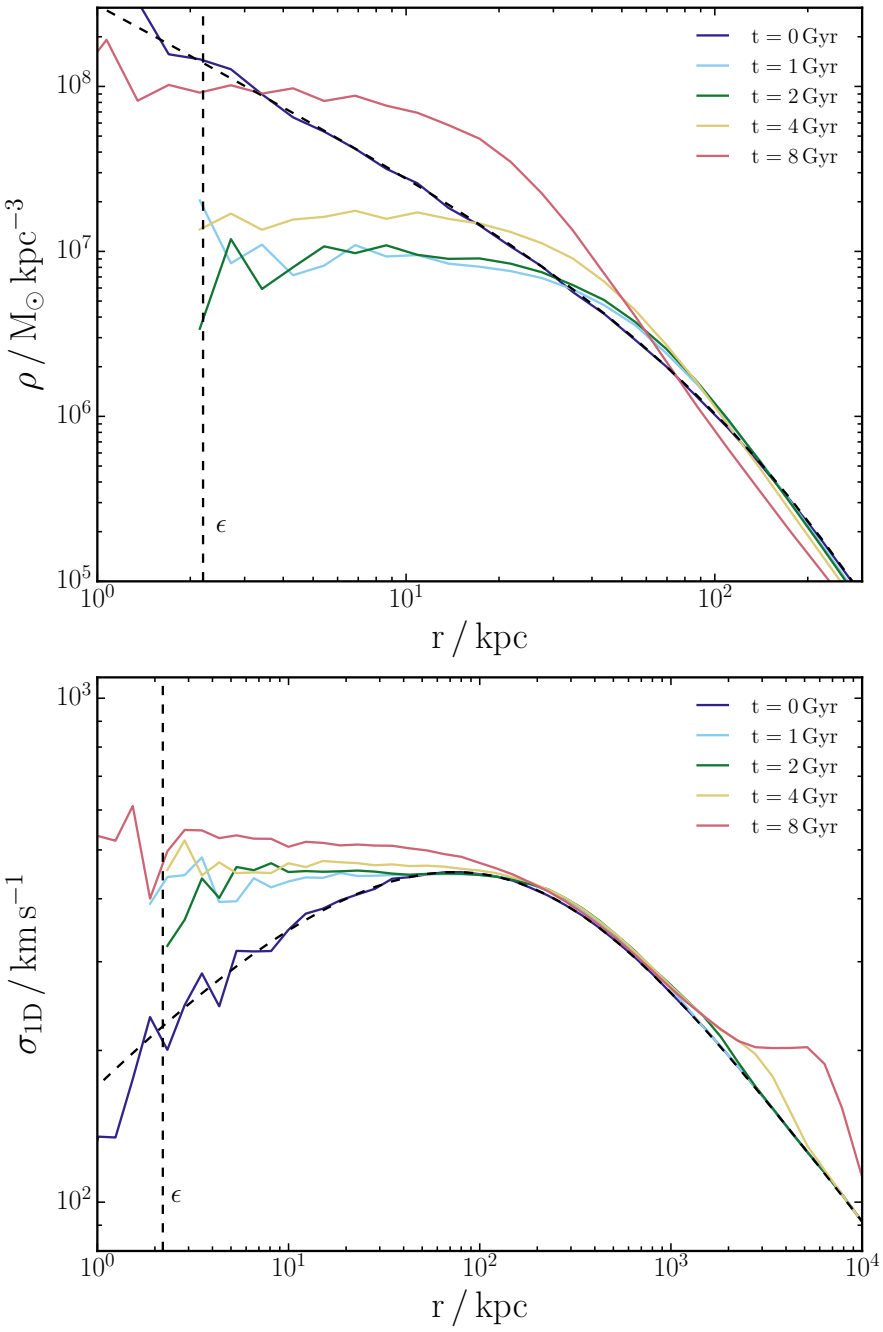


Figure 4.11: The density profile and velocity dispersion at different times in a simulation of a halo undergoing core collapse. The black dashed lines show the analytical profiles for the Hernquist halo used as initial conditions, while the solid lines show how these profiles evolve in time. The halo was the same one as in Fig. 4.10, with the cross-section corresponding to the $\hat{\sigma} = 10$ case. Initially the central density is lowered by DM scattering, but with a large cross-section the interactions transport heat so efficiently that the core undergoes gravothermal collapse and ends up extremely dense. This corresponds to the shrinking of r_c in Fig. 4.10, though the density profiles are not well described by cored-Hernquist profiles once core collapse has begun.

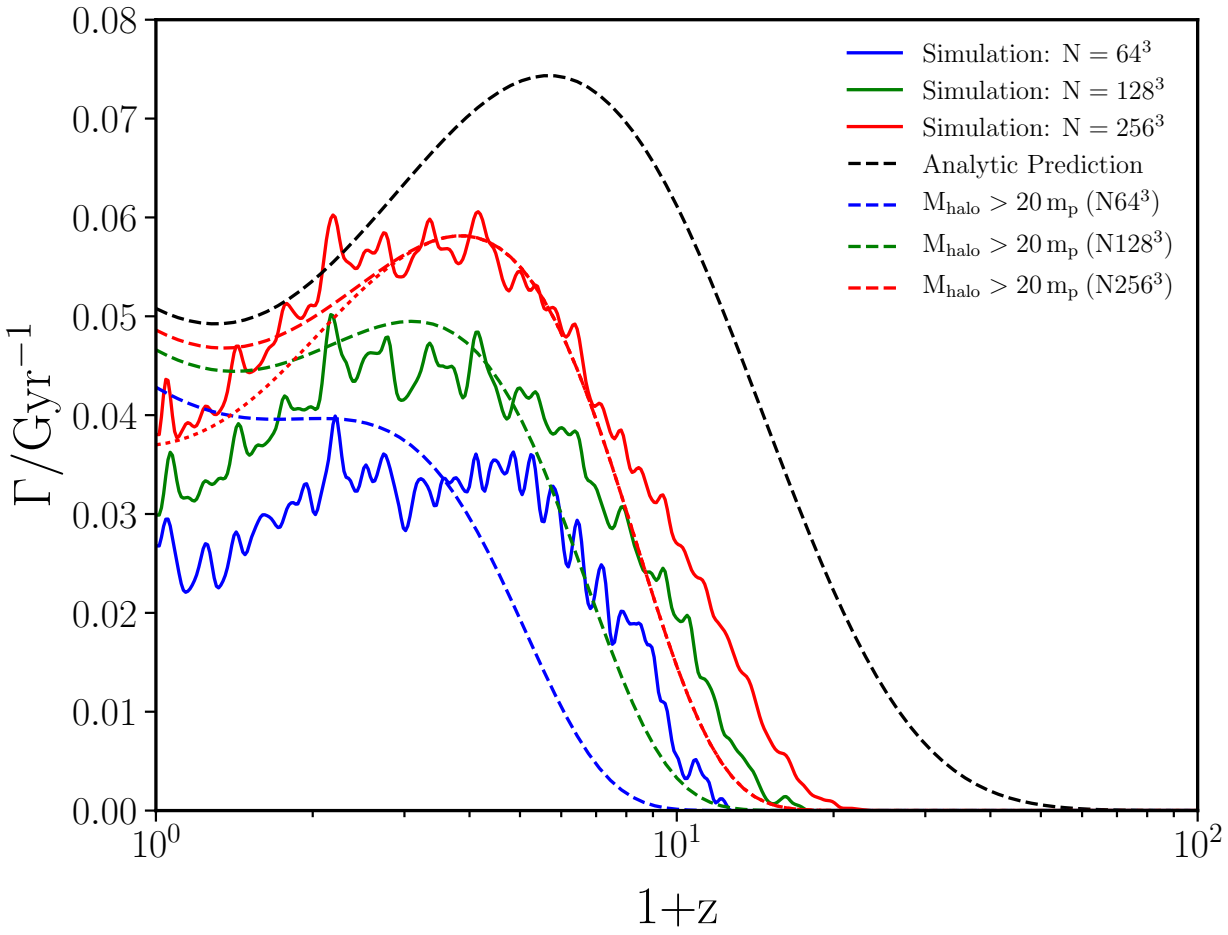


Figure 4.12: The scattering rate per particle in simulations of a $25\,h^{-1}\,\mathrm{Mpc}$ cosmological box, with a Planck 2013 cosmology (Planck Collaboration et al., 2014), and an SIDM cross-section of $\sigma/m = 1\,\mathrm{cm}^2\,\mathrm{g}^{-1}$. The simulations were run at three different mass resolutions, corresponding to particle numbers of 64^3, 128^3 and 256^3. The black dashed line shows the analytical predictions using an ST mass function and the Diemer & Kravtsov (2015) concentration-mass relation, following the method from Chapter 2. In our simulations we do not expect scattering in haloes that are not well resolved, and so we include the analytical prediction counting only those scattering events in haloes more massive than some minimum mass. These are drawn as the coloured dashed lines, with the minimum halo mass corresponding to $20m_\mathrm{p}$, with m_p the particle masses at our three resolutions. Finally, the red dotted line is the same as the red dashed line, but also not including scattering in haloes with $M_{200} > 10^{14}\,\mathrm{M}_\odot$.

resolution increases, the rate of scattering goes up. This is to be expected, because with increased resolution we resolve smaller-scale overdensities that host DM–DM scattering. The way in which $\Gamma(z)$ changes in response to changing resolution can be approximately reproduced by modifying the analytical prediction to only include scattering in haloes more massive than twenty times the particle masses used in our simulations.

Finally, we mention that the simulation results drop below the analytical predictions for $z \lesssim 1$. By inspecting the bottom panel of Fig. 2.1, we can see that at low redshift a significant contribution to $\Gamma(z)$ comes from massive ($> 10^{14}\,\mathrm{M}_\odot$) DM haloes. Given the relatively small box size of our simulations, we do not expect massive haloes to form. In Fig. 4.12 we therefore show, for one of the analytical prediction lines, the result if we also ignore scattering in objects more massive than $10^{14}\,\mathrm{M}_\odot$. This is largely illustrative, but the good agreement between this line and the result from the corresponding simulation is reassuring.

Chapter 5

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

5.1 Introduction

The massive galaxy cluster 1E 0657–56 (the 'Bullet Cluster') acts as a 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).

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 $1000\,\mathrm{km\,s^{-1}}$, 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 et al., 2012; Tulin et al., 2013a), and constraining such models requires a handle on the cross-section at different velocity scales (Kaplinghat et al., 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% 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%. This suggests that the significance of the $\sigma/m < 0.7\,\mathrm{cm^2\,g^{-1}}$ result derived in R08 is over-stated, 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 merg-

ers 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 $\sigma/m < 0.47\,\mathrm{cm}^2\,\mathrm{g}^{-1}$ for the DM elastic scattering cross-section.

In this chapter 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). In this chapter 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.

This chapter is structured as follows. In §5.2 we describe the initial conditions we use for our simulations. In §5.3 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 §5.4. Finally, we give our conclusions in §5.5. 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 ($z = 0.296$) $1\,\mathrm{kpc}$ corresponds to 0.23 arcsec.

5.2 Simulation 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 ide-

alised 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 idealised 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.

5.2.1 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 fit 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.[1]

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 $\sim 100\,\mathrm{kpc}$, 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 over-predicting 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.

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

$$\rho(r) = \frac{M}{2\pi}\frac{a}{r}\frac{1}{(r+a)^3}. \tag{5.1}$$

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. We matched our NFW halo parameters to Hernquist profiles using the method discussed in §4.2. This method matches the normalisation of the density in the central regions, for which $\rho \propto r^{-1}$ for both NFW and Hernquist profiles, as well as the mass within r_{200}. With a mass, M, and scale radius, a, for each of our two haloes, we generated equilibrium Hernquist profile initial conditions using the method from §4.2.1.

5.2.2 Relative velocity of the DM haloes

The relative velocity between the two DM haloes in the Bullet Cluster was originally estimated to be $4700\,\mathrm{km\,s^{-1}}$, 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 $4\,\mathrm{Mpc}$ 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.

5.2.3 Summary of initial conditions

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

The mass within each halo is 99% DM, and 1% 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 §5.4.4. The gas initially follows the same density profile as the DM and stars, with the halo mass being 83% DM, 16% gas, and 1% 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 $8.4\,\mathrm{keV}$, in agreement with the temperature of the pre-shocked gas in the Bullet Cluster (Markevitch, 2006).

5.2.4 Comparison to other SIDM studies

In Fig. 5.1 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 $\sigma/m = 1\,\mathrm{cm}^2\,\mathrm{g}^{-1}$ passing through the main halo of our fiducial model at a projected radius of $200\,\mathrm{kpc}$, where the projected surface density is $\sim 0.15\,\mathrm{g\,cm}^{-2}$, would have a $\sim 15\%$ chance of scattering off a particle in the main halo.

In the right-hand panel of Fig. 5.1 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. 5.1 to calculate the fraction of particles in the bullet halo that scatter with a particle from the main halo. For our fiducial model with $\sigma/m = 1\,\mathrm{cm}^2\,\mathrm{g}^{-1}$, we expect $\sim$ 23 (33)% 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)% respectively. Considering all particles in the bullet halo, the number goes down to 12% for our fiducial model, in good agreement with the value of 13% that we get in our simulations (see §5.4.3).

Stability of an isolated halo

In Fig. 5.2 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 $4\,\mathrm{Mpc}$ results in core passage taking place $\sim 1.1\,\mathrm{Gyr}$ 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 im-

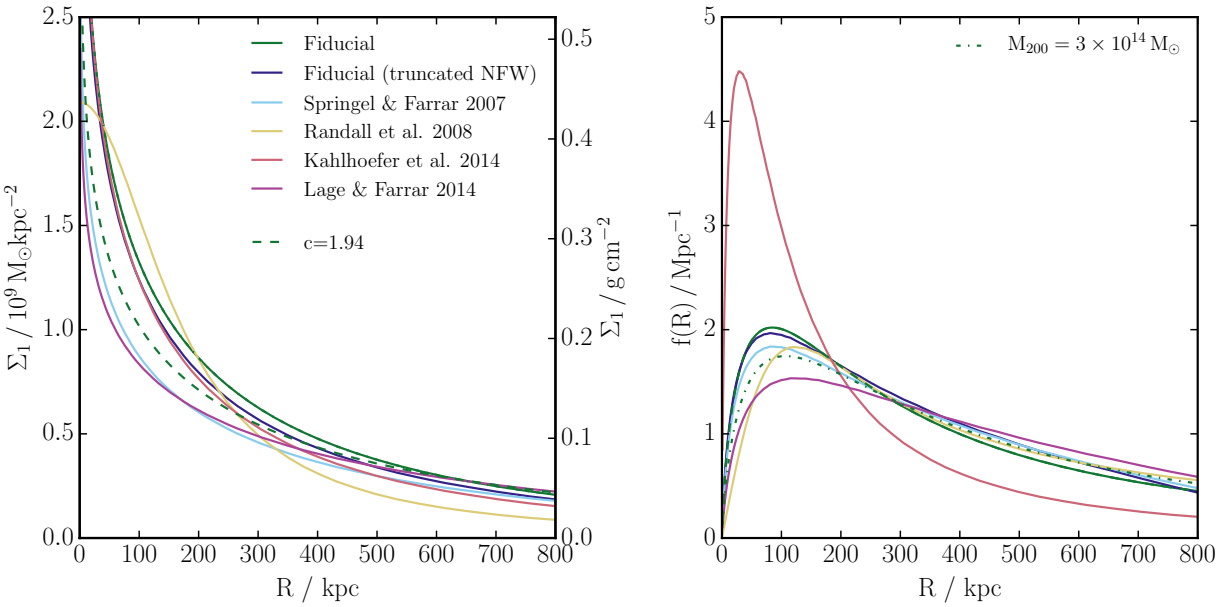


Figure 5.1: Left panel: the projected mass density through the main halo of the Bullet Cluster, as a function of projected radius. Right panel: the distribution of DM mass at different projected radii in the subcluster of the Bullet Cluster, normalised so that $\int_0^{1\,\mathrm{Mpc}} f(R)\mathrm{d}R = 1$. Different line styles and colours correspond to different choices for the density profiles. Our fiducial model is described in §5.2.3 while two variations to our fiducial model ($c = 1.94$ and $M_{200} = 3 \times 10^{14}\,\mathrm{M}_\odot$) are described in §5.4.2. 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}. Springel & Farrar 2007 was the fiducial model used in that paper, while Randall et al. 2008 is for the density profiles used in their simulation with $\sigma/m = 1.25\,\mathrm{cm}^2\,\mathrm{g}^{-1}$ (their initial conditions were changed slightly for different cross-sections). Kahlhoefer et al. 2014 only simulated one model for the Bullet Cluster, which had a particularly concentrated bullet halo as evident in the right panel. Lage & Farrar 2014 was the best-fit model found from running over a thousand simulations with different initial conditions and comparing the results to several observational datasets.

pact on our results, we experimented with a different initial separations between the two haloes. Starting the haloes with a separation of $9\,\mathrm{Mpc}$, the haloes have evolved for $3.4\,\mathrm{Gyr}$ before they collide. We found this only had a small impact on our results, changing the best-fit separation between stars and DM (with a scattering cross-section of $1\,\mathrm{cm}^2\,\mathrm{g}^{-1}$) from $9.2\,\mathrm{kpc}$ to $8.4\,\mathrm{kpc}$ at the time of the observed Bullet Cluster in our fiducial model. This change is small compared to the effects discussed in §5.4.

5.3 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.

5.3.1 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_{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 $R_{\mathrm{min}} = 200\,\mathrm{kpc}$. 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

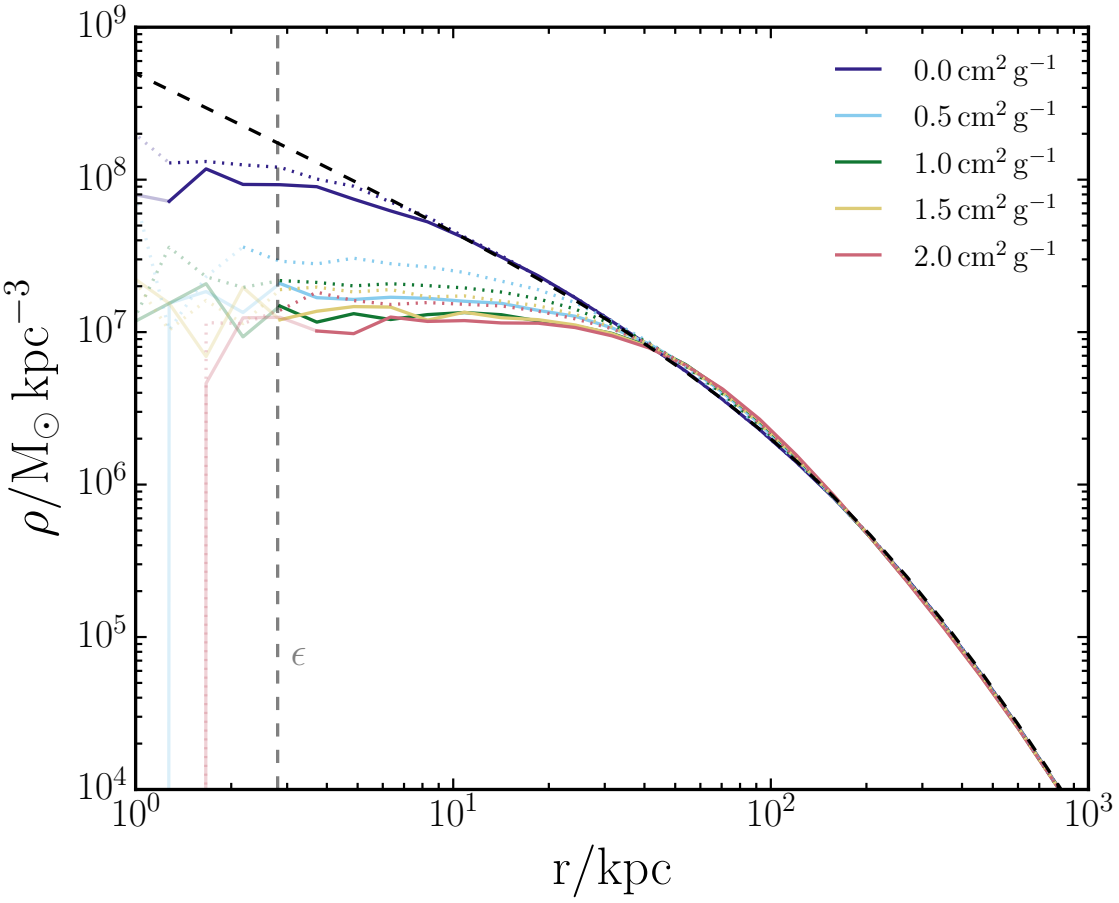


Figure 5.2: The radial density profile of an isolated halo with collisionless DM as well as SIDM with isotropic cross-sections ranging from 0.5 to $2\,\mathrm{cm}^2\,\mathrm{g}^{-1}$. The dotted lines show the average profile between 1 and 2 Gyr after the start of the simulation, while the solid lines show the period 5–6 Gyr after the start of the simulation. Lines are semi-transparent 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.

particles belonging to the different haloes, or be started with the circles already shrunk to a size where they only contain one DM peak.

5.3.2 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

$$\rho(r) = \frac{\rho_0}{(1 + r^2/r_{\rm core}^2)(1 + r^2/r_{\rm cut}^2)}; \quad r_{\rm cut} > r_{\rm core}. \tag{5.2}$$

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

$$\begin{aligned} \Sigma(R) &= 2\int_R^\infty \frac{\rho(r) r}{\sqrt{r^2 - R^2}}\,{\rm d}r \\ &= \Sigma_0 \frac{r_{\rm core}\, r_{\rm cut}}{r_{\rm cut} - r_{\rm core}} \left(\frac{1}{\sqrt{r_{\rm core}^2 + R^2}} + \frac{1}{\sqrt{r_{\rm cut}^2 + R^2}} \right), \end{aligned} \tag{5.3}$$

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

$$\Sigma_0 = \pi \rho_0 \frac{r_{\rm core}\, r_{\rm cut}}{r_{\rm cut} + r_{\rm core}}. \tag{5.4}$$

As described in Kassiola & Kovner (1993), the axially symmetric projected density profile in equation (5.3) can be made elliptical by substituting $R \to \tilde{R}$, where

$$\tilde{R}^2 = \frac{\tilde{x}^2}{(1+\epsilon)^2} + \frac{\tilde{y}^2}{(1-\epsilon)^2}, \tag{5.5}$$

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 semi-major and semi-minor 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 (5.3) 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 (5.5). 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:

$$\mathcal{L}(\theta = \{X, Y, \rho_0, r_{\rm core}, r_{\rm cut}, \phi, \epsilon\}) = \prod_i \frac{m_i^{d_i}\, \mathrm{e}^{-m_i}}{d_i!}. \tag{5.6}$$

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 (MCMC) 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. 5.3 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. 5.3 shows visually the level of noise associated with having a discrete set of particles and using $20\,{\rm kpc}$ 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 (5.6) 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 $\sim 2\,{\rm kpc}$.

5.3.3 Parametric fits to shear maps

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

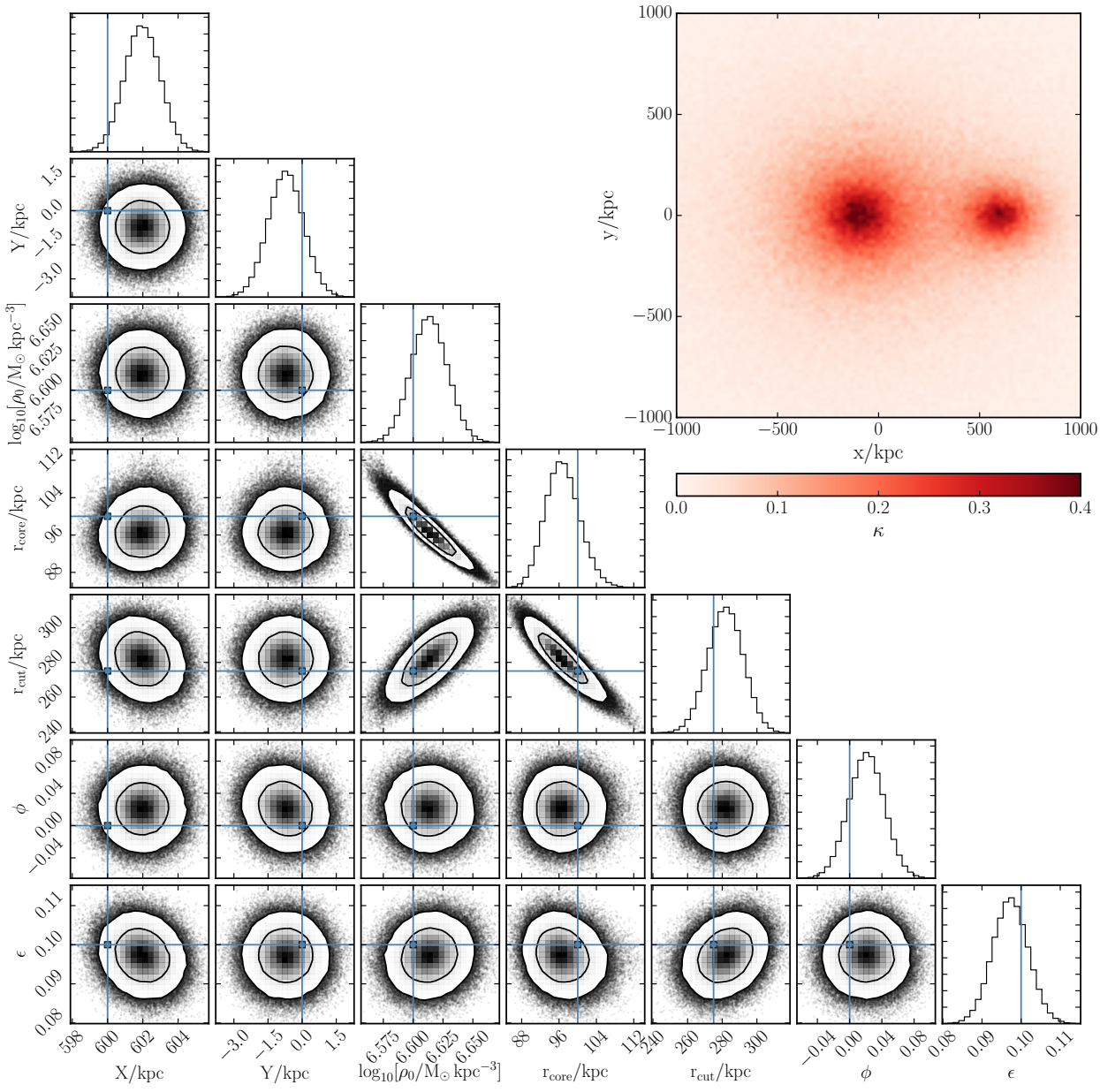


Figure 5.3: The posterior distributions for the model parameters of the smaller halo (on the right in the convergence map), found from simultaneously fitting two PIEMDs to the projected density generated from two model PIEMDs. The contours show 68 and 95% confidence intervals. The model values used to 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).

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

$$\kappa = \frac{1}{2}\left(\frac{\partial^2\Psi}{\partial x^2} + \frac{\partial^2\Psi}{\partial y^2}\right), \tag{5.7}$$

$$\gamma_1 = \frac{1}{2}\left(\frac{\partial^2\Psi}{\partial x^2} - \frac{\partial^2\Psi}{\partial y^2}\right), \tag{5.8}$$

and

$$\gamma_2 = \frac{\partial^2\Psi}{\partial x \partial y}. \tag{5.9}$$

The convergence is also given by the scaled projected density

$$\kappa(x,y) = \frac{\Sigma(x,y)}{\Sigma_{\rm crit}}, \tag{5.10}$$

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

$$\Sigma_{\rm crit} = \frac{c^2}{4\pi G}\frac{D_{\rm s}}{D_{\rm l} D_{\rm ls}}, \tag{5.11}$$

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 (5.10) and (5.11), 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 $\Sigma_{\rm crit} = 2.85 \times 10^9\ {\rm M}_\odot\,{\rm kpc}^{-2}$.

Once we have a κ map, we can generate maps of γ_1 and γ_2 by making use of equations (5.7 - 5.9). Taking the Fourier transform of these equations, we find

$$\hat{\kappa} = -\frac{1}{2}(k_x^2 + k_y^2)\hat{\Psi}, \tag{5.12}$$

$$\hat{\gamma_1} = -\frac{1}{2}(k_x^2 - k_y^2)\hat{\Psi}, \tag{5.13}$$

and

$$\hat{\gamma_2} = -k_x k_y \hat{\Psi}, \tag{5.14}$$

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

$$\hat{\gamma_1} = \frac{k_x^2 - k_y^2}{k_x^2 + k_y^2}\hat{\kappa}, \tag{5.15}$$

and

$$\hat{\gamma_2} = \frac{2k_x k_y}{k_x^2 + k_y^2}\hat{\kappa}. \tag{5.16}$$

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 $1 - \kappa$.

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

$$\mathcal{L}(\theta) = \prod_i \exp\left\{\left(\frac{(g^{\rm d}_{1,i} - g^{\rm m}_{1,i})^2}{2\sigma_\gamma^2}\right)\right\} \exp\left\{\left(\frac{(g^{\rm d}_{2,i} - g^{\rm m}_{2,i})^2}{2\sigma_\gamma^2}\right)\right\}, \tag{5.17}$$

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 §5.3.3. 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 (5.17), 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. 5.3 is shown in Fig. 5.4. 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. 5.4) 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 realisations 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. 5.4 suggest that using gravitational shear with 80 galaxies arcmin^{-2} we cannot determine the position of the bullet DM halo to better than $\sim \pm 40\,\mathrm{kpc}$. This is consistent with Harvey et al. (2015) who found a typical 1σ error of $60\,\mathrm{kpc}$ on the DM halo positions determined from weak gravitational lensing with ~ 60 galaxies arcmin^{-2} (D. Harvey, private communication).

5.4 Results

5.4.1 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 §5.3.2. Doing this independently for the DM

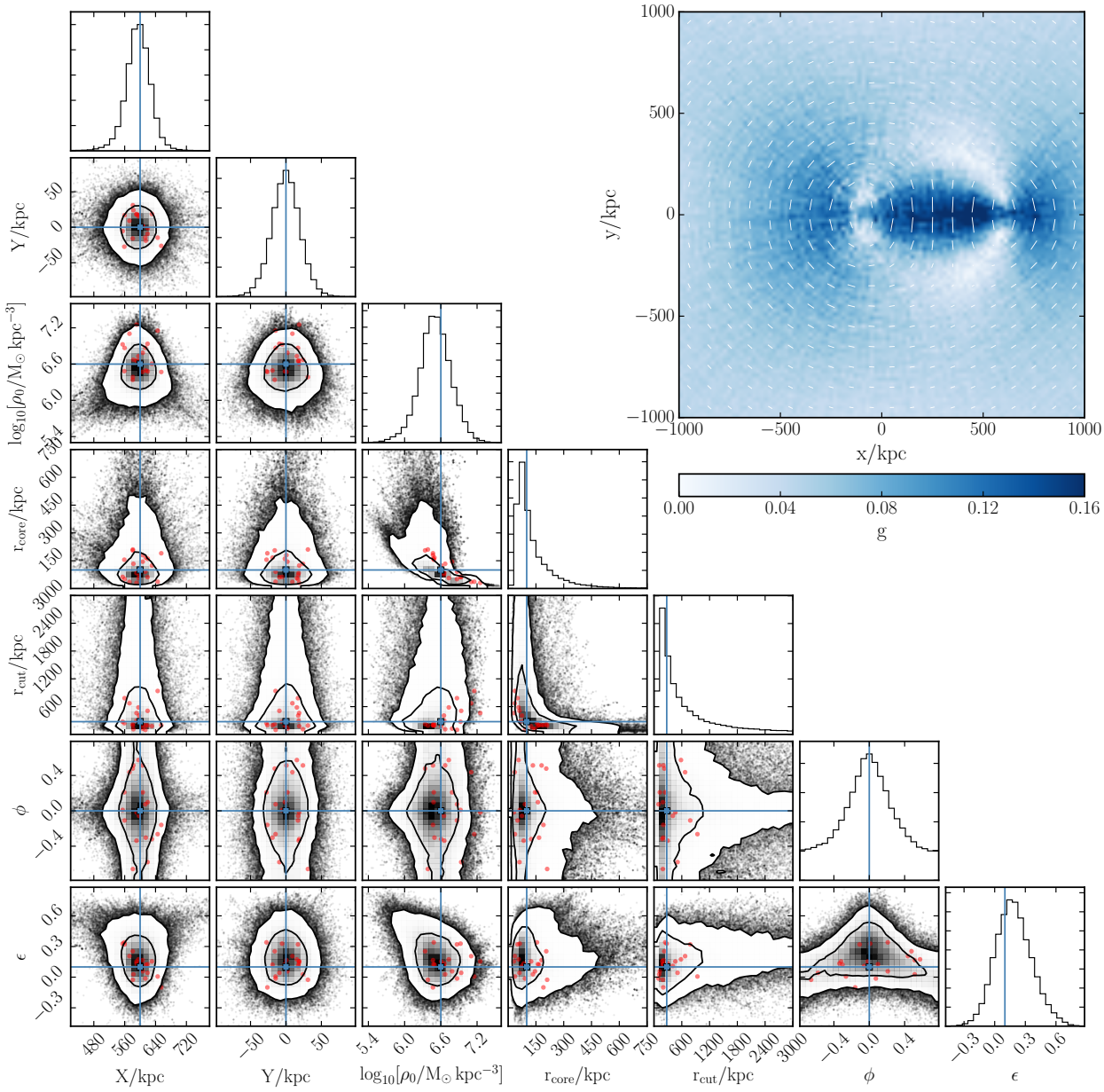


Figure 5.4: Similar to Fig. 5.3, but for the case of fitting to reduced gravitational shear. The red points indicate the maximum likelihood parameter values found from fitting to the underlying shear map from the model with the addition of 20 different realisations of noise from source-galaxy intrinsic 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.

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. 5.5. 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 $720\,\mathrm{kpc}$ occurs when the bullet halo is at $X_{\mathrm{DM}} \approx 600\,\mathrm{kpc}$.

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% smaller than in K14. For $\sigma/m = 1\,\mathrm{cm}^2\,\mathrm{g}^{-1}$ the offset at the time of the observed Bullet Cluster is $\sim 10\,\mathrm{kpc}$, 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 $25 \pm 29\,\mathrm{kpc}$ R08 placed constraints on the DM cross-section of $\sigma/m < 1.25\,\mathrm{cm}^2\,\mathrm{g}^{-1}$, 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.

5.4.2 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.

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

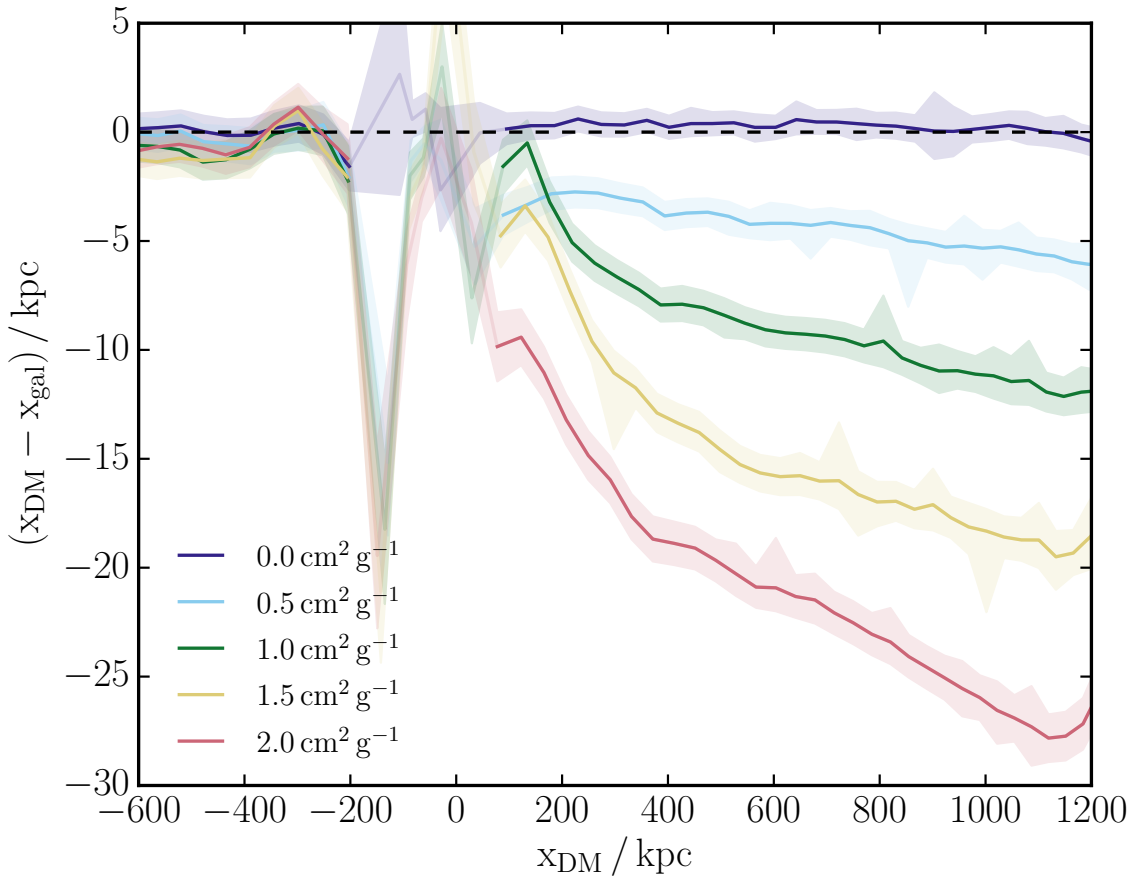


Figure 5.5: The offset between the stellar (galaxy) and DM component of the bullet halo for different SIDM cross-sections, where both the stellar and DM positions were determined by simultaneously fitting two PIEMDs to the respective projected maps. The offsets scale linearly with DM cross-section, and at the time of the observed bullet cluster the DM trails the galaxies by $\sim 10\,\mathrm{kpc}$ for $\sigma/m = 1\,\mathrm{cm}^2\,\mathrm{g}^{-1}$. Lines are semi-transparent around the time of core-passage (which due to tidal forces happens at $X_{\mathrm{DM}} \approx -80\,\mathrm{kpc}$ 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.

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 fit concentration for the main halo measured after the collision via weak lensing.

The resulting offsets with collisionless DM and SIDM with $\sigma/m = 1\,\mathrm{cm}^2\,\mathrm{g}^{-1}$ are shown in Fig. 5.6. 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 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 that scatter with particles from the main halo drops from 13% for our fiducial model to 10%, in broad agreement with estimates of the scattering fraction that can be made from Fig. 5.1.

Relative velocity between haloes

As discussed in §5.2.2 the shock velocity in the observed Bullet Cluster is $4700 \pm 600\,\mathrm{km\,s}^{-1}$. In previous work 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 is significantly higher than the relative velocity of the DM haloes (SF07, Milosavljević et al., 2007; Lage & Farrar, 2014).

In Fig. 5.6 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 $4\,\mathrm{Mpc}$, 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

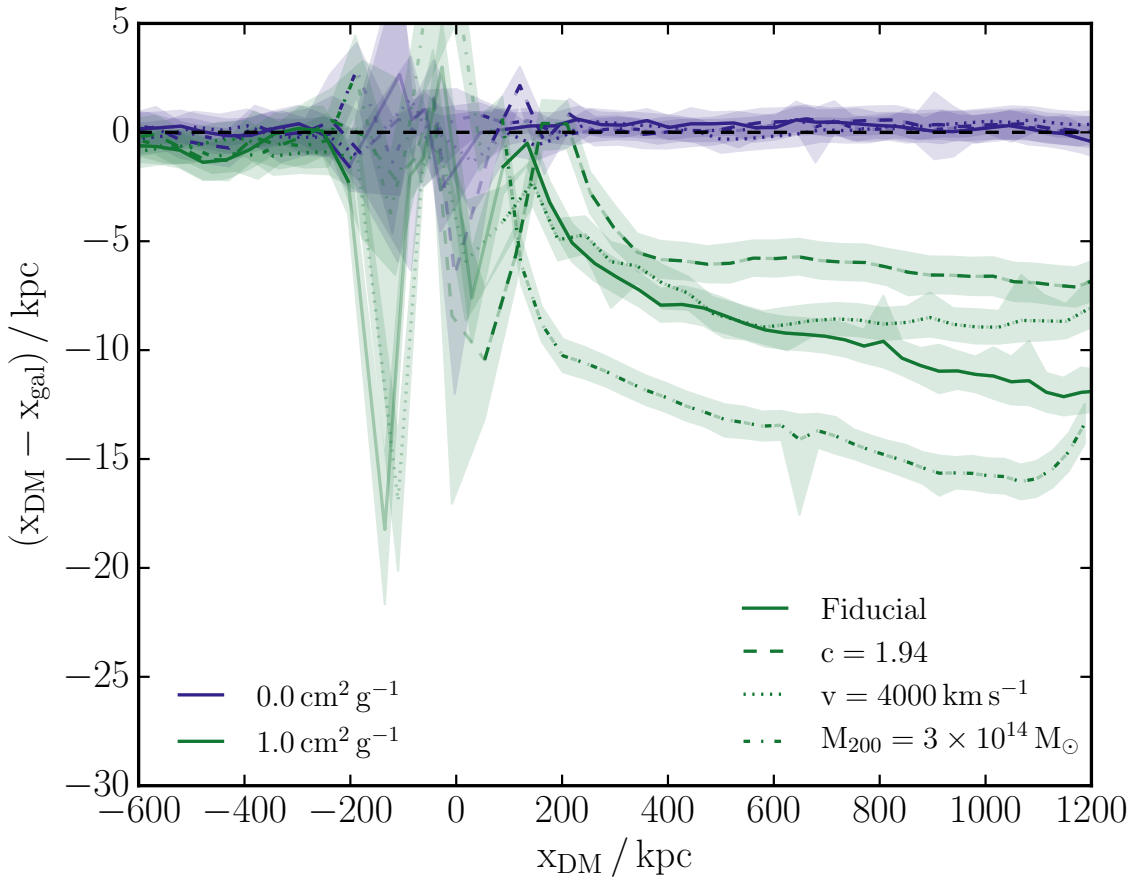


Figure 5.6: DM–galaxy offsets as in Fig. 5.5, but with four different sets of initial conditions, each run with collisionless DM and SIDM with $\sigma/m = 1\,\mathrm{cm}^2\,\mathrm{g}^{-1}$. These initial conditions are described in §5.4.2, but in summary are as follows: compared to our fiducial model '$c = 1.94$' has a lower concentration for the main halo, '$v = 4000\,\mathrm{km\,s}^{-1}$' has an increased relative velocity between the two haloes, and '$M_{200} = 3 \times 10^{14}\,\mathrm{M}_\odot$' has a more massive bullet halo.

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.

Mass of bullet halo

The weak lensing derived mass for the bullet halo of $M_{200} = 1.5 \times 10^{14}\,\mathrm{M_\odot}$ is low in comparison with the strong lensing results (Bradač et al., 2006) that calculate the mass in a $300\,\mathrm{kpc}$ radius cylinder centred on the bullet halo's BCG to be $3 \times 10^{14}\,\mathrm{M_\odot}$. 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 fit 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 $M_{200} = 3 \times 10^{14}\,\mathrm{M_\odot}$, 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 $10\,\mathrm{kpc}$ for our fiducial model to $14\,\mathrm{kpc}$.

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 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 $4\,\mathrm{Mpc}$ and on a zero energy orbit, but rotate the velocities of the haloes by θ_{init} with respect to the x-axis that connects the two halo centres (keeping the velocities of the two haloes anti-parallel). We choose θ_{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

Table 5.1: Summary of non-zero impact parameter simulations.

$r_*/\,\mathrm{kpc}$	$\theta_{\mathrm{init}}/^\circ$	$b_{4\,\mathrm{Mpc}}/\,\mathrm{kpc}$	$r_{\mathrm{min}}/\,\mathrm{kpc}$	$\theta_{\mathrm{obs}}/^\circ$
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

minimum separation between the halo centres, r_{min}, is significantly larger than r_*.

We summarise our different impact parameter runs in Table 5.1. As well as r_*, θ_{init}, and r_{min}, we include the perpendicular distance between the two haloes' velocities when they are separated by $4\,\mathrm{Mpc}$, $b_{4\,\mathrm{Mpc}}$, and the angle between the halo-halo separation and the bullet halo velocity at the time of the observed Bullet Cluster, θ_{obs}. Assuming that gas is stripped in the opposite direction to the direction of motion, θ_{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. 5.7 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 $r_* < 12.5\,\mathrm{kpc}$ to avoid a gas distribution that is more asymmetric than that observed, while Mastropietro & Burkert (2008) found that an impact parameter, $b = 150\,\mathrm{kpc}$, 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%.

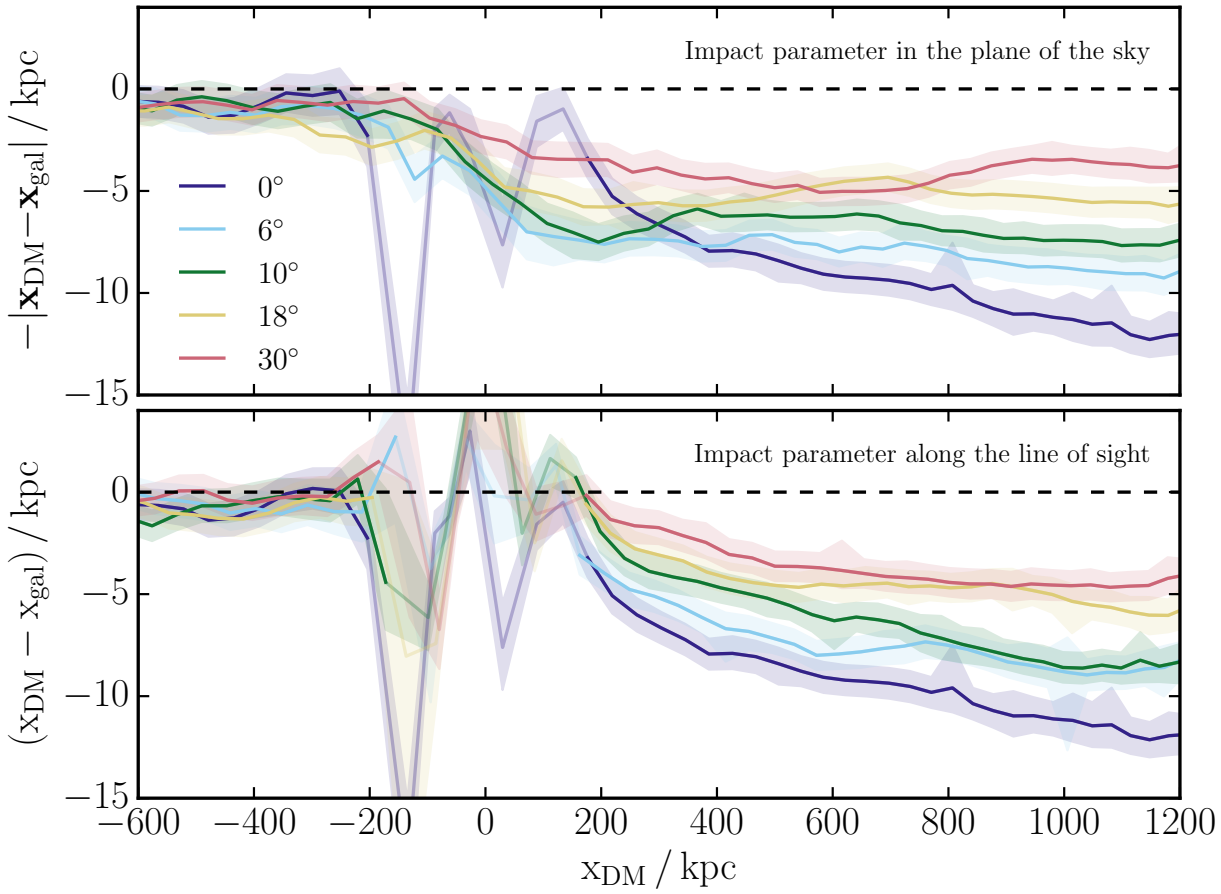


Figure 5.7: DM-galaxy offsets with $\sigma/m = 1\,\mathrm{cm}^2\,\mathrm{g}^{-1}$ and four different impact parameters, as well as a head-on collision. The runs are labelled by the angle between the separation of the two haloes and the velocity of the bullet halo measured at the time of the observed Bullet Cluster, $\theta_{\rm obs}$. 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.

5.4.3 Offsets with different position measures

Having found that the offset results are reasonably insensitive to the choice of initial conditions, in Fig. 5.8 we show the effect of using different methods to measure positions. For both collisionless DM and SIDM with $\sigma/m = 1\,\mathrm{cm}^2\,\mathrm{g}^{-1}$ we measured the separation between the stellar and DM components of the bullet halo using the methods described in §5.3. 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 measured by fitting to the projected density and the DM halo measured using reduced gravitational shear. Fig. 5.4 demonstrates that with 80 galaxies/arcmin^2 the position of the bullet halo can only be determined to $\pm 40\,\mathrm{kpc}$. 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. 5.8 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 §5.3.1 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 $R_{\mathrm{min}} = 200\,\mathrm{kpc}$ 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 $500\,\mathrm{kpc}$.

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 $\sigma/m = 1\,\mathrm{cm}^2\,\mathrm{g}^{-1}$ using different methods can be as different as the offsets for the different cross-sections shown in Fig. 5.5, 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.

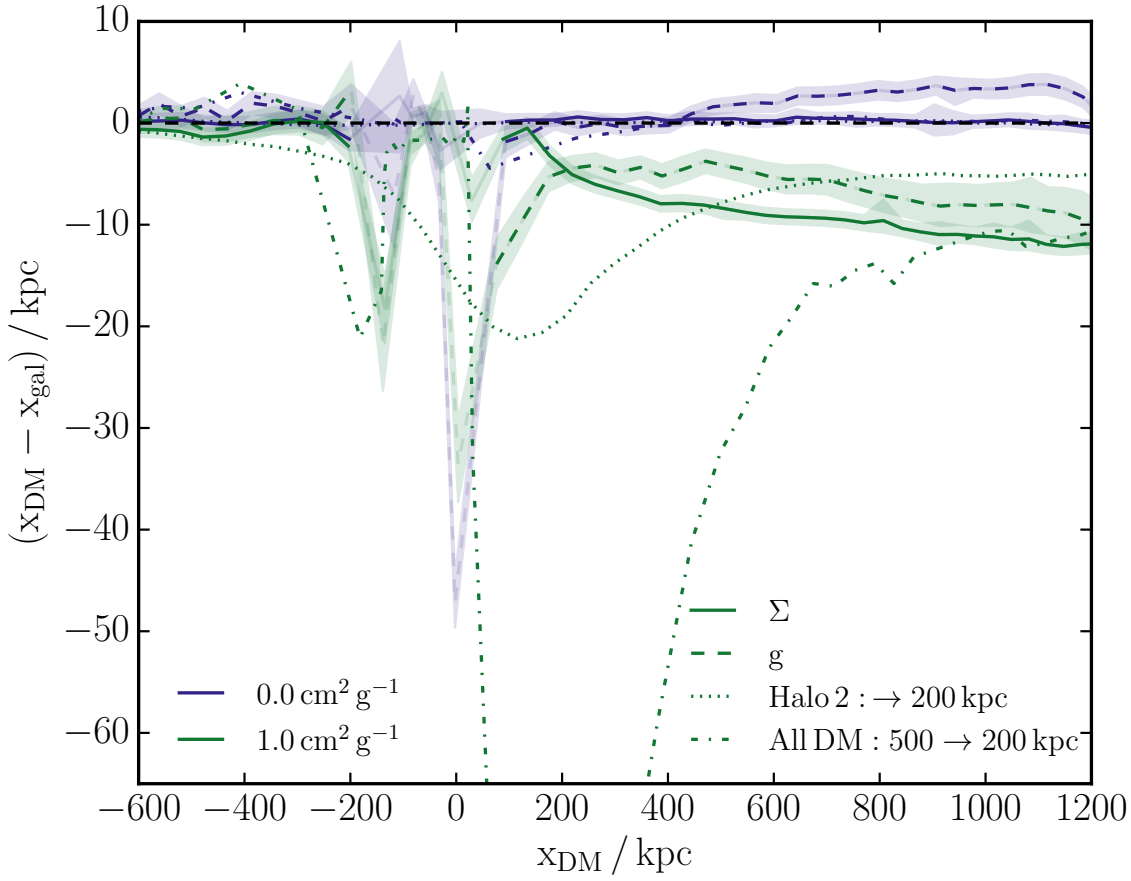


Figure 5.8: DM - galaxy offsets as in Fig. 5.5, but measured using different methods: fitting to the projected surface density (Σ), the reduced gravitational shear (g) and two different shrinking circles techniques. For all methods but g, the same method was used to find the position of both the stars 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 §5.4.3.

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 the other halo.[2] 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 un-scattered 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 $\sigma/m = 1\,\mathrm{cm}^2\,\mathrm{g}^{-1}$ we show the projected DM density at the time of the observed Bullet Cluster in Fig. 5.9, 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% 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% of the total projected density at the same location.

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. 5.10, run on our fiducial simulation with $\sigma/m = 1\,\mathrm{cm}^2\,\mathrm{g}^{-1}$ at the time of the observed Bullet Cluster (the same snapshot used for Fig. 5.9). 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

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

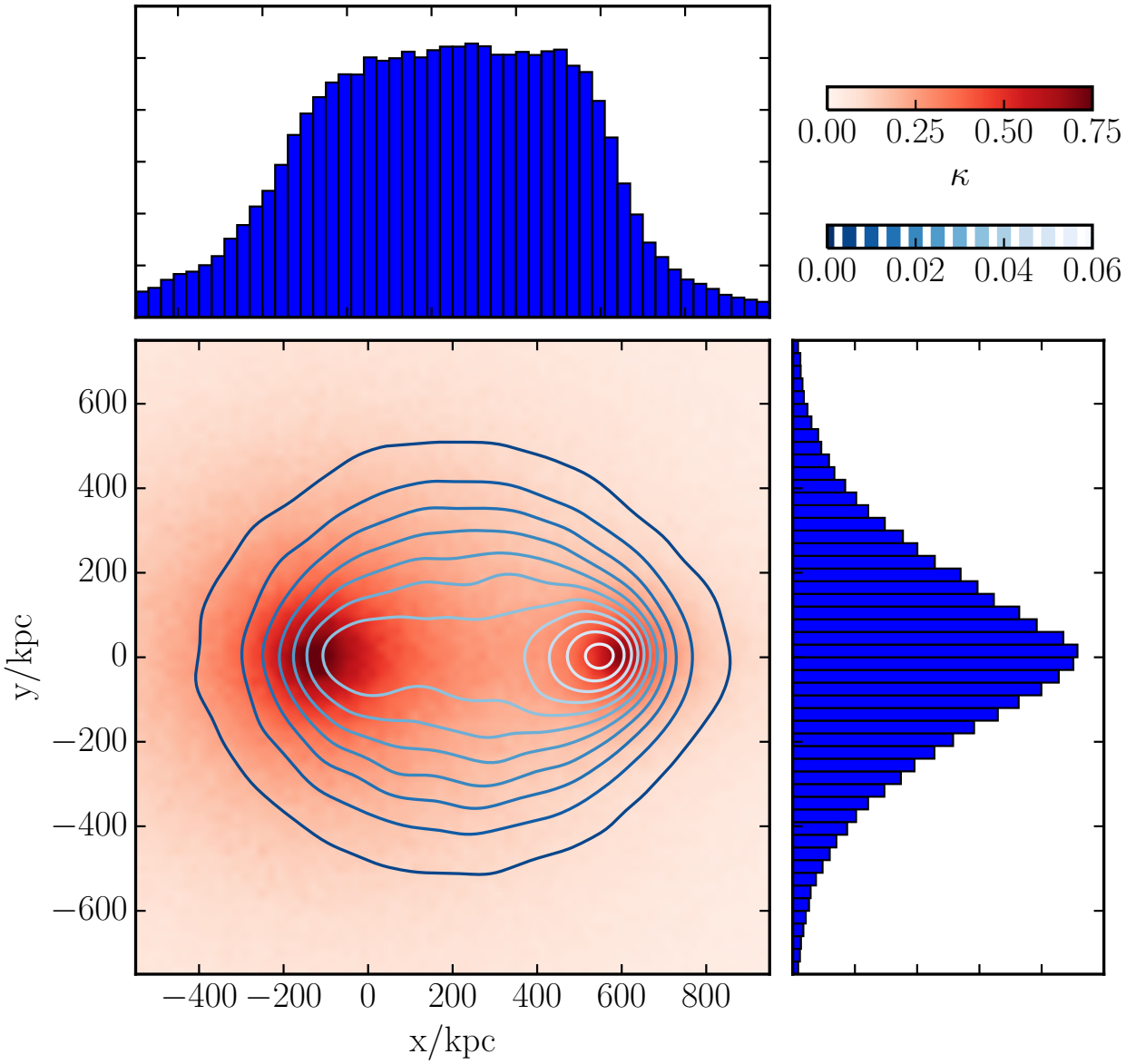


Figure 5.9: For SIDM with $\sigma/m = 1\,\mathrm{cm}^2\,\mathrm{g}^{-1}$ the projected density of all DM in red, with the projected density of DM particles that have scattered with a particle originally from the other DM halo shown in blue contours, and also projected along the axes and shown as 1D histograms. The total mass in these scattered particles is $6.5 \times 10^{13}\,\mathrm{M}_\odot$, corresponding to 13% of particles from the bullet halo scattering with particles in the main halo.

conditions. The initial radius used was $400\,\mathrm{kpc}$, a bit over half of the separation between the two DM haloes. Initially as the circles are shrunk and re-centred 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 $R_{\mathrm{min}} \lesssim 50\,\mathrm{kpc}$ the results become noisy as the number of particles involved in the position estimate decreases, and there is no clear density peak (with $\sigma/m = 1\,\mathrm{cm}^2\,\mathrm{g}^{-1}$ the core size of the bullet halo is $\sim 100\,\mathrm{kpc}$, though this is less obvious in the top of Fig. 5.10 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 regularisation scheme (e.g. Bradač et al., 2006), such that the diffuse cloud of scattered particles (Fig. 5.9) could shift the derived DM peak back and lead to a measurable offset.

From the bottom panel of Fig. 5.10 it is clear that $R_{\mathrm{min}} = 200\,\mathrm{kpc}$ 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_{min}, as the results do not converge as R_{min} is decreased. For these reasons we fit parametric models to our haloes in this chapter, 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 $0.1 - 1\,\mathrm{cm}^2\,\mathrm{g}^{-1}$, 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 idealised initial conditions, our work is not suited to testing whether the presence of strong lensing features can constrain the DM cross-section.

5.4.4 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% of the total halo mass.

The resulting offsets between stars and DM are shown in Fig. 5.11. The offsets measured for $\sigma/m = 1\,\mathrm{cm}^2\,\mathrm{g}^{-1}$ 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 $\sim 16\%$ 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.

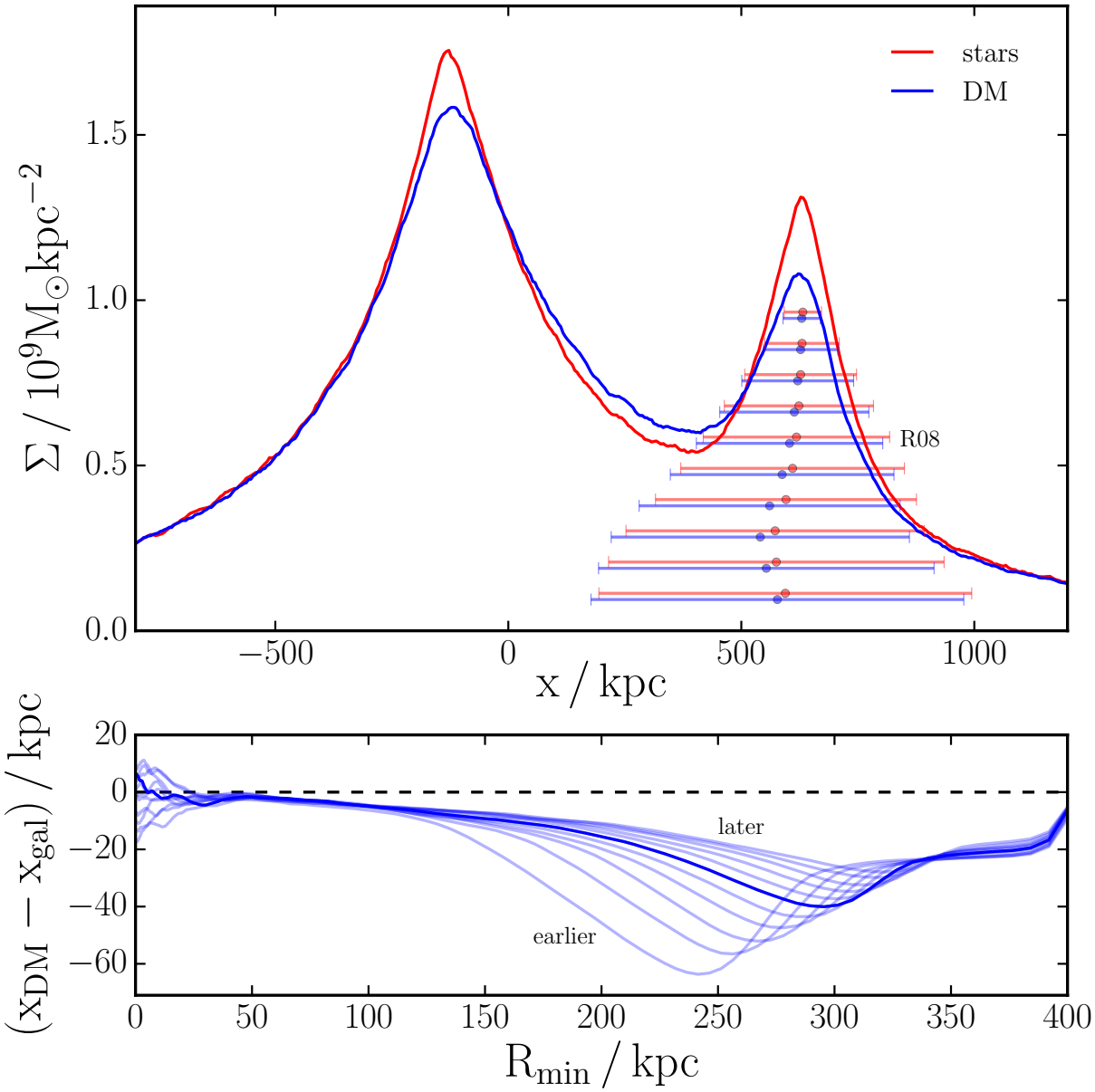


Figure 5.10: Top panel: the average projected surface density measured in a $400\,\mathrm{kpc}$ strip centred on the collision axis, for both the stars and DM (the stellar surface densities have been scaled up so that the mass in stars matches that in DM). The points show the position of the bullet halo 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 $\sim 720\,\mathrm{kpc}$, and with $\sigma/m = 1\,\mathrm{cm}^2\,\mathrm{g}^{-1}$), while the lighter lines are for successive snapshots separated by 10 Myr.

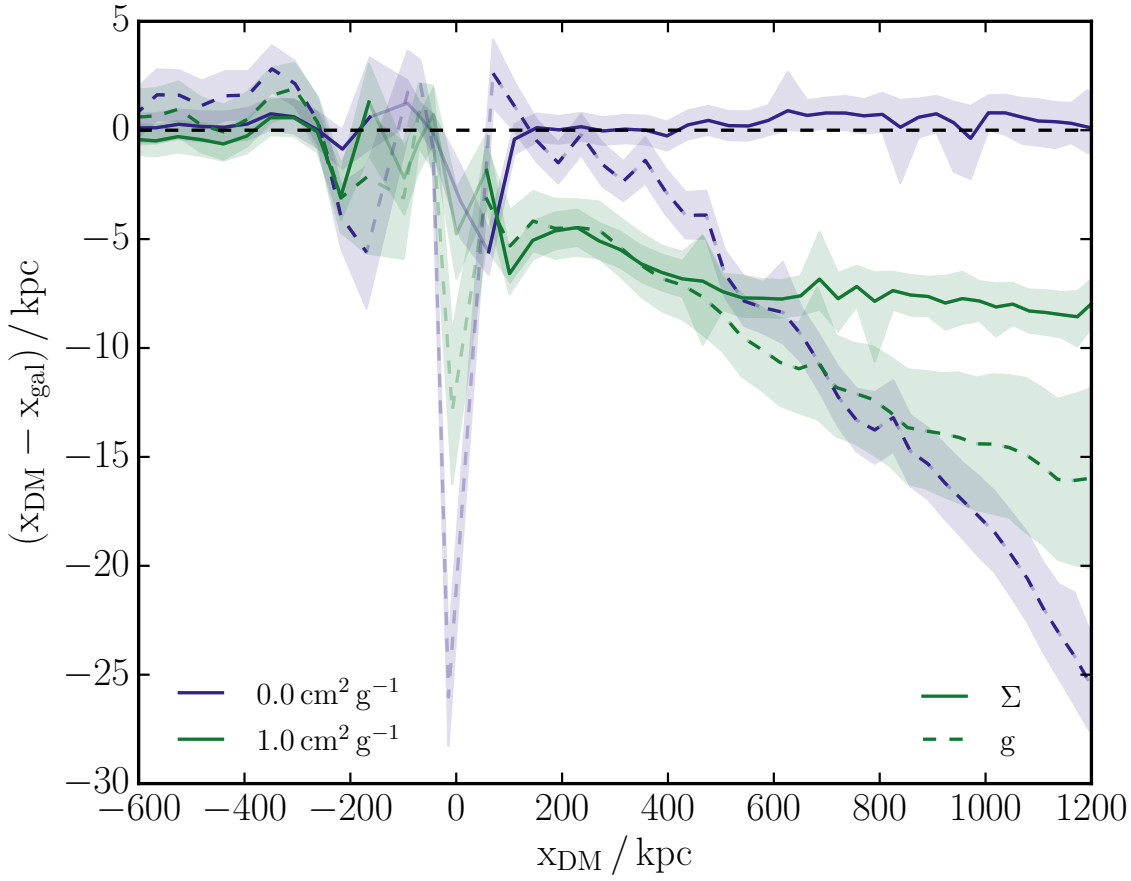


Figure 5.11: The offset between the stars and DM from simulations including adiabatic gas. As in Fig. 5.8, the g measurement is the offset between the DM position measured using reduced shear and the stellar position measured by fitting to the projected density of stars. As such, the large offset with 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.

Explaining offsets with collisionless DM

In the top row of Fig. 5.12 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 of mass measured on larger scales. This is quantified in the top panel of Fig. 5.13 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. 5.11 show the best-fit 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. 5.13 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 C. We find that this quantity peaks at $R \sim 60\,\mathrm{kpc}$ for reduced shear and $R \sim 230\,\mathrm{kpc}$ 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.

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. In Fig. 5.14 we show the press release image of the Bullet Cluster, with a similarly rendered output from our fiducial CDM + gas simulation shown in Fig. 5.15. The gas morphology of the bullet, as well as the offset between the X-ray bright bullet and its associated DM peak, are in good agreement with the observed Bullet Cluster.

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 $\sigma/m = 2\,\mathrm{cm}^2\,\mathrm{g}^{-1}$, 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 $70\,\mathrm{kpc}$ with collisionless DM to $110\,\mathrm{kpc}$ with $\sigma/m = 2\,\mathrm{cm}^2\,\mathrm{g}^{-1}$. While this latter value is in better agreement with the observed distance between the shock front and contact discontinuity ($\sim 140\,\mathrm{kpc}$), 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.

5.5 Conclusions

We have used our SIDM code, described in Chapter 3, to perform idealised 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.

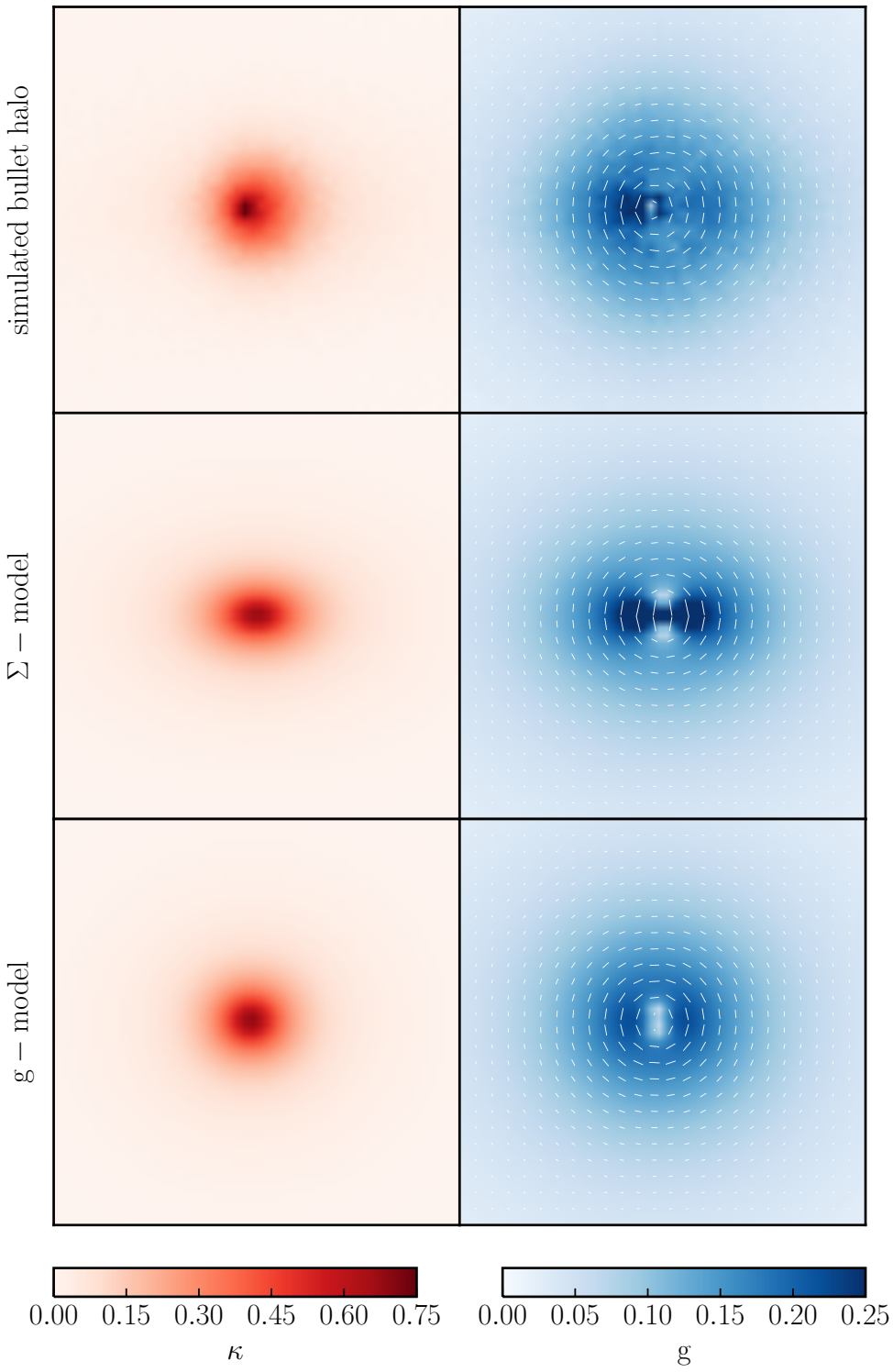


Figure 5.12: The convergence (left column) and reduced shear (right column) due to the bullet halo DM, for a simulation with collisionless DM and non-radiative gas. The top row shows the simulation output (only including DM particles that are part of the bullet halo in the initial conditions), while the middle and bottom rows show the best fit maps generated by fitting to the projected surface density and reduced gravitational shear respectively. Each panel is 1 Mpc across.

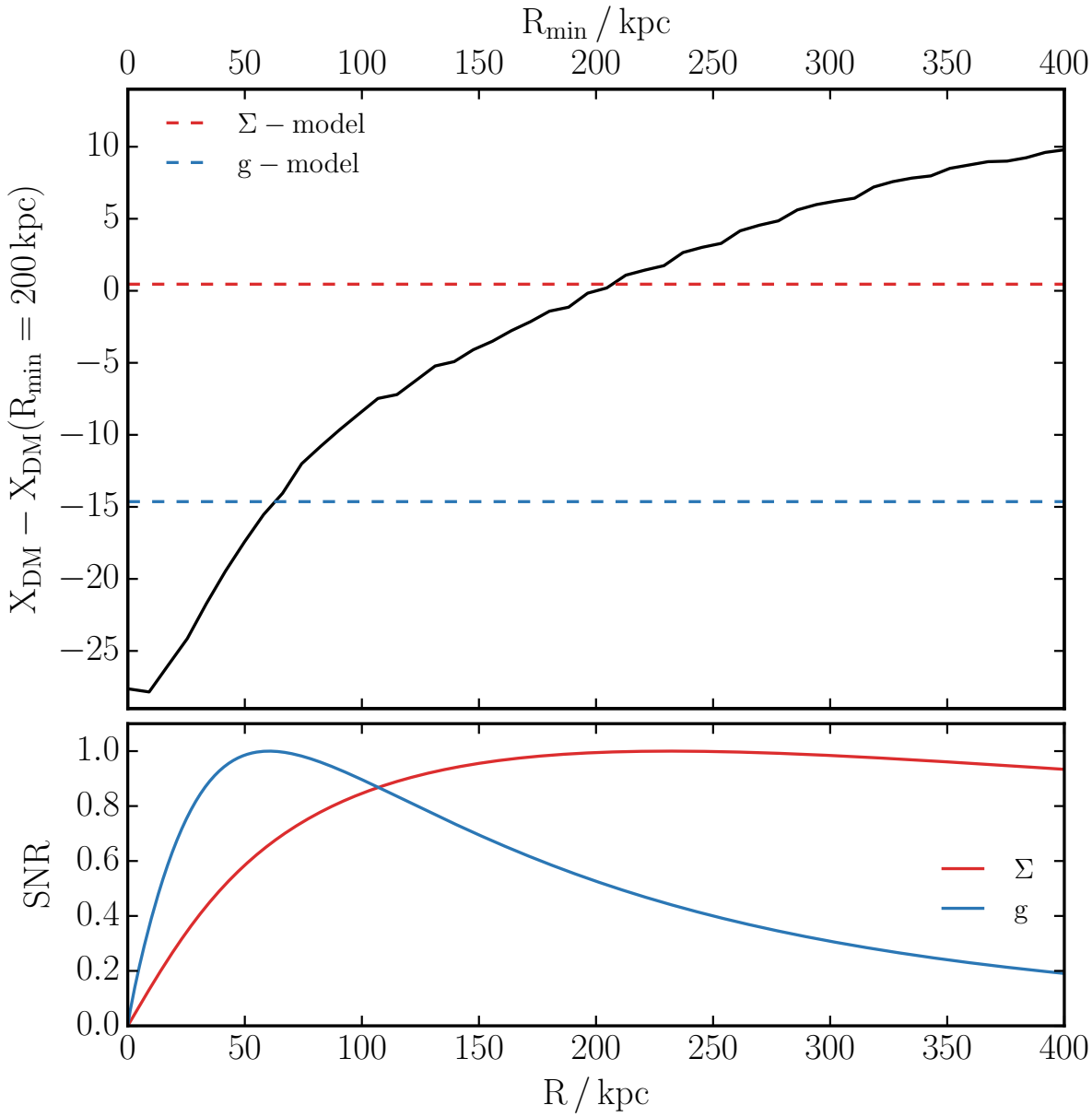


Figure 5.13: Top panel: the x-position of the bullet DM halo measured using shrinking circles on only the DM particles that are part of the bullet halo in the initial conditions. The position is shown as a function of the radius to which the circle is shrunk, with the DM halo shifting to the left as the measurement is made on smaller scales. The best-fit 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-fit model to the projected density. Further details are in Appendix C.

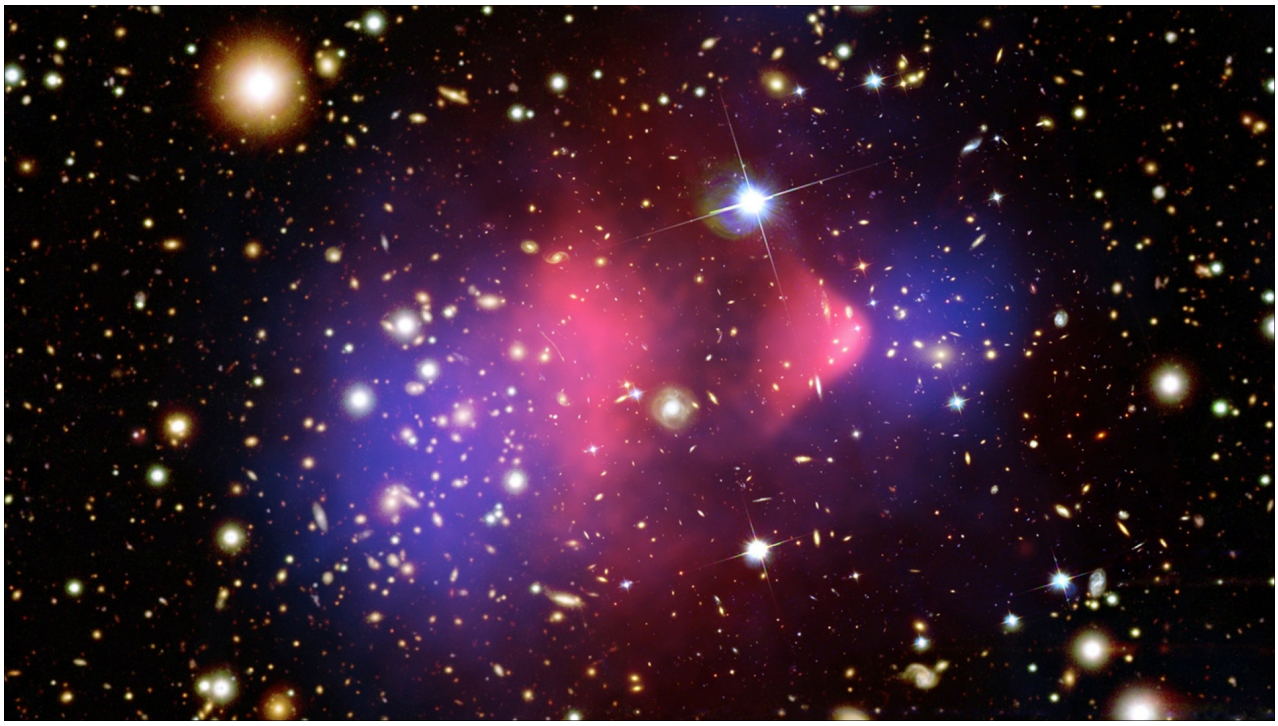


Figure 5.14: Press release image of the Bullet Cluster. Credit: X-ray: NASA/CXC/CfA/Markevitch (2006); Optical: NASA/STScI; Magellan/U.Arizona/Clowe et al. (2006); Lensing Map: NASA/STScI; ESO WFI; Magellan/U.Arizona/Clowe et al. (2006).

Figure 5.15: The projected DM density (blue) and X-ray luminosity (pink) from a simulation run from our fiducial initial conditions, with CDM and gas. The optical image from the press release has been overlaid. A video of the simulation evolving (including animated galaxies) is available at https://www.youtube.com/watch?v=rLx_TXhTXbs

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 $\sigma/m < 1.25\,\mathrm{cm^2\,g^{-1}}$ 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 $\sigma/m = 2\,\mathrm{cm^2\,g^{-1}}$, the DM–galaxy offset at the time of the observed Bullet Cluster is $\sim 20\,\mathrm{kpc}$, which is allowed by the $25 \pm 29\,\mathrm{kpc}$ 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.

Chapter 6

Simulations of self-interacting dark matter with anisotropic scattering

6.1 Introduction

The tightest constraints on DM's self-interaction cross-section have come from galaxy cluster scales (Miralda-Escudé, 2002; Randall et al., 2008; Rocha et al., 2013; Peter et al., 2013; Harvey et al., 2015; Kim et al., 2017), while the astrophysical motivation for SIDM predominantly comes from dwarf galaxies (Vogelsberger et al., 2012; Zavala et al., 2013; Elbert et al., 2015; Vogelsberger et al., 2016). The typical velocities of DM particles within galaxy clusters are of the order of $1000\,\mathrm{km\,s^{-1}}$, while in dwarf galaxies they can be 10 to 100 times lower. Given that it is common for scattering cross-sections to have a strong velocity dependence (such as $\sigma \propto v^{-4}$ in the case of Rutherford scattering) and that this is true also of many particle physics based models for SIDM (Ackerman et al., 2009; Feng et al., 2010; Buckley & Fox, 2010), it is not unreasonable for the cross-section in dwarf galaxies to be orders of magnitude larger than in galaxy clusters. This has led to such particle candidates being simulated (Vogelsberger et al., 2012; Zavala et al., 2013; Vogelsberger & Zavala, 2013; Vogelsberger et al., 2014b) in a bid to alleviate tensions on small scales, while evading constraints that come from larger scales.

The parameters governing such velocity-dependent cross-sections (DM mass, mediator mass and coupling strength) can in principle be constrained by estimating the cross-section

for DM–DM scattering at different velocities. Kaplinghat et al. (2016) recently estimated the DM mass and dark photon mass, assuming that the inferred core sizes in observed galaxy clusters (Newman et al., 2013), low surface brightness galaxies (Kuzio de Naray et al., 2008) and dwarf galaxies (Oh et al., 2011) are due to SIDM, and using a coupling strength equal to the electromagnetic fine structure constant $\alpha' = \alpha \approx 1/137$.

What is often ignored when simulating these velocity-dependent cross-sections is that the scattering is usually anisotropic. This is because the velocity dependence results from a term in the scattering cross-section that depends on the exchanged momentum, which depends on both the collision velocity and the scattering angle. This angular dependence has not been included in previous simulations, which have instead simulated the scattering as isotropic but with a cross-section modified such that the effects of DM scattering should be similar to what would result from a faithful simulation using the underlying particle interaction. This has been done by matching the momentum transfer cross-section as a function of collision velocity, $\sigma_T(v)$, between the true particle interaction and that used in the simulations. While this may work well when the DM velocity distribution is close to isotropic, Kahlhoefer et al. (2014, hereafter K14) found that for the case of colliding galaxy clusters the momentum transfer cross-section is insufficient to fully characterise the effects of DM scattering. This is not surprising; when galaxy clusters collide there is a strongly preferred direction along which DM particles collide and the angular distribution of scattered DM becomes important.

The goal of this chapter is to explore the effects of anisotropic DM scattering, by simulating scattering processes faithfully to their underlying particle physics models. We explore how the results from anisotropic scattering compare with the case of isotropic scattering for the evolution of an isolated DM halo, as well as in a galaxy cluster collision. By also simulating these systems using an appropriately matched isotropic cross-section, as has been done in the past, we can test the validity of this approximate scheme.

This chapter is organised as follows. In §6.2 we discuss the physics of anisotropic scattering and introduce two examples of anisotropic scattering cross-sections. We then investigate how core-formation in an isolated DM halo depends on the angular-dependence of the scattering cross-section in §6.3. In §6.5 we show the results of simulations of the Bullet

Cluster run with anisotropic scattering, and contrast this system that has a clear directionality to the case of an isolated halo. Finally, we summarise the chapter in §6.6.

6.2 Angular Dependent Scattering

The key particle physics quantity for a study of the astrophysical effects of SIDM is the differential cross-section, $\mathrm{d}\sigma/\mathrm{d}\Omega$. This quantifies the rate at which particles are scattered into different patches of solid angle, and can vary as a function of the collision velocity, v.

In this section we briefly describe the particle physics that leads to angular-dependent scattering, and then describe different integrated cross-sections, which condense a scattering probability that varies with angle into a single number. We end the section by introducing two different anisotropic cross-sections, which we later use in our simulations.

When dealing with particle scattering we assume that DM particles are indistinguishable and work in the centre of momentum frame of the two interacting particles, with the velocities of the DM particles defined to be $\pm v/2$ in the z-direction. To conserve both energy and momentum, both particles leave the collision with a velocity of $v/2$ at a polar angle θ with respect to their incoming direction, where θ must be the same for the two particles. Assuming that the scattering potential is spherically symmetric, the differential cross-section is independent of the azimuthal angle ϕ. The DM particle mass is m_χ and the DM scattering is mediated by a particle with mass m_ϕ, with a coupling strength (analogous to the fine structure constant) α_χ.

6.2.1 Particle physics of angular dependent scattering

As mentioned in this chapter's introduction, most efforts to simulate SIDM have treated the DM scattering as isotropic. This isotropic scattering, also commonly referred to as 'contact interactions', 'hard sphere scattering' and 'billiard ball scattering'[1] results from an interaction with a massive mediator, leading to a short range force. For DM particles with a

[1]Readers may be interested to note that the cross-section per unit mass of a World Pool-Billiard Association ball is $0.64\,\mathrm{cm}^2\,\mathrm{g}^{-1}$, of similar magnitude to commonly studied SIDM cross-sections, though it is unclear how resilient pool balls would be to $\sim 1000\,\mathrm{km\,s}^{-1}$ collisions.

relative velocity v, scattering will be isotropic when the mediator mass is much heavier than the DM particle momenta, $c\,m_\phi \gg v\,m_\chi$. When this is not the case, the interaction cross-section will typically depend on the exchanged momentum, which increases with an increased collision velocity or an increased scattering angle, leading to velocity-dependent anisotropic scattering. This second case, with long range interactions due to a light or massless mediator, arises in models of mirror DM (Blinnikov & Khlopov, 1983; Berezhiani et al., 1996; Foot, 2004) and atomic DM (Cline et al., 2012; Cyr-Racine & Sigurdson, 2013), as well as some other hidden sector DM models (Feng et al., 2009; Foot & Vagnozzi, 2015; Boddy et al., 2016).

6.2.2 Integrated cross-sections

Given an azimuthally-symmetric differential cross-section, $\mathrm{d}\sigma/\mathrm{d}\Omega$, the total cross-section is

$$\sigma \equiv \int \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\mathrm{d}\Omega = 2\pi \int \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\sin\theta\,\mathrm{d}\theta. \tag{6.1}$$

While this is the relevant quantity when considering the rate at which particles interact, it does not fully describe the consequences of these interactions, as the effect of scattering by a large angle (and so transferring a large amount of momentum between the two particles) is greater than the effect of scattering by a small angle.

A useful concept when comparing the macroscopic consequences of particle interactions with different angular dependencies for the differential cross-section is the momentum-transfer cross-section. For a scattering angle of θ, the momentum transfer along the direction of the collision is

$$\Delta p_z = p(1-\cos\theta), \tag{6.2}$$

where p is the magnitude of each of the incoming particles' momenta in the centre of momentum frame. We therefore define the momentum-transfer cross section as

$$\sigma_T \equiv \int (1-\cos\theta)\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\mathrm{d}\Omega. \tag{6.3}$$

This is similar to the definition of σ, except that interactions that lead to a large amount of momentum transfer contribute more, while those that transfer little momentum are down-weighted. For the case of isotropic scattering, where $\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{\sigma}{4\pi}$ is independent of angle, the momentum-transfer cross-section and the cross-section are equal, i.e. $\sigma_T = \sigma$.

K14 point out that this definition of σ_T overestimates the momentum transfer due to scattering with $\theta > \pi/2$, as in these cases the particles, which we assume to be identical, could be relabelled in such a way that they had scattered with $\theta < \pi/2$. If we weight scatters by the amount of momentum transfer, but relabel particles if they scatter by an angle greater than $\pi/2$, then we get the integrated cross-section

$$\sigma_{\tilde{T}} \equiv \int_{\theta=0}^{\pi/2} (1 - \cos\theta) \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} \mathrm{d}\Omega + \int_{\theta=\pi/2}^{\pi} (1 + \cos\theta) \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} \mathrm{d}\Omega, \tag{6.4}$$

which we call the modified momentum-transfer cross-section. For isotropic scattering $\sigma_{\tilde{T}} = \sigma/2$, while for cross-sections with a negligible amount of large-angle scattering $\sigma_{\tilde{T}} \approx \sigma_T$.[2]

While integrated cross-sections such as σ_T and $\sigma_{\tilde{T}}$ do not fully describe a scattering process, they are useful as a way to compare different scattering cross-sections, and have been used in cosmological simulations of SIDM. The reason for this is computational efficiency. For anisotropic cross-sections, where the vast majority of scattering events involve a low momentum transfer (such as scattering from a Coulomb potential), there can be a large number of interactions, each having very little effect. A less computationally intensive way to simulate a similar effect is to simulate the scattering as isotropic, where most scattering events involve a large amount of momentum transfer, but with a total cross-section scaled down so that the rate of momentum transfer matches the momentum transfer expected from the underlying particle physics model for the DM. This reduces the number of interactions that need to be calculated, while attempting to adequately capture the effects of particle scattering.

6.2.3 A velocity-independent, anisotropic cross-section

In §6.2.1 we discussed that anisotropic scattering usually occurs when the cross-section is also velocity-dependent. However, studying an anisotropic cross-section without velocity dependence is useful to gain intuition for what might happen with more complicated cross-sections, and if realised in nature could have interesting effects in merging galaxy clusters

[2]For most cross-sections $\sigma_{\tilde{T}}$ has a similar value to the viscosity (or conductivity) cross section $\sigma_V \equiv \int \sin^2\theta \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} \mathrm{d}\Omega$ advocated by Tulin et al. (2013a), Cline et al. (2014a) and Boddy et al. (2016) for reasons similar to those for introducing $\sigma_{\tilde{T}}$.

(K14). As an example of such a cross-section we use:

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{\alpha^2}{2m_\chi^2}\frac{1+\cos^2\theta}{1-\cos^2\theta} \tag{6.5}$$

for which both σ and σ_T diverge and which we call Kahlhoefer velocity-independent (KVI). In the case of the σ_T divergence, this is because of the divergence in the differential cross-section as $\theta \to \pi$. While scattering by $\sim \pi$ leads to a significant amount of momentum transfer between the two particles, it leaves the system relatively unchanged, as two identical particles just swap velocities with each other. For this reason, $\sigma_{\tilde{T}}$ is a more sensible choice to describe the scattering. For this differential cross-section

$$\sigma_{\tilde{T}} = \frac{\pi\alpha^2}{m_\chi^2}\left(\ln 16 - 1\right). \tag{6.6}$$

For the KVI cross-section, the divergence in the differential cross-section as $\theta \to 0$ and $\theta \to \pi$, means that we cannot simulate the cross-section completely faithfully. However, the modified momentum-transfer cross-section is finite for this differential cross-section because the divergence in cross-section at low angles is accompanied by a suitably rapid decline in the effectiveness of these scatters to transfer momentum. This means that for a small $\theta_{\min}$, one should expect that ignoring scattering with $\theta < \theta_{\min}$ and $\theta > \pi - \theta_{\min}$, should lead to negligible changes to the effect of this cross-section.

Introducing a cut-off, such that the differential cross-section follows equation (6.5) for $\theta_{\min} < \theta < \pi - \theta_{\min}$ and is 0 outside of this, the cross-section is then finite,

$$\sigma(\theta_{\min}) = \frac{2\pi\alpha^2}{m_\chi^2}\left\{\ln\left(\frac{1+\cos\theta_{\min}}{1-\cos\theta_{\min}}\right) - \cos\theta_{\min}\right\}. \tag{6.7}$$

The modified momentum-transfer cross-section becomes

$$\sigma_{\tilde{T}}(\theta_{\min}) = \frac{\pi\alpha^2}{m_\chi^2}\left\{\cos^2\theta_{\min} - 2\cos\theta_{\min} + 4\ln(1+\cos\theta_{\min})\right\} \tag{6.8}$$

which can be compared with equation (6.6) to see how much momentum-transfer we expect to miss by introducing $\theta_{\min}$. In particular, for $\theta_{\min} \ll 1$

$$\sigma_{\tilde{T}} - \sigma_{\tilde{T}}(\theta_{\min}) = \frac{\pi\alpha^2}{m_\chi^2}\left\{\theta_{\min}^2 + \mathcal{O}(\theta_{\min}^4)\right\}. \tag{6.9}$$

As an example, with $\theta_{\min} = 0.1$ we only lose 0.6% of $\sigma_{\tilde{T}}$.

6.2.4 Yukawa-potential SIDM

A general result for scattering mediated by a massive mediator particle is that it is equivalent to having a Yukawa potential. Loeb & Weiner (2011) noted that such a cross-section would display interesting astrophysical signatures because the rate of scattering peaks at a particular pairwise velocity, falling at smaller or larger velocities. This could lead to scattering being important in DM haloes of a particular mass (and so a particular velocity dispersion), while being negligible in the more massive haloes that have thus far provided the tightest constraints on the DM cross-section. Simulations including such a model for DM scattering have been performed, but have simulated the scattering as isotropic, using the momentum-transfer cross-section of the underlying particle physics model, as the cross-section for isotropic scattering (Vogelsberger et al., 2012; Vogelsberger & Zavala, 2013; Zavala et al., 2013; Vogelsberger et al., 2014b). We call this method of simulating DM models with anisotropic cross-sections, σ_T-match.

There is no analytical form for the differential scattering cross-section due to a Yukawa potential, but by using the Born-approximation (Jelley, 1990), valid when the scattering potential can be treated as a small perturbation, we can find an analytical form that approximates the true differential cross-section. For an interaction potential given by

$$V(r) = -\frac{\alpha_\chi e^{-m_\phi r}}{m_\phi r}, \tag{6.10}$$

the differential cross-section assuming the Born approximation is (Ibe & Yu, 2010)

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{\alpha_\chi^2}{m_\chi^2 \left(m_\phi^2/m_\chi^2 + v^2 \sin^2 \frac{\theta}{2}\right)^2}, \tag{6.11}$$

where we have used natural units with $\hbar = c = 1$. This can be re-written as

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{\sigma_0}{4\pi(1 + \frac{v^2}{w^2} \sin^2 \frac{\theta}{2})^2}, \tag{6.12}$$

where $w = m_\phi c/m_\chi$ is a characteristic velocity, below which the scattering is roughly isotropic with $\sigma \approx \sigma_0$. At higher velocities, the scattering has an angular dependence that tends to that from scattering with a Coulomb potential, with a cross-section that decreases with increasing velocity.

From the differential cross-section we can calculate the integrated cross-sections

$$\sigma = \frac{\sigma_0}{1 + \frac{v^2}{w^2}}, \tag{6.13}$$

$$\sigma_T = \sigma_0 \frac{2w^4}{v^4} \left\{ \ln\left(1 + \frac{v^2}{w^2}\right) - \frac{v^2}{w^2 + v^2} \right\}, \tag{6.14}$$

and

$$\sigma_{\tilde{T}} = \sigma_0 \frac{2w^4}{v^4} \left\{ 2\ln\left(1 + \frac{v^2}{2w^2}\right) - \ln\left(1 + \frac{v^2}{w^2}\right) \right\}. \tag{6.15}$$

We note that at low velocities the scattering is non-perturbative and the Born approximation is no longer valid. This is the reason why the behaviour of σ_T in equation (6.14) differs from that found by numerically solving for orbits in a classical Yukawa potential (Khrapak et al., 2004, and Appendix B where we sketch out their calculations) at low velocities. The Born approximation result tends towards isotropic scattering with a velocity-independent cross-section, while the results of the full calculation have a cross-section that logarithmically increases towards low velocities. We therefore expect slightly different results compared with a full calculation of scattering through a Yukawa potential. However, using the Born approximation is useful as it gives us an analytical differential cross-section that we can faithfully simulate, allowing us to test the procedure of using isotropic scattering to capture the effects of a more complicated scattering process. With a known differential cross-section, we can simulate the scattering in a fully consistent manner, and compare that with simulating it with a suitably matched $\sigma(v)$ and isotropic scattering.

6.3 Implementing DM scattering

Our implementation of DM scattering builds upon the method introduced in Chapter 3. To briefly summarise: at each time-step particles search for neighbouring particles, separated by a distance less than the search radius h, and scatter from each of their neighbouring particles with a probability

$$P_{\rm scat} = \frac{(\sigma/m) m_{\rm p}\, v\, \Delta t}{\frac{4\pi}{3} h^3}, \tag{6.16}$$

where σ/m is the DM scattering cross-section divided by the DM particle mass, $m_{\rm p}$ is the mass of a simulation particle, v is the relative velocity of the two particles, and Δt is the

size of the time-step. h is a numerical parameter that we keep fixed for all particles, with a size similar to the gravitational softening length ϵ.

All simulations discussed in §6.4 and §6.5.1 used $h = 2\epsilon = 5.6\,\mathrm{kpc}$, while in §6.5.2 larger h (up to 4ϵ) were used. In general, a smaller h is better, because scattering is then more local and the scattering rate resolves small-scale density peaks. However, with smaller h, fewer neighbour particles are found, so the probability of scattering from those particles must increase to achieve the correct rate of scattering. If h is too small, such that $P_{\mathrm{scat}} > 1$, then the rate of scattering will no longer be correctly calculated. To avoid this, small time-steps can be used, but this makes the simulations computationally expensive. For this reason, we use larger h for our simulations that have large cross-sections, keeping the maximum value of P_{scat} below 0.1 in all cases. A more detailed discussion about the choice of h can be found in §4.2.3.

6.3.1 Implementation of velocity-dependent scattering

For DM interactions that lead to a velocity-dependent cross-section, equation (6.16) can be modified by letting σ vary with velocity. For simple models of DM, $\sigma(v)$ may have an analytical representation, while for more complicated models a lookup table can be used.

6.3.2 Implementation of angular-dependent scattering

From considerations of the solid angle at different polar angles, the probability density function for scattering by an angle θ is

$$p(\theta) = \frac{2\pi \sin\theta}{\sigma} \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}. \tag{6.17}$$

Integrating this, we get the cumulative distribution function,

$$P(\theta) = \int_0^\theta p(\theta')\,\mathrm{d}\theta', \tag{6.18}$$

which is the probability that a particle scatters by an angle less than θ.

For particles due to scatter, a polar scattering angle can be drawn from $p(\theta)$ as the θ that satisfies

$$P(\theta) = X \tag{6.19}$$

where X is a random variable with a uniform distribution in the interval $[0, 1]$.

For a general differential cross-section, the inverse of $P(\theta)$ is not necessarily analytical. To allow us to simulate cross-sections with general angular dependence, we numerically find solutions to equation (6.19) at N_θ values of X distributed uniformly in the interval $[0, 1]$. The X_i take the values $X_i = \frac{i-1/2}{N_\theta}$ where $i = \{1, 2, 3, ..., N_\theta\}$, and we label the angles uniformly drawn from $p(\theta)$, θ_i (i.e. $P(\theta_i) = X_i$). For two particles that scatter, finding a polar scattering angle is then just a case of drawing an integer i from the interval $[1, N_\theta]$ and setting $\theta = \theta_i$.

6.3.3 Implementation of velocity-dependant angular-dependence

In general the angular and velocity dependence of a scattering cross-section need not be separable, and $P(\theta)$ can vary with velocity. For these cases, the discussion in §6.3.2 can be easily extended by generating a set of θ_i for each of N_v velocities, where N_v must be large enough that $p(\theta)$ does not vary substantially from v_i to v_{i+1}. Using this as well as a velocity-dependent $\sigma(v)$ in equation (6.16) allows us to simulate particle scattering with a general differential cross-section.

Throughout the rest of this work, we use $N_\theta, N_\mathrm{v} = 1000$ with the v_i logarithmically spaced from 0.01 to $10\,000\,\mathrm{km\,s^{-1}}$. At velocities below $0.01\,\mathrm{km\,s^{-1}}$ the cross-section and angular-dependence were set as if $v = 0.01\,\mathrm{km\,s^{-1}}$, while at velocities greater than $10\,000\,\mathrm{km\,s^{-1}}$ the cross-section was set to zero. This was to reflect the fact that for Yukawa-like models the cross-section does not vary at low velocities, and falls off rapidly at high velocities. The values of N_θ and N_v could be increased if required, as could the range of velocities covered, but these values were found to be sufficient for the cross-sections and systems we simulated here.

6.3.4 Testing generalised scattering

To test our implementation of SIDM with velocity-dependent anisotropic cross-sections we repeated our test case from §4.1, a cube of particles moving through a uniform slab of stationary particles. Particles in the cube all moved with a common velocity v_cube through

the slab and there were no gravitational forces. We used the differential cross-section for Yukawa scattering described by equation (6.12) and ran the test at five different $v_{\rm cube}$, ranging from $0.1w$ to $10w$. The cross-section normalisation σ_0 and the projected density of slab particles were chosen such that 10% of the $N_{\rm cube} = 10^6$ cube particles would be scattered if the scattering was in the isotropic regime ($v \ll w$) and particles were not allowed to scatter more than once.

In Fig. 6.1 we show the results of these test cases, plotted as the number of particles that scatter, $N_{\rm s}$, per unit polar angle. These agree with the predictions that were made using $\frac{{\rm d}\sigma}{{\rm d}\Omega}(v_{\rm cube})$ and the projected density of DM through the slab. To make these predictions, the number of expected scatters was calculated using $\sigma(v_{\rm cube})$. Their angular distribution was then calculated by transforming the relevant $p(\theta)$ into the frame of the slab, from the centre of momentum frame of the collisions where it is defined. As well as the predicted distribution at velocity $v_{\rm cube}$, we also plot the predicted distribution at $(v_{i+5}/v_i)v_{\rm cube}$, where the v_i were determined using $N_{\rm v} = 1000$ and velocities in the range 0.01 to $10\,000\,{\rm km\,s^{-1}}$. Changing velocity by only one bin led to imperceptibly small changes in the scattered distribution, and the small change in the predicted distribution when increasing velocity by five velocity-bins justifies our choice of $N_{\rm v}$.

6.4 Core growth in isolated haloes

In an isolated halo with an isotropic velocity distribution, there is no preferred direction for particle scattering, and a suitably matched isotropic cross-section may be able to mimic the effects of an anisotropic one. To test the efficacy of the σ_T-match procedure defined in §6.2.4, we investigate the rate at which cores form in an isolated Hernquist profile (Hernquist, 1990) with anisotropic scattering and compare to isotropic scattering. We simulate a halo with parameters corresponding to the bullet halo in the Bullet Cluster, namely $M = 2.46 \times 10^{14}\,{\rm M}_\odot$ and $a = 279\,{\rm kpc}$ (Chapter 5).

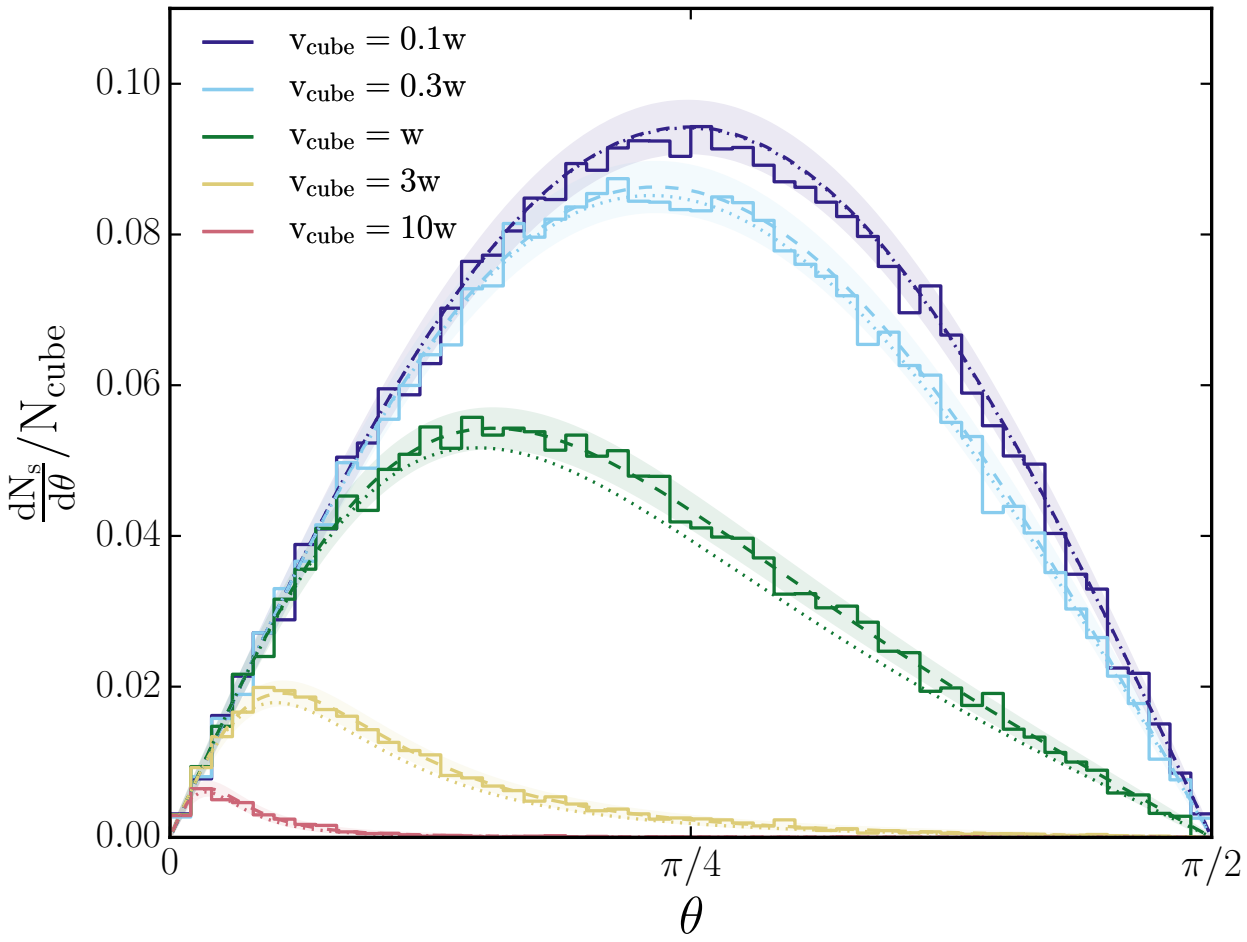


Figure 6.1: The number of particles that scatter by different polar angles for a cube of DM particles moving through a uniform slab of particles at a speed $v_{\rm cube}$. The scattering cross-section was Yukawa scattering under the Born approximation - see equation (6.12). Different line colours correspond to different $v_{\rm cube}$, which changes the normalisation and angular dependence of the Yukawa cross-section. The solid lines show scatters into different bins of angle measured in our test simulations, while the dashed lines and shaded regions show the analytically predicted distribution of scattering angles for each of these simulations and the expected 2σ Poisson variation. The dotted lines show the prediction for a cube velocity of $(v_{i+5}/v_i)v_{\rm cube}$ and are described further in §6.3.4. With increasing $v_{\rm cube}$ the number of scatters drops, and the scattering becomes more anisotropic.

6.4.1 Determining core sizes

Following §4.2.4, we find the core size, $r_{\rm core}$, by fitting a cored Hernquist profile

$$\rho(r) = \frac{M}{2\pi} \frac{a}{(r^\beta + r_{\rm core}^\beta)^{1/\beta}} \frac{1}{(r+a)^3} \tag{6.20}$$

to the radial density distribution. During the fitting procedure we allow M, a, and $r_{\rm core}$ to vary, while holding $\beta = 4$ fixed. We measure the core evolution in terms of a dimensionless time, $T/T_{\rm dyn}$, where the dynamical time follows the definition in Kochanek & White (2000),

$$T_{\rm dyn} = 4\pi \sqrt{\frac{a^3}{GM}}, \tag{6.21}$$

and is $1.8\,{\rm Gyr}$ for our chosen Hernquist profile. We run simulations with different cross-sections, which again are defined to be dimensionless

$$\hat{\sigma} = \frac{\sigma}{m} \frac{M}{a^2}, \tag{6.22}$$

such that $\hat{\sigma} = 1$ corresponds to $\sigma/m \approx 1.5\,{\rm cm^2\,g^{-1}}$ for our simulated halo.

6.4.2 KVI scattering

With a KVI cross-section we cannot use σ_T-match, as the momentum transfer cross-section diverges. Instead we match $\sigma_{\tilde{T}}$ between KVI scattering and isotropic scattering. In Fig. 6.2 we demonstrate the effectiveness of $\sigma_{\tilde{T}}$-match, showing that when the KVI cross-section is normalised such that it has the same $\sigma_{\tilde{T}}$ as a particular isotropic cross-section, then the evolution of the core size is close to that for the matched isotropic cross-section. We plot the results using $\theta_{\rm min} = 0.025$ and $\theta_{\rm min} = 0.1$, which respectively correspond to 0.04% and 0.6% of $\sigma_{\tilde{T}}$ being truncated. For $\theta_{\rm min} = 0.025(0.1)$ the total cross-section, σ, is 4.4(2.8) times larger than for the $\sigma_{\tilde{T}}$-matched isotropic cross-section.

The similarity between the $r_{\rm core}$ evolution with isotropic scattering and with $\sigma_{\tilde{T}}$-matched anisotropic scattering suggests that at least in locally isotropic systems $\sigma_{\tilde{T}}$ is a useful way to characterise DM scattering. Using only isotropic scattering and a calculation of $\sigma_{\tilde{T}}$ (and not the full, underlying differential cross-section that leads to it) we can predict how a system would evolve with anisotropic scattering.

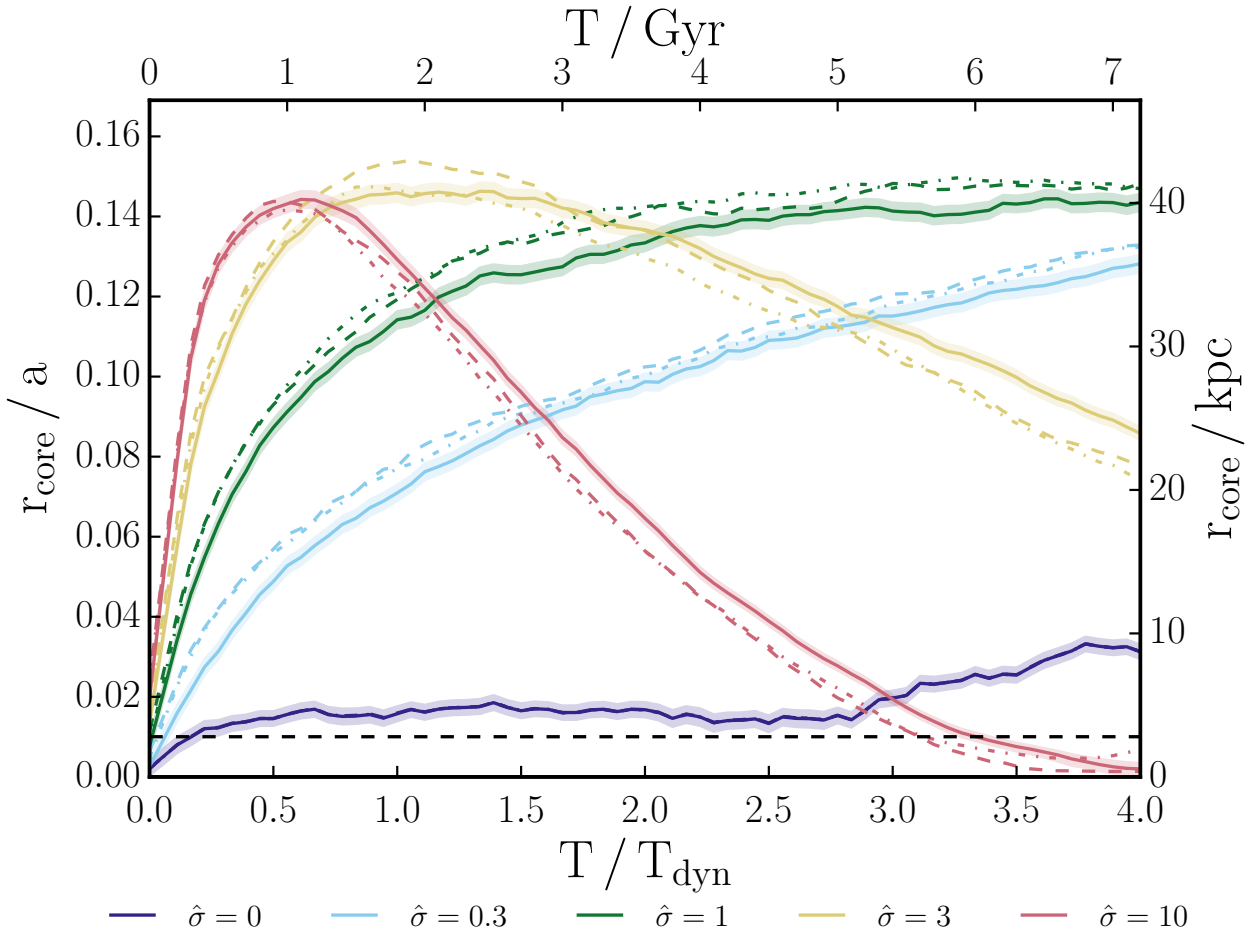


Figure 6.2: The evolution of core size for an isolated Hernquist profile, evolved with different SIDM scattering cross-sections. The solid lines are for isotropic DM scattering, while the dashed and dot-dashed lines are for anisotropic KVI scattering with $\theta_{\rm min} = 0.025$ and 0.1 respectively, matched to the isotropic cross-sections using $\sigma_{\tilde{T}}$. In the case of the dashed line, there is between 4 and 5 times as much scattering as for the equivalent solid line, but due to the angular-dependence of those interactions the resulting evolution is similar. The shaded regions around the solid lines show the 1σ error on $r_{\rm core}$ and the horizontal dashed line shows the size of the gravitational softening length.

6.4.3 Yukawa-potential scattering

In order to further test the ability of an integrated cross-section to capture the effects of anisotropic scattering, we simulate the same Hernquist profile as in §6.4.2, this time with scattering from a Yukawa potential assuming the Born approximation. The DM scattering follows equation (6.12) and we simulate cross-sections with three different w, with a variety of σ_0. The modified momentum-transfer cross-sections for the different simulated cross-sections are shown in the left panel of Fig. 6.3.

In the right panel of Fig. 6.3 we show how the core sizes evolve for the different particle models shown in the left panel. As the Hernquist halo in question has a typical velocity for particles moving within the halo of $v_g = \sqrt{GM/a} \approx 1950\,\mathrm{km\,s^{-1}}$, most scattering for the $w = 3000\,\mathrm{km\,s^{-1}}$ models is in the isotropic regime, and the core evolution is similar to that seen in Fig. 6.2 with isotropic scattering. For all of the particle models, the evolution of the core-size is approximately determined by the value of $\sigma_{\tilde{T}}(v_g)$. For example the $\hat{\sigma}_0 = 3$, $w = 1000\,\mathrm{km\,s^{-1}}$ cross-section and $\hat{\sigma}_0 = 1$, $w = 3000\,\mathrm{km\,s^{-1}}$ cross-section show similar evolution of r_{core}, while having similar values of $\sigma_{\tilde{T}}(v_g)$.

Having simulated these particle models using the full differential cross-section, we can also test both σ_T-match and $\sigma_{\tilde{T}}$-match. We simulated each of the particle models from Fig. 6.3, using isotropic scattering and the appropriate $\sigma(v)$ for σ_T-match and $\sigma_{\tilde{T}}$-match. Our results show that the momentum transfer cross-section, σ_T, is not a good quantity to use to match an isotropic cross-section to an anisotropic one. Instead we found that when cross-sections with different angular dependencies are matched by $\sigma_{\tilde{T}}$, the rate of core formation is very similar.

Fig. 6.4 shows the core size from the σ_T-match and $\sigma_{\tilde{T}}$-match simulations, divided by the core size from the simulation with the full differential cross-section. For clarity we show only the $\hat{\sigma}_0 = 10$ cross-sections, but found that σ_T-match systematically underpredicts the rate of scattering for all of the simulated cross-sections. For the examples shown in Fig. 6.4, this underprediction in scattering rates with σ_T-match manifests itself in core sizes being smaller than in the full differential cross-section simulations with $w = 300\,\mathrm{km\,s^{-1}}$, and larger when $w = 3000\,\mathrm{km\,s^{-1}}$. This change in behaviour is because with $w = 3000\,\mathrm{km\,s^{-1}}$ the halo undergoes core-collapse, and so a lower rate of scattering leads to larger cores at fixed

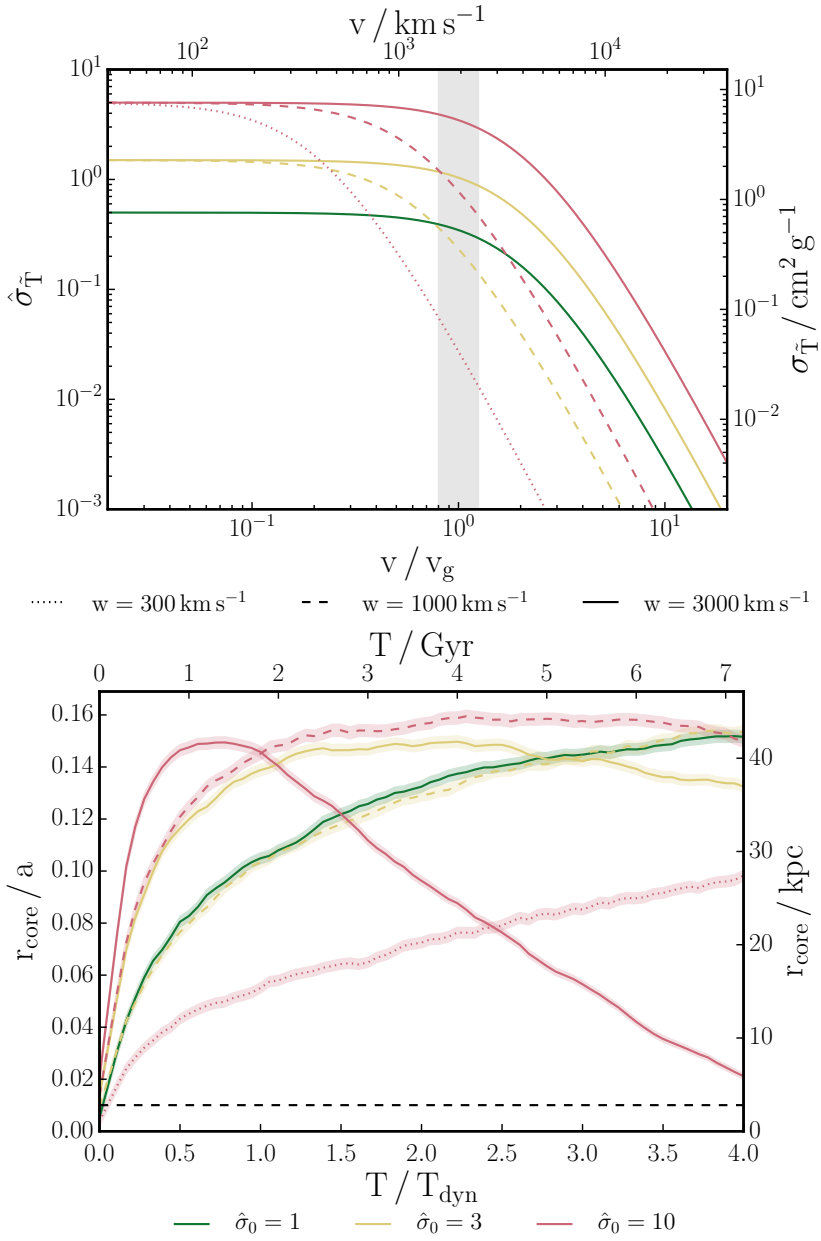


Figure 6.3: Top panel: The modified momentum-transfer cross-section as a function of velocity for particles scattering through a Yukawa potential, assuming the Born approximation. The different lines correspond to different values for the DM and mediator particle masses, as well as the coupling strength for their interaction. These three parameters lead to the two astrophysically important parameters that describe the scattering: σ_0, the cross-section at low velocities when the scattering is isotropic, and w, the velocity around which the cross-section transitions from being isotropic and velocity-independent to anisotropic with a cross-section that drops rapidly with increasing velocity. Bottom panel: the evolution of core size in an isolated Hernquist profile for the different cross-sections shown in the left panel. The evolution of core sizes is approximately captured by the value of $\sigma_{\tilde{T}}$ at $v = v_g$, where v_g is a typical velocity for particles within the halo, marked by the vertical shaded region in the left panel.

time. In contrast, $\sigma_{\tilde{T}}$-match correctly predicts the evolution of the core size for all of our simulated cross-sections. At early times the ratio of core sizes using $\sigma_{\tilde{T}}$-match is not unity, however this is because the small cores lead to large errors on this ratio. This applies also to the $\hat{\sigma}_0 = 10$ model at late times, when $r_{\rm core}$ is again small. We emphasise that σ_T-match has been used in previous work, that has incorrectly estimated the effects of DM models with anisotropic particle scattering (e.g. Vogelsberger et al., 2012; Zavala et al., 2013; Cyr-Racine et al., 2016).

That $\sigma_{\tilde{T}}$-match leads to an enhanced rate of scattering compared with σ_T-match can be understood from considering what happens when scattering is anisotropic, with a large fraction of scattering being by small angles. In this case, σ_T and $\sigma_{\tilde{T}}$ will be similar, as they only differ in how they treat scattering by angles $\theta > \pi/2$. However, for isotropic scattering, $\sigma_{\tilde{T}} = \sigma_T/2$, so matching with $\sigma_{\tilde{T}}$ will lead to twice the rate of scattering as matching with σ_T. Fig. 6.4 demonstrates that $\sigma_{\tilde{T}}$ more accurately captures the effects of scattering by anisotropic cross-sections. This should be expected given that two particle models could have different σ_T, while having indistinguishable particle interactions, purely due to how particles are labelled.

6.5 DM-galaxy offsets in the Bullet Cluster

Having demonstrated that the effects in an isolated halo of SIDM with an anisotropic cross-section can be understood by considering the modified momentum-transfer cross-section, we now go on to investigate whether this is still the case in a system with strong directionality. The system we use is based on the merging galaxy cluster 1E 0657-56 (the Bullet Cluster), using the fiducial mass model from §5.2.3. The initial conditions used for all the simulations contain two Hernquist profiles, separated by $4\,\mathrm{Mpc}$, and with a relative velocity of $2970\,\mathrm{km\,s^{-1}}$ along the line joining the two cluster centres. The main halo and bullet halo have Hernquist density profiles with masses and scale radii $M = 3.85 \times 10^{15}\,\mathrm{M}_\odot$, $a = 1290\,\mathrm{kpc}$, and $M = 2.46 \times 10^{14}\,\mathrm{M}_\odot$, $a = 279\,\mathrm{kpc}$ respectively. The mass within each halo is 99% DM, and 1% 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 profile.

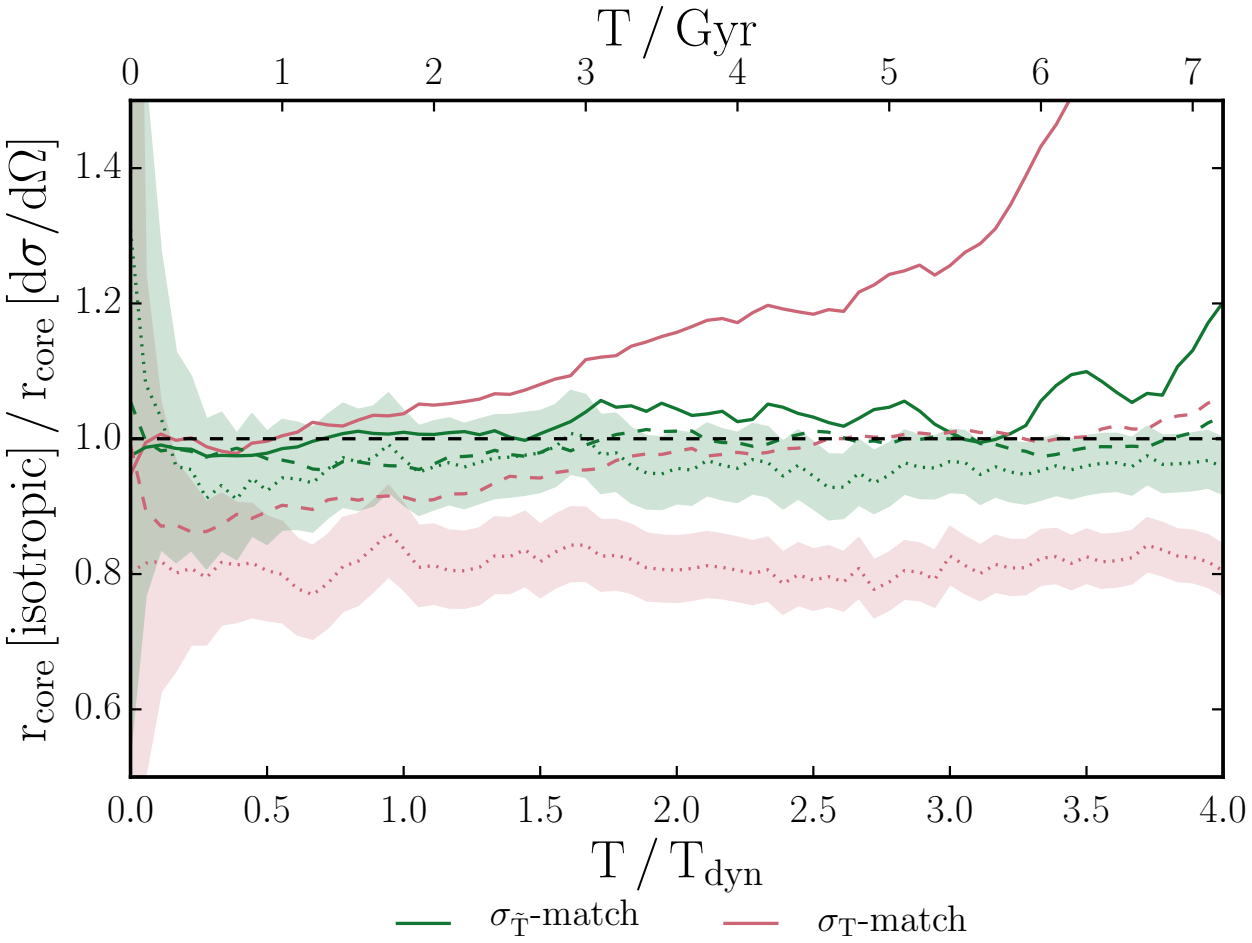


Figure 6.4: The core size using σ_T or $\sigma_{\tilde{T}}$-match, normalised by the core size from faithfully simulating the underlying particle model, for the three cross-sections with $\hat{\sigma} = 10$ from Fig. 6.3. The line styles represent the value of w, in the same manner as Fig. 6.3, and the shaded regions around the $w = 300\,\mathrm{km\,s^{-1}}$ lines are representative of the errors on this ratio – coming from the widths of the posterior distributions for $r_{\rm core}$. $\sigma_{\tilde{T}}$-match works well in all cases, while σ_T-match incorrectly predicts the effects of anisotropic scattering on an isolated halo.

The position estimates for the DM and galaxies were performed following the method described in §5.3.2. This involves simultaneously fitting parametric models for the two haloes to the projected surface density, modelling each halo with a Pseudo Isothermal Elliptical Mass Distribution (PIEMD), which has a 3D density profile

$$\rho(r) = \frac{\rho_0}{(1 + r^2/r_{\rm core}^2)(1 + r^2/r_{\rm cut}^2)}; \quad r_{\rm cut} > r_{\rm core}. \tag{6.23}$$

We choose to fit parametric models because it is often 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) and because of its advantages over local position estimates discussed in §5.4.3.

6.5.1 KVI scattering

In Fig. 6.5 we show that with KVI scattering the measured offsets between DM and galaxies are $\sim 50\%$ larger than for the isotropic cross-section to which they are matched. This matching was done using $\sigma_{\tilde{T}}$-match, which as demonstrated in Fig. 6.2 leads to core formation rates in isolated haloes that are very similar for isotropic scattering and a matched KVI cross-section. This $\sim 50\%$ increase in offsets was seen throughout the evolution of the merger for each of the four cross-sections simulated: 0.5, 1, 1.5 and 2 $\rm cm^2\,g^{-1}$. This result supports the findings of K14, who found that $\sigma_{\tilde{T}}$ was not enough to fully characterise the effects of DM scattering.

Intuitively this can be understood: for isotropic scattering, only a small fraction of bullet halo DM particles scatter with a particle from the main halo, and those that do are typically ejected from the bullet halo. The unscattered DM is coincident with the collisionless galaxies, and any measured offset is a result of fitting to the wake of scattered particles. This is not the case with anisotropic scattering, where many more particles can scatter, but each receives only a small momentum kick. The DM particles that have received such a kick lag behind the collisionless galaxies, leading to an offset between the galaxies and DM.

The KVI cross-sections were simulated using $\theta_{\rm min} = 0.025$. Given that Fig. 6.5 demonstrates that cross-sections with the same $\sigma_{\tilde{T}}$ can lead to different DM-galaxy offsets, one might worry that the results in Fig. 6.5 are dependent on $\theta_{\rm min}$. We tested for convergence with respect to $\theta_{\rm min}$ by running the KVI cross-section $\sigma_{\tilde{T}}$-matched to isotropic $1\,\rm cm^2\,g^{-1}$ with

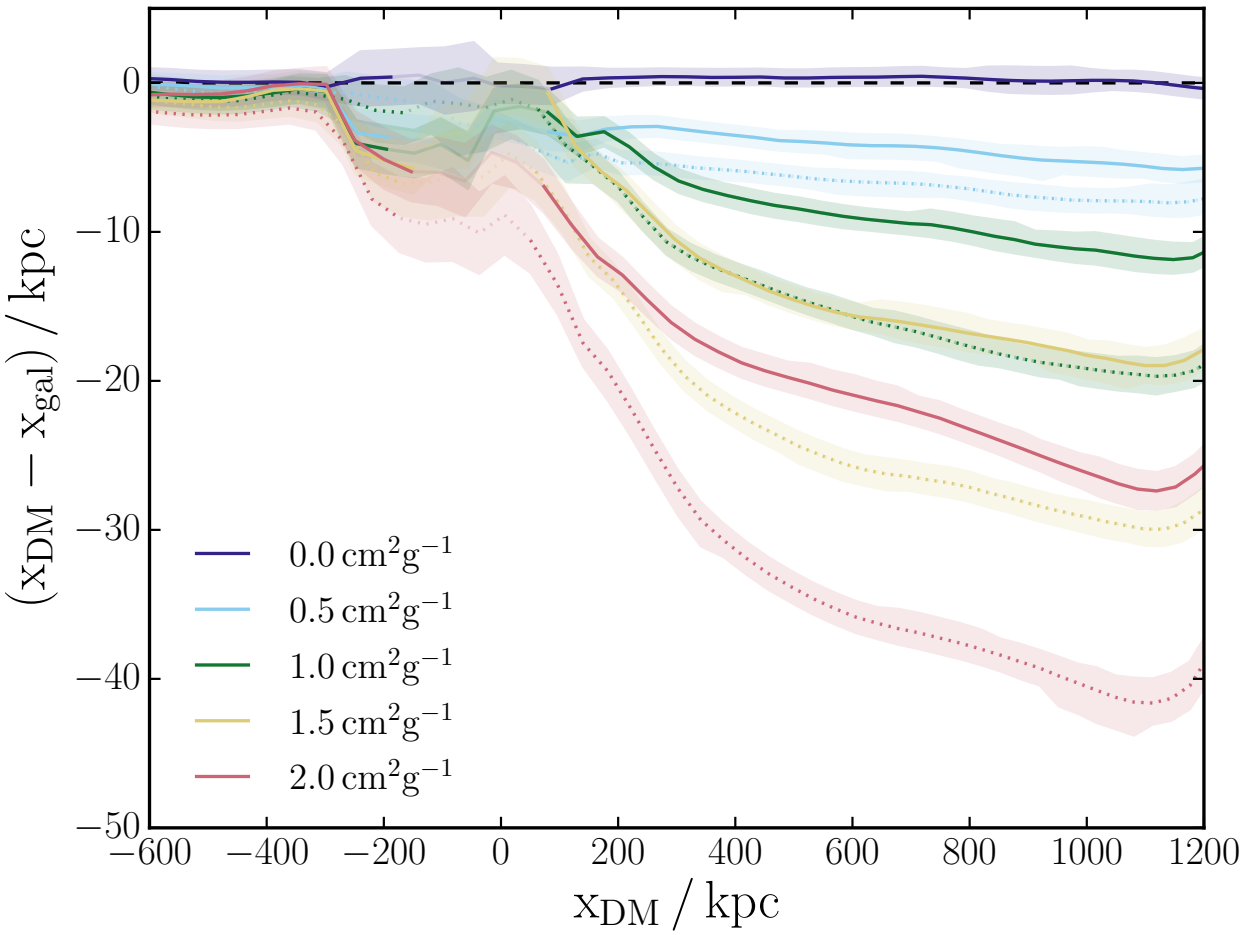


Figure 6.5: The DM-galaxy offset in the bullet halo of the Bullet Cluster as a function of the bullet halo position (in the centre of mass frame). The solid lines are for isotropic scattering with a cross-section as given in the legend. The dotted lines are for scattering with a KVI cross-section that has the same $\sigma_{\tilde{T}}$ as the corresponding isotropic cross-section. The DM and galaxy positions were determined by fitting two parametric model haloes to the respective projected densities. The separations were calculated every $10\,\mathrm{Myr}$, and are plotted using a $50\,\mathrm{Myr}$ moving average. Lines are shown as faint around core passage because measurements of the best-fit halo positions become noisy. In the observed Bullet Cluster the two haloes are separated by $\sim 720\,\mathrm{kpc}$, which happens at $x_{DM} \approx 600\,\mathrm{kpc}$ with our mass model.

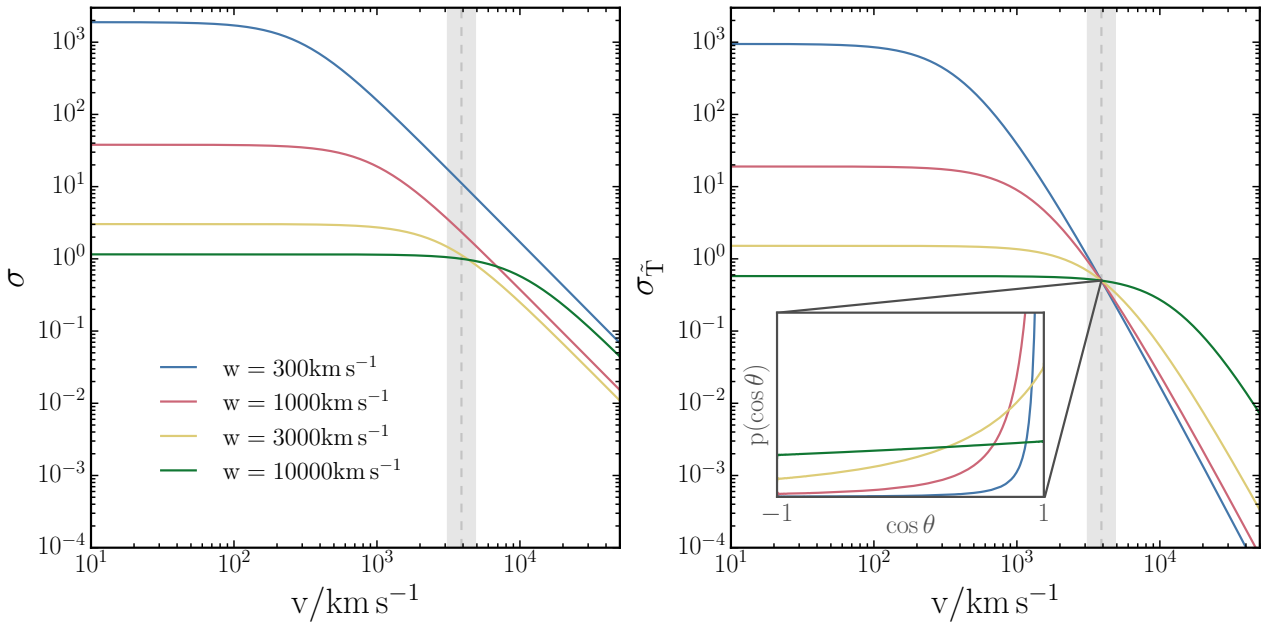


Figure 6.6: The total cross-section (left panel) and modified momentum-transfer cross-section (right panel) for four differential cross-sections. These four cross-sections correspond to the differential cross-section in equation (6.12), with four different values of w. The normalisation of the cross-section, σ_0, was chosen such that the modified momentum-transfer cross-section at a velocity of $3900\,\mathrm{km\,s^{-1}}$ (the relative velocity between the two DM haloes during core passage in our Bullet Cluster simulations) was $0.5\,\mathrm{cm^2\,g^{-1}}$, the same as for isotropic scattering with $\sigma/m = 1\,\mathrm{cm^2\,g^{-1}}$. Inset in the right panel is a plot showing the angular dependence of the four different cross-sections at a velocity of $3900\,\mathrm{km\,s^{-1}}$. The $w = 300\,\mathrm{km\,s^{-1}}$ line shows highly anisotropic scattering, with a majority of low-θ scattering events, while for $w = 10\,000\,\mathrm{km\,s^{-1}}$ the scattering is almost isotropic ($p(\theta) \propto \sin\theta$).

$\theta_{\rm min} = 0.00625$ and 0.1. The results from these tests were in agreement with each other and the $\theta_{\rm min} = 0.025$ results.

6.5.2 Yukawa-potential scattering

Having shown that using isotropic scattering to emulate the effects of the KVI anisotropic scattering would lead to an underprediction in the measured offset between DM and galaxies, we now perform similar tests with the more complicated differential cross-section described by equation (6.12). As well as having a velocity dependence, this cross-section has an angular dependence that changes with velocity. This means that the behaviour of such a particle in galaxy cluster collisions could be very different from in the cores of dwarf galaxies, due to the very different velocity scales and the anisotropic nature of a cluster collision.

Simulated cross-sections

We simulate four different cross-sections, with $w = 300, 1000, 3000$ and $10\,000\,{\rm km\,s^{-1}}$. The relative velocity between the two DM haloes in our simulations is $3900\,{\rm km\,s^{-1}}$ at the time of core passage, so this range of w values was chosen to bracket inter-halo scattering in the isotropic regime ($w = 10\,000\,{\rm km\,s^{-1}}$) all the way down to Rutherford-like scattering ($w = 300\,{\rm km\,s^{-1}}$). With fixed σ_0 the low-w cross-sections would have much lower $\sigma_{\tilde{T}}$ at $\sim 3900\,{\rm km\,s^{-1}}$ than those with high-w. In order to keep the offsets with the different cross-sections measurable, we normalise the different cross-sections such that $\sigma_{\tilde{T}}(v = 3900\,{\rm km\,s^{-1}}) = 0.5\,{\rm cm^2\,g^{-1}}$, the same $\sigma_{\tilde{T}}$ as isotropic scattering with $\sigma/m = 1\,{\rm cm^2\,g^{-1}}$.

The four simulated cross-sections are displayed in Fig. 6.6, which shows both $\sigma(v)$ and $\sigma_{\tilde{T}}(v)$, as well as the angular dependence, $p(\cos\theta)$, at $v = 3900\,{\rm km\,s^{-1}}$. In the $v > w$ regime $\sigma_{\tilde{T}}$ rises rapidly towards low velocities, such that for these particle models, tests on smaller scales may provide better constraints. Nevertheless, simulations of dwarf galaxies have shown that at velocities $\sim 40\,{\rm km\,s^{-1}}$ the cross-section could be as large as $50\,{\rm cm^2\,g^{-1}}$ (Elbert et al., 2015) without being in tension with observations of Milky Way or Local Field dwarf galaxies, such that these models may not be as outlandish as they first appear. Even if

such large cross-sections cannot be accommodated at low velocities, these differential cross-sections correspond to a perturbative treatment of Yukawa scattering. Tulin et al. (2013a) have shown that quantum mechanical and non-perturbative effects can become important when $\alpha_\chi m_\chi / m_\phi \gtrsim 1$ and $m_\chi v / m_\phi c \lesssim 1$ respectively. In this "resonant regime", quasi-bound states in the potential can lead to resonances or antiresonances that could alter the cross-section substantially at low velocities. Using a model such as the $w = 300\,\mathrm{km\,s^{-1}}$ one is therefore interesting as it probes what would happen in a galaxy cluster collision, where the collision speed places inter-halo scatters deep within the anisotropic regime.

DM-galaxy offsets

Fig. 6.7 shows the measured DM-galaxy offsets for the different Yukawa cross-sections at the time of the observed Bullet Cluster – defined as the snapshot where the separation between the two haloes is closest to $720\,\mathrm{kpc}$. These offsets are calculated using a shrinking circles approach down to different final radii. We stress that the offsets found through shrinking circles can be anomalously large due to the bias that comes from the presence of a nearby halo discussed in §5.4.3, and so these offsets should not be compared with observations of DM-galaxy offsets. However, the shrinking circles procedure is used as shrinking to different radii provides insight into the 2D distribution on different scales, and is useful for comparing the effects of different scattering cross-sections. To allow for a comparison with observed offsets we also plot in Fig. 6.7 the offset measured by fitting parametric models to the projected mass distribution. For all simulated cross-sections these are less than $10\,\mathrm{kpc}$ and decrease with increasing angular dependence.

The largest offset arises when the cross-section is closest to isotropic, which is surprising given that these cross-sections were matched to have the same $\sigma_{\tilde{T}}$ at the collision velocity of the two DM haloes and in Fig. 6.5 we demonstrated that the more anisotropic scattering cross-section (KVI) lead to larger DM-galaxy offsets than the $\sigma_{\tilde{T}}$-matched isotropic cross-section. To investigate this apparent discrepancy further, we now isolate the effects of angular and velocity dependence by running $\sigma_{\tilde{T}}$-matched isotropic versions of our Yukawa cross-sections, as well as Yukawa cross-sections with the velocity dependence removed.

$\sigma_{\tilde{T}}$-matched Yukawa scattering

In an isolated halo, Fig. 6.4 demonstrated that using $\sigma_{\tilde{T}}$-match allows us to use velocity-dependent isotropic scattering to predict the effects of Yukawa scattering. To test whether this still works in a system with a strong directionality, we plot the DM-galaxy offsets when using $\sigma_{\tilde{T}}$-match in Fig. 6.7. For the high-w cross-sections (where scattering at $v \approx 3900\,\mathrm{km\,s^{-1}}$ is fairly isotropic anyway) this procedure is effective and the results are similar to those from using the full differential cross-section. For the low-w cross-sections $\sigma_{\tilde{T}}$-match underpredicts the separations from using the full differential cross-section.

The effects of velocity-dependent scattering on DM-galaxy offsets

The DM-galaxy offsets with $\sigma_{\tilde{T}}$-matched Yukawa scattering are in agreement with our findings in §6.5.1 that isotropic scattering leads to smaller offsets than anisotropic scattering when the cross-sections have the same $\sigma_{\tilde{T}}$. However, they leave the question of why the most isotropic cross-section ($w = 10\,000\,\mathrm{km\,s^{-1}}$) leads to offsets that are substantially larger than for the more anisotropic cases. There are two primary reasons for this, both related to the velocity dependence of the anisotropic cross-sections:

1. While the relative velocity between the centres of mass of the two DM haloes is $\sim 3900\,\mathrm{km\,s^{-1}}$ during core passage, the velocity of particles within their own haloes transverse to the collision axis means that for inter-halo pairs of particles the mean pairwise velocity is larger than $3900\,\mathrm{km\,s^{-1}}$. Assuming isotropic velocity dispersions in the two haloes, with 1D velocity dispersions of $1200\,\mathrm{km\,s^{-1}}$ and $600\,\mathrm{km\,s^{-1}}$ for the main and bullet halo respectively leads to an average pairwise velocity of $\sim 4350\,\mathrm{km\,s^{-1}}$. At this velocity, the most anisotropic cross-sections, which have the steepest $\sigma_{\tilde{T}}(v)$, have the lowest $\sigma_{\tilde{T}}$ as they were normalised to have the same $\sigma_{\tilde{T}}$ at a lower velocity, $3900\,\mathrm{km\,s^{-1}}$.

2. The steep gradient of $\sigma(v)$ around $v = 3900\,\mathrm{km\,s^{-1}}$ in the low-w models means that pairs of particles with low pairwise velocities are significantly more likely to scatter than those with high pairwise velocities. This means that of the particles in the bullet halo, it is those moving in the opposite direction to the motion of the bullet through

the main halo that are most likely to scatter with a particle from the main halo. Preferentially scattering those particles travelling backwards over those travelling forwards leads to a forward shift in the position of the DM halo compared to all particles scattering with equal probability. This in turn reduces the DM-galaxy separation.

The second reason is elucidated in Fig. 6.8 where we show the DM-galaxy offsets with isotropic scattering and power-law velocity-dependent cross-sections. We simulate cross-sections of the form

$$\frac{\sigma(v)}{m} = \left(\frac{v}{4350\,\mathrm{km\,s^{-1}}}\right)^{-\alpha}\,\mathrm{cm^2\,g^{-1}}, \tag{6.24}$$

such that all cross-sections have $\sigma/m = 1\,\mathrm{cm^2\,g^{-1}}$ at the average pairwise velocity for particles drawn randomly from the two different haloes. To cut-off the low velocity divergence in the cross-section, we capped the cross-sections from equation (6.24) to $100\,\mathrm{cm^2\,g^{-1}}$.

We find that despite having the highest rate of inter-halo scattering, the $\alpha = 4$ case also has the lowest DM-galaxy offsets, even though scattering is isotropic in all cases. This is explained by the selection effect of a velocity dependent cross-section, such that those particles that scatter preferentially had certain properties (see bottom panel of Fig. 6.8). Using terminology whereby the bullet halo moves to the 'right': not just are bullet halo particles that scatter more likely to have been moving left relative to the bullet halo, the particles from the main halo with which they scatter are likely to be moving right relative to the main halo. This means that with large α the majority of scatters take place with a relative velocity lower than the mean pairwise velocity, and so transfer less momentum between the two haloes. Also, if scattered particles are ejected from the bullet halo, and these scattered particles were preferentially moving left, the remaining DM particles will preferentially be moving right with respect to the bullet halo, which pushes the measured position of the DM halo right and reduces the DM-galaxy offset.

Isolating Yukawa scattering's angular-dependence

Returning to Fig. 6.7, we isolate the effects of the angular-dependence of Yukawa scattering at different v/w by looking at the DM-galaxy separations in the Bullet Cluster including DM scattering that uses the angular dependence and normalisation of our Yukawa models at

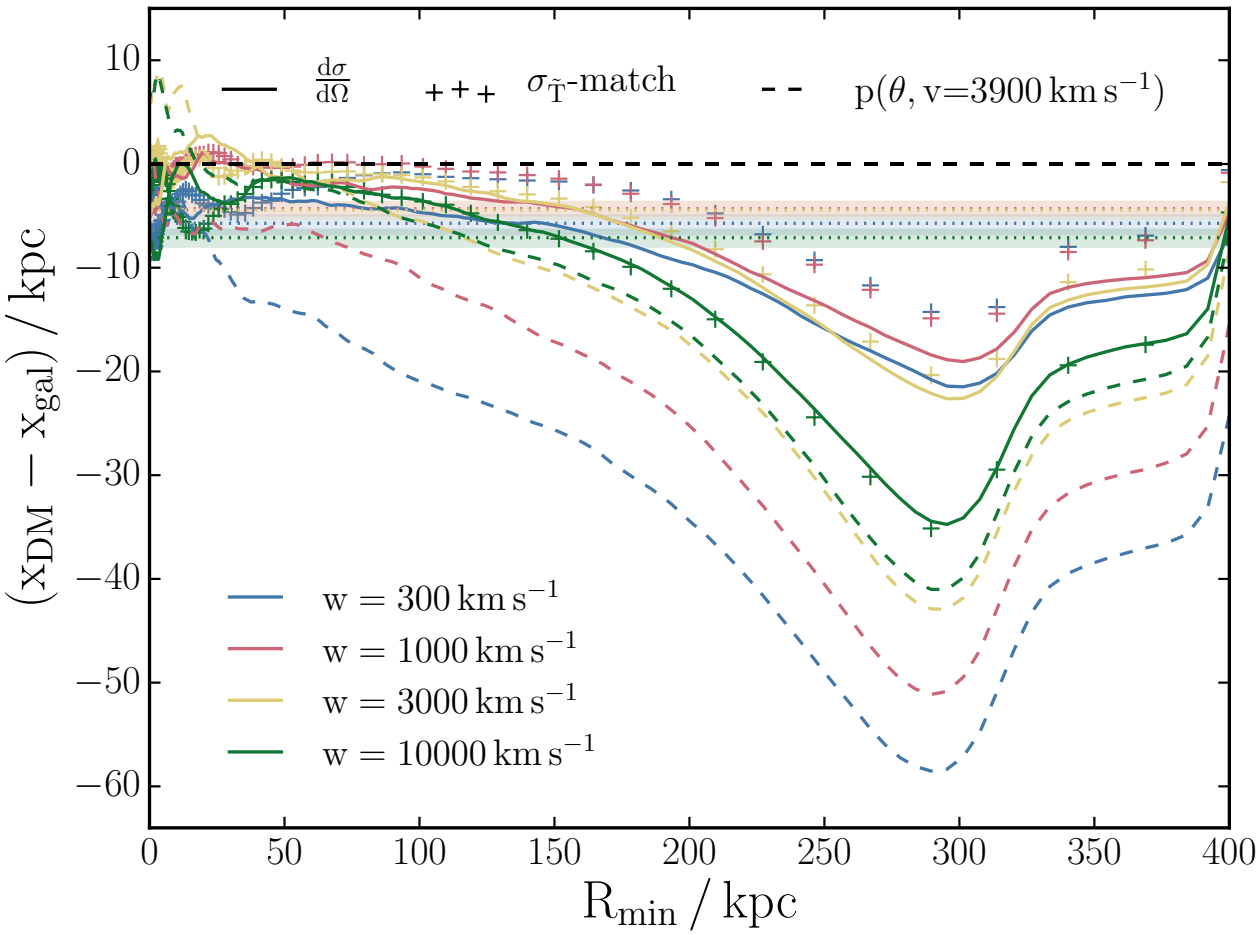


Figure 6.7: The shrinking circles offsets for different simulated particle physics models, showing how the DM-galaxy separation varies as a function of the final radius to which the circles are shrunk. The solid lines show the results using the full differential cross-section for Yukawa scattering under the Born approximation, while the crosses show the results of trying to mimic this scattering using suitably matched isotropic scattering (matched using $\sigma_{\tilde{T}}$). The dashed lines show what happens when we remove the velocity dependence of the Yukawa scattering models, by using the Yukawa differential cross-section at $v = 3900\,\mathrm{km\,s^{-1}}$ at all velocities. Finally, the horizontal bands represent the DM-galaxy offsets for the full differential cross-section cases, as measured by fitting parametric models to the projected surface density.

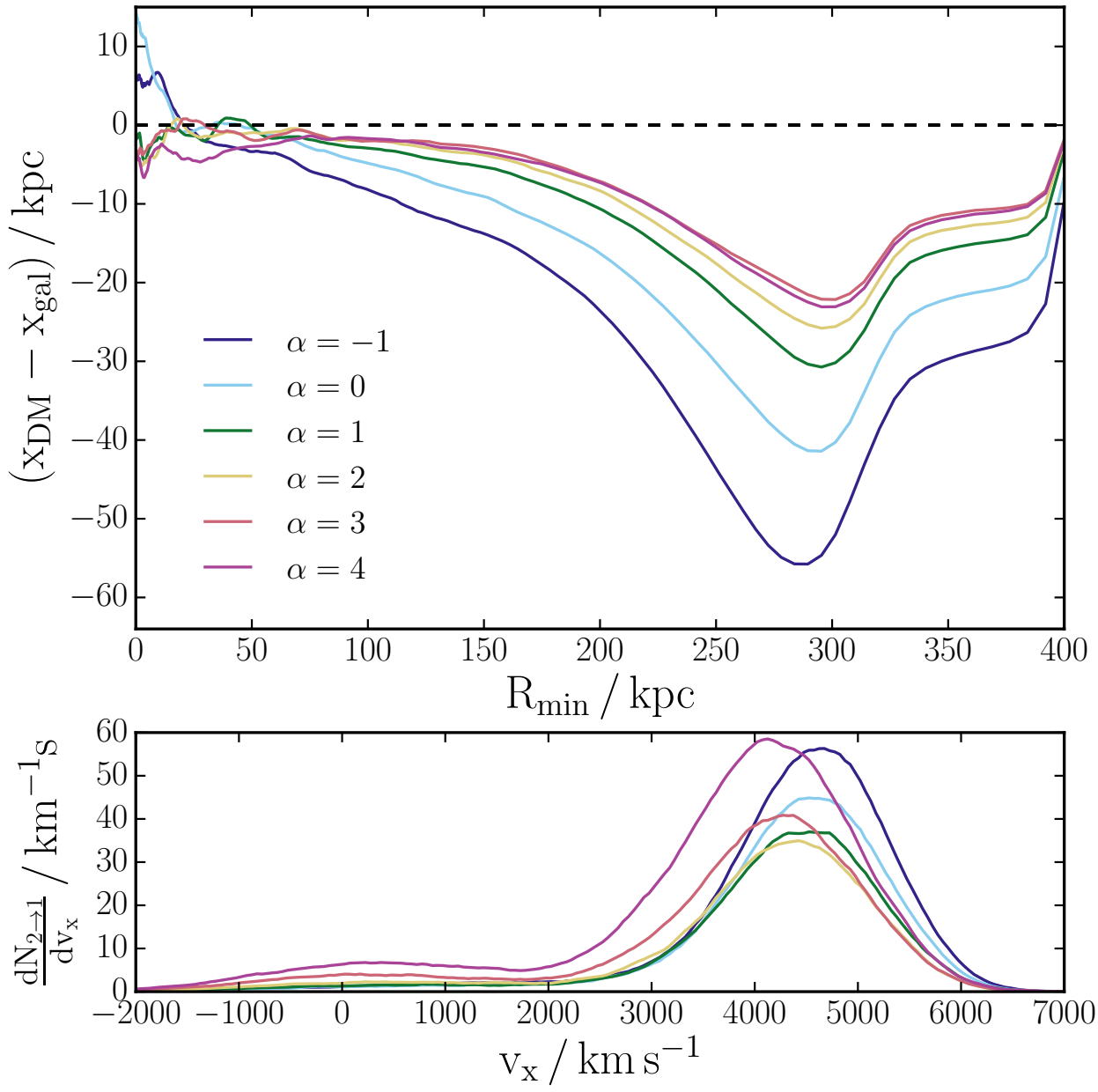


Figure 6.8: Top panel: The DM-galaxy offset at the time of the observed Bullet Cluster with isotropic DM scattering with a power-law velocity-dependence: $\sigma(v)/m = \left(v/4350\,\mathrm{km\,s^{-1}}\right)^{-\alpha}\ \mathrm{cm^2\,g^{-1}}$. The cross-section with the strongest velocity-dependence ($\alpha = 4$) has a smaller offset than the velocity-independent case ($\alpha = 0$), despite the former having a greater fraction of bullet halo particles that scatter from particles in the main halo. Bottom panel: The distribution of v_x, the velocity along the collision axis, for particles from the bullet halo before they scatter with a particle from the main halo. As well as changing the total number of scattered particles, increasing α shifts the distribution towards lower v_x.

$v = 3900\,\mathrm{km\,s^{-1}}$ at all velocities. For these cases we recover our previous result that more anisotropic scattering, in this case lower w, results in larger offsets at fixed $\sigma_{\tilde{T}}$. The smaller offsets with smaller w seen for the full differential cross-section runs do not contradict our findings in §6.5.1. Rather, the effects of a steeply decreasing $\sigma_{\tilde{T}}(v)$ more than compensate for the increased angular dependence.

6.6 Conclusions

We have explored simulations of SIDM where the scattering is anisotropic. Anisotropic cross-sections arise when the rate of scattering depends on the amount of exchanged momentum, and are natural in models with a velocity-dependent DM scattering cross-section. We considered two different models of anisotropic scattering, one without any velocity-dependence (KVI) and one that has a total cross-section and angular dependence that varies with velocity and corresponds to Yukawa scattering under the Born approximation.

For both of these anisotropic models the evolution of an isolated halo could be adequately captured by treating the scattering as isotropic (Fig. 6.2 and Fig. 6.4), provided that the isotropic cross-section is suitably matched to the underlying model. We find that what needs to be matched between different cross-sections in order for them to behave in a similar way is $\sigma_{\tilde{T}}$, defined in equation (6.4). This is similar to the momentum transfer cross-section, σ_T, that has been used by previous authors (Vogelsberger et al., 2012; Vogelsberger & Zavala, 2013; Zavala et al., 2013; Vogelsberger et al., 2014b) to match an underlying particle physics model on to a velocity-dependent but isotropic scattering cross-section that is more easily simulated. For cross-sections that are close to isotropic the matching scheme chosen is not particularly important, but when the scattering is highly anisotropic (with the majority of particles scattering by $\theta \ll \pi$) there is a factor of 2 difference in the σ_T-matched and $\sigma_{\tilde{T}}$-matched isotropic cross-sections. This is because σ_T overestimates the ability of isotropic scattering to alter dynamics, because scattering by large angles ($\sim \pi$) leads to a large amount of momentum transfer, despite leaving the system relatively unchanged (the two particles have simply switched places). In Fig. 6.4 we demonstrate that this results in σ_T-matched isotropic scattering underpredicting the effects of an anisotropic cross-section,

with cores using σ_T-matched isotropic scattering evolving slower than using $\sigma_{\tilde{T}}$-matched isotropic scattering, which in turn agreed with the results of using the full anisotropic cross-section.

We went on to investigate how the $\sigma_{\tilde{T}}$ matching scheme works in a system that has a strong directionality, namely Bullet Cluster-like galaxy cluster collisions. With an anisotropic but velocity-independent cross-section, we found that the distribution of DM was not correctly captured by using matched isotropic scattering, which underpredicted the size of DM-galaxy offsets induced by KVI scattering by $\sim 33\%$. For the case of Yukawa scattering in a galaxy cluster collision we found that the strong velocity dependence of the cross-section in regimes where the cross-section is anisotropic, leads to a suppression of the DM-galaxy offsets. Using matched isotropic scattering still underpredicts the DM-galaxy offset (crosses in Fig. 6.8), but these offsets are small anyway due to the velocity dependence. This suppression of DM-galaxy offsets is not simply because velocity dependent cross-sections must be small at typical galaxy cluster velocities to be reasonable at lower velocities. In fact, the small offsets result even when the velocity dependent cross-sections are boosted to have a substantial $\sigma_{\tilde{T}}$ at cluster velocities. The small offsets are a result of the gradient in $\sigma(v)$, which results in particle pairs with low relative velocities being more likely to scatter than others. These low velocity pairs are made of particles that move within their halo in the opposite direction to the bulk velocity of their halo, and preferentially scattering these particles leaves a population of unscattered particles moving faster than the bulk velocity of the halo. This shifts the measured DM position forwards reducing any DM-galaxy offset.

Ignoring the angular-dependence of SIDM models and instead using suitably matched isotropic cross-sections appears to work well in isotropic systems such as an isolated halo, but can lead to differences from the true result in anisotropic systems. Despite these differences, merging galaxy clusters do not appear to be a good place to constrain Yukawa-like DM scattering, as the cross-section at cluster velocities would be lower than in smaller objects, and the increased DM-galaxy separation due to the anisotropic nature of the scattering is more than compensated for by the decreased DM-galaxy separation coming from the gradient in $\sigma(v)$ about the collision velocity of the clusters. Previous results that have simulated an anisotropic scattering model using appropriately matched isotropic scatter-

ing have typically focused on the density profiles of dwarf galaxies. Our results in isolated haloes suggest these results are probably robust to changing from isotropic scattering to using the underlying differential cross-section. That being said, a cosmologically formed DM halo evolves through numerous mergers, and it is unclear if incorrectly modelling the effects of SIDM in these mergers could lead to differences in the final density profile. This will need to be addressed in the future by including anisotropic scattering in cosmological simulations.

Chapter 7

Conclusions

Despite an ever growing amount of astrophysical evidence pointing towards the existence of DM as the dominant matter component in the Universe (e.g. Planck Collaboration et al., 2016), we still do not know what the DM is. Terrestrial experiments hoping to detect DM (Bertone et al., 2005), offer hope that one day we may know the identity of the DM particle(s), but for now DM has only been detected through its gravitational effects on the things that we can see. It is therefore prudent to ask what we can learn about the nature of DM particles from the way they behave astrophysically. In this thesis, we investigated some of the consequences of a dark sector in which DM particles can elastically scatter from one another – known as SIDM. A significant cross-section for DM–DM scattering is one of the few properties of the DM that can be probed by its astrophysical behaviour, alongside some DM–SM interactions (Bœhm et al., 2014) and the mass of DM particles (e.g. Hu et al., 2000; Alcock et al., 2000; Lovell et al., 2014).

We began in Chapter 2 by introducing a method to estimate the rate of SIDM scattering in collapsed structures throughout the history of the universe, for different particle physics models of SIDM. For the simplest (velocity-independent) cross-sections, our results match the canonical picture in which most scattering occurs in massive structures $\gtrsim 10^{12}\,\mathrm{M}_{\odot}$ at late times $z \lesssim 1$. The main source of uncertainty in this calculation was the concentration-mass relation for DM haloes, with different relations predicting a total number of interactions over the history of the Universe that vary by a factor of $\lesssim 2$.

We then explored how this prediction changes, if the DM has a well-motivated velocity-

dependent cross-section. This moves most scattering into low mass objects $\lesssim 10^4$ $\mathrm{M}_\odot$ at early times $z \gtrsim 7$. The scattering rate as a function of redshift peaks at $z \sim 20$, with a pronounced peak two orders of magnitude higher than the scattering rate at the present day. This has important consequences for attempts to simulate velocity-dependent SIDM cosmologies. Scattering in low mass haloes at high redshift would not be resolved in typical cosmological simulations, but the unresolved interactions could be important in establishing the later dynamics of particles. Not resolving these interactions may therefore lead to incorrect predictions for the structure of haloes formed from SIDM particles with a velocity-dependent cross-section.

In Chapter 3 we gave a theoretical overview describing how DM–DM scattering can be incorporated into an N-body code. We then gave a detailed account of our specific implementation of SIDM with an isotropic scattering cross-section in the GADGET code. We put particular emphasis on rarely discussed aspects of simulating SIDM; including how to handle particles scattering multiple times within a single time-step, how to implement scattering across processors, when (within a leapfrog integration scheme) to implement the scattering, and how scattering probabilities are calculated from program values in a cosmological simulation where the internal variables are not physical distances and velocities.

Having described our SIDM implementation, we then tested it in Chapter 4. We started with a simple test case of a cube of SIDM particles moving through a uniform background of stationary particles. For this system, we demonstrated that both the rate of DM scattering and the distribution of velocities of scattered particles agreed with analytical predictions. We also explained the mechanism that incorrectly lead Rocha et al. (2013) to conclude that their SIDM implementation would under-predict the scattering rate in low density regions.

We then investigated the more astrophysically relevant situation of scattering in a DM halo, demonstrating that the radial distribution of scattering events within a halo agreed with the analytical prediction at all but the smallest of radii, where the predicted rate is high owing to the divergent central density of our test (Hernquist) halo. The scattering rate dropped below the analytical prediction at radii smaller than the neighbour search radius used to find potential scattering partners. This is because the neighbour search radius acts as a smoothing scale on which the local density is estimated for the purposes of scattering.

With this in mind, we then discussed how to choose the neighbour search radius in order to give the best results, finding that a search radius similar in size to the gravitational softening was a good choice. We allowed our Hernquist halo to evolve under the influence of DM scattering, and demonstrated that for moderate cross-sections the density profile evolved a constant density core, while for sufficiently large cross-sections the halo would undergo core collapse, resulting in an extremely dense halo centre.

We ended Chapter 4 by showing the rate of scattering in a simulation of a cosmological periodic box, compared with the predicted rate of scattering from Chapter 2. The two were in good agreement, acting as a sanity check of both our SIDM implementation and our framework for analytically predicting the cosmic scattering rate.

Chapter 5 then used the SIDM implementation introduced in Chapter 3 to simulate the merging pair of galaxy clusters known as the Bullet Cluster. If DM particles can interact with each other, the scattering leads to a transfer of momentum between the two colliding DM haloes, causing the DM haloes to lag behind the (collisionless) galaxies. Analytical models in the literature have regarded this as a drag force, which decelerates all of the DM. However, with isotropic DM scattering, this transferred momentum is not evenly distributed between the particles in each halo. Most DM particles do not scatter with a particle from the other halo, while a subset of them do and receive large momentum kicks. These scattered particles lag behind the galaxies and unscattered DM, shifting the measured position of the DM halo relative to the galaxies, but in a way that depends sensitively on the method used to measure the positions of the different components. In particular, shrinking circles methods similar to those used by Randall et al. (2008) 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 $\sigma/m < 1.25\,\mathrm{cm}^2\,\mathrm{g}^{-1}$ constraint placed on the cross-section for DM scattering by Randall et al. (2008) is overstated, with cross-sections as large as $\sigma/m = 2\,\mathrm{cm}^2\,\mathrm{g}^{-1}$ producing offsets consistent with observations of the Bullet Cluster.

Finally, in Chapter 6, we extended our implementation of SIDM scattering to include velocity-dependent and anisotropic cross-sections, with a method that allows SIDM with an arbitrary differential cross-section to be simulated. We considered two different models

of anisotropic scattering: one without any velocity-dependence (Kahlhoefer et al., 2014) and one with a total cross-section and angular dependence that varies with velocity, in a manner similar to that predicted by various particle models of SIDM (Loeb & Weiner, 2011; Tulin et al., 2013b; Cyr-Racine et al., 2016).

For both of these anisotropic models, we discovered that the evolution of an isolated halo could be adequately captured by treating the scattering as isotropic, so long as the cross-section (hence rate of scattering) is appropriately reduced to account for the fact that isotropic scattering is more efficient at transporting energy and momentum than predominantly small-angle scattering. In practice, this means matching the momentum transfer cross-section between the anisotropic cross-section and the isotropic one used to mimic it. We stress that this momentum transfer cross-section should take into account that particles scattering by large angles ($> \pi/2$), could be re-labelled so as to have scattered by a smaller angle and have transferred less momentum. In previous simulations that have emulated the effects of an anisotropic cross-section using an isotropic one (Vogelsberger et al., 2012; Vogelsberger & Zavala, 2013; Zavala et al., 2013; Vogelsberger et al., 2014b), this re-labelling has not been considered, leading to effective scattering rates incorrect by up to a factor of two from the ones that would most faithfully reproduce the effects of the particle models they were trying to simulate.

We then investigated the effects of anisotropic scattering in a system that is itself highly anisotropic, running simulations of the Bullet Cluster that were similar to those in Chapter 5. We found that for velocity-independent cross-sections, matched to have the same rate of core formation in an isolated halo, scattering that was more anisotropic produced larger offsets between the DM and galaxies. However, when the cross-section was a steeply declining function of velocity, the offsets shrank because particle pairs with low relative velocities are more likely to scatter than others. These low velocity pairs are made of particles that move within their halo in the opposite direction to the bulk velocity of their halo, and preferentially scattering these particles leaves a population of unscattered particles moving faster than the bulk velocity of the halo. This shifts the measured DM position forwards, reducing any DM-galaxy offset. A strong velocity-dependence in the cross-section typically arises for the same particle physics reasons as anisotropic scattering, such that one tends to

imply the other. Unfortunately this means that the large offsets with anisotropic scattering will tend to be suppressed by the velocity-dependence. The net effect is that, even for these models of SIDM, existing constraints in the literature have again been overestimated. SIDM remains a viable theory, not only consistent with all current observations, but potentially able to explain cosmology's remaining small-scale inconsistencies.

Appendix A

The expected number of scattering events for a cube moving through a uniform background

As mentioned in §4.1, the expected number of scattering events in our test case of a cube moving through a uniform background is more complicated than that given in equation (4.4). This is because the first cube particles to pass through the background (which initially has a number density of particles n_{b}), scatter with some of the background particles, lowering the background density seen by later cube particles. The system can be modelled by breaking the cube up into infinitesimal slabs of thickness $\mathrm{d}x$, with x the distance of a particular

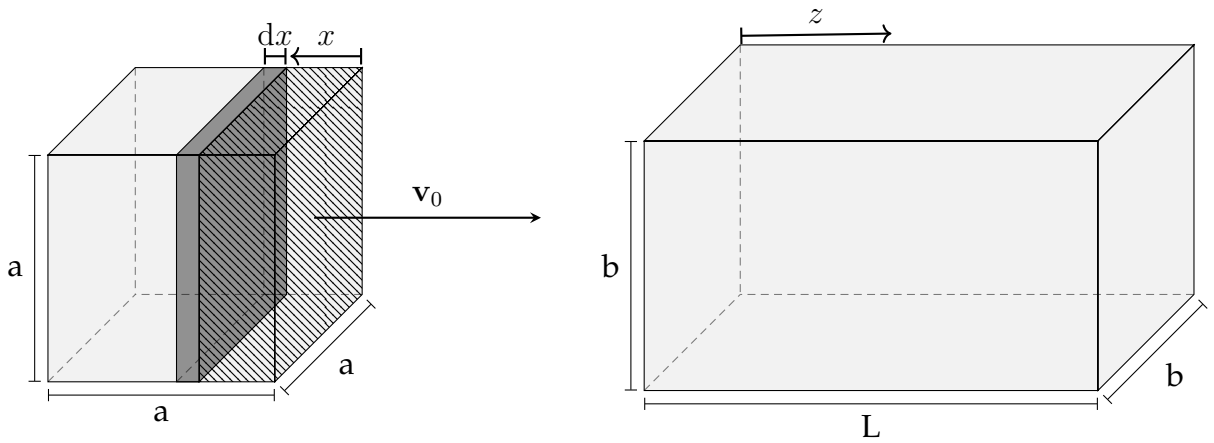


Figure A.1: The test-case setup in Fig. 4.1, but additionally defining z, x and $\mathrm{d}x$.

slab from the front edge of the cube. This setup is shown in Fig. A.1.

The expected number of scattering events as the slab $\mathrm{d}x$ passes through the background cuboid is still given by an equation similar to (4.4), but now accounting for the fact that the number density of background particles is not constant with z, and depends on x i.e. what fraction of cube particles have already passed through the background. The expected number of scattering events involving particles from the slab $\mathrm{d}x$ is

$$\mathrm{d}N_{\mathrm{exp}} = \frac{N_{\mathrm{c}}\,\mathrm{d}x}{a}\left[1 - \exp\left(-\sigma_{\mathrm{p}}\int n(z,x)\,\mathrm{d}z\right)\right], \tag{A.1}$$

where $N_{\mathrm{c}}\,\mathrm{d}x/a$ is the number of particles initially in the slab, and $n(z,x')$ is the background number density as a function of z,[1] after cube particles with $x < x'$ have passed through the background cuboid. We define $N_{\mathrm{exp}}(x')$, to be the expected number of scattering events involving particles with $x < x'$, i.e. those particles from the hatched region in Fig. A.1. Each scattering event decreases the number of background particles by one, such that

$$\int n(z,x)\,\mathrm{d}z = n_{\mathrm{b}}L - N_{\mathrm{exp}}(x)/a^2. \tag{A.2}$$

Combining equations (A.1) and (A.2)

$$\frac{\mathrm{d}N_{\mathrm{exp}}}{\mathrm{d}x} = \frac{N_{\mathrm{c}}}{a}\Big[1 - \exp\left\{-\sigma_{\mathrm{p}}\left(n_{\mathrm{b}}L - N_{\mathrm{exp}}(x)/a^2\right)\right\}\Big], \tag{A.3}$$

which integrates to

$$\int_0^{N_{\mathrm{exp}}(x)} \frac{\mathrm{d}N'_{\mathrm{exp}}}{1 - \exp\left\{-\sigma_{\mathrm{p}}\left(n_{\mathrm{b}}L - N'_{\mathrm{exp}}/a^2\right)\right\}} = \int_0^x \frac{N_{\mathrm{c}}}{a}\,\mathrm{d}x'. \tag{A.4}$$

The integral on the left hand side of equation (A.4), can be solved by making the substitution $u = \sigma_{\mathrm{p}}(n_{\mathrm{b}}L - N'_{\mathrm{exp}}/a^2)$, and using the indefinite integral $\int(1-\mathrm{e}^{-u})^{-1}\mathrm{d}u = \ln(\mathrm{e}^u - 1)$. Using these, we arrive at

$$\frac{-a^2}{\sigma_{\mathrm{p}}}\ln\left[\frac{\exp\left\{\sigma_{\mathrm{p}}\left(n_{\mathrm{b}}L - N_{\mathrm{exp}}(x)/a^2\right)\right\} - 1}{\exp\left\{\sigma_{\mathrm{p}}n_{\mathrm{b}}L\right\} - 1}\right] = \frac{N_{\mathrm{c}}}{a}x, \tag{A.5}$$

which we can rearrange for $N_{\mathrm{exp}}(x)$ and evaluate at $x = a$, with $N_{\mathrm{exp}} \equiv N_{\mathrm{exp}}(a)$

$$N_{\mathrm{exp}} = n_{\mathrm{b}}a^2L - \frac{a^2}{\sigma_{\mathrm{p}}}\ln\left[1 + \exp\left(\frac{-\sigma_{\mathrm{p}}N_{\mathrm{c}}}{a^2}\right)\left\{\exp\left(\sigma_{\mathrm{p}}n_{\mathrm{b}}L\right) - 1\right\}\right]. \tag{A.6}$$

Evaluating this with the parameters used in §4.1, predicts an N_{exp} that is 0.9% lower than when using equation (4.4). This difference is smaller than the expected Poisson noise.

[1] Specifically, the number density in the central $a \times a$ cross-section of the cuboid, through which the cube passes.

Appendix B

A guide to SIDM with a Yukawa potential

Attempts to explain the small-scale challenges to ΛCDM by invoking SIDM are hampered by the tight constraints on the DM–DM scattering cross-section in galaxy clusters. However, these cluster constraints can be evaded if the cross-section decreases with increasing pairwise velocity, such that scattering rates are negligible in cluster sized haloes with typical velocities of $1000\,\mathrm{km\,s^{-1}}$, but can be significant in the smaller haloes that host dwarf galaxies. A natural particle physics model for the DM that would give rise to such a cross-section is scattering mediated by a new gauge boson of mass m_ϕ, resulting in an attractive Yukawa potential.

In this appendix, we introduce a simple model of scattering through a Yukawa potential. While this is not the model we use in either Chapter 2 or 6, it hopefully serves to increase the reader's intuition for velocity-dependent and anisotropic SIDM models. We also derive results for scattering in a Coulomb potential, that we make use of in §3.1.2.

B.1 Rutherford scattering

To introduce some notation and gain some intuition, we will first look at the case of Rutherford scattering, originally the scattering of alpha particles from nuclei (Rutherford, 1911). If we consider a particle with mass m and charge e scattering from a more massive particle (a nucleus) with charge Ze then the repulsive Coulomb force between these two particles

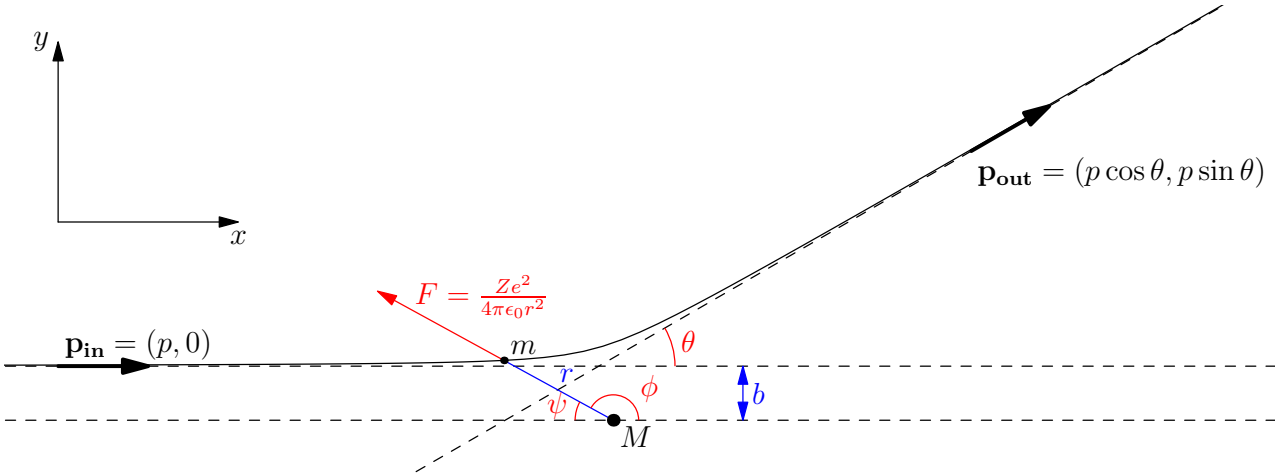


Figure B.1: The motion of a particle in a central repulsive Coulomb potential follows a hyperbolic trajectory. From the impact parameter, b, and initial momentum, p, one can calculate the polar scattering angle, θ.

is

$$F_C(r) = \frac{Ze^2}{4\pi\epsilon_0 r^2}. \tag{B.1}$$

We assume that the more massive particle is infinitely massive, and work in the frame of that particle, in which we have a particle of mass m moving in a central potential. The incoming particle starts infinitely far from the origin, with momentum $p = mv$ along the x-axis, and impact parameter b as shown in Fig. B.1. As the scattering is elastic, the particle will leave with $|\mathbf{p}_{\rm out}| = |\mathbf{p}_{\rm in}| \equiv p$, but now at an angle θ from the x-axis, such that $\mathbf{p}_{\rm out} = (p\,\cos\theta, p\,\sin\theta)$[1].

We label the angle between the x-axis and the line joining the centre of the potential to the position of the particle at time t, $\phi(t)$, with $\phi(\infty) \equiv \theta$. To calculate how the scattering angle θ depends on the impact parameter and particle momentum we can equate the time-integral of the y-component of the force on the particle, to the change in momentum along the y-axis

$$p\,\sin\theta = \int_{-\infty}^{\infty} F_C(r)\sin\phi\,\mathrm{d}t = \int_{-\infty}^{\infty} \frac{Ze^2}{4\pi\epsilon_0 r^2}\sin\phi\,\mathrm{d}t. \tag{B.2}$$

We proceed by changing integration variable to ϕ, and use conservation of angular momentum to relate the initial angular momentum to that at some later time, $mvb = -mr^2\dot{\phi}$,

[1]The motion of a particle in a central potential remains in the plane defined by its initial position and velocity, so we can work in just two dimensions.

so that

$$\frac{1}{r^2}\frac{\mathrm{d}t}{\mathrm{d}\phi} = \frac{-1}{vb}. \tag{B.3}$$

Using this in equation (B.2)

$$p\,\sin\theta = \int_{\pi}^{\theta} \frac{-Ze^2}{4\pi\epsilon_0 bv}\sin\phi\,\mathrm{d}\phi = \frac{Ze^2}{4\pi\epsilon_0 bv}(1+\cos\theta). \tag{B.4}$$

Then using $p = mv$ and the identity $\frac{\sin\theta}{1+\cos\theta} = \tan\frac{\theta}{2}$

$$b = \frac{Ze^2}{4\pi\epsilon_0 mv^2\tan\frac{\theta}{2}} \tag{B.5}$$

and

$$\left|\frac{\mathrm{d}b}{\mathrm{d}\theta}\right| = \frac{Ze^2}{8\pi\epsilon_0 mv^2\sin^2\frac{\theta}{2}}. \tag{B.6}$$

From Fig. B.2, the infinitesimal annulus containing trajectories with impact parameters in the range $[b, b+\mathrm{d}b]$, has an area $\mathrm{d}\sigma = 2\pi b\,\mathrm{d}b$. Employing the identity $\tan\frac{\theta}{2} = 2\sin^2\frac{\theta}{2}/\sin\theta$ and $\mathrm{d}\Omega = 2\pi\sin\theta\,\mathrm{d}\theta$, we find the differential cross section

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \left(\frac{Ze^2}{8\pi\epsilon_0 m}\right)^2 \frac{1}{v^4\sin^4\frac{\theta}{2}}. \tag{B.7}$$

Integrating over solid angle, the total cross-section

$$\sigma \equiv \int \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\mathrm{d}\Omega = 2\pi\int \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\sin\theta\,\mathrm{d}\theta \tag{B.8}$$

diverges in this case, a result of the infinite range of the Coulomb force.

Performing a similar calculation for an attractive Coulomb potential one finds that equations (B.5)-(B.7) remain unchanged if θ is now defined to be below the x-axis. In the attractive case the particle gets closer to the centre of the potential than for the repulsive case, leading to larger forces, but this is countered by the change of sign of the y-component of the force on the particle during the encounter. These effects cancel one another, and (for a given impact parameter and initial momentum) $|\Delta p_y| = p\sin\theta$ is the same in the attractive case as in the repulsive case, with only the sign of Δp_y changed.

Rutherford scattering momentum transfer cross-section

A useful concept when comparing the macroscopic consequences of particle interactions with different angular dependencies for the differential cross section is the momentum-

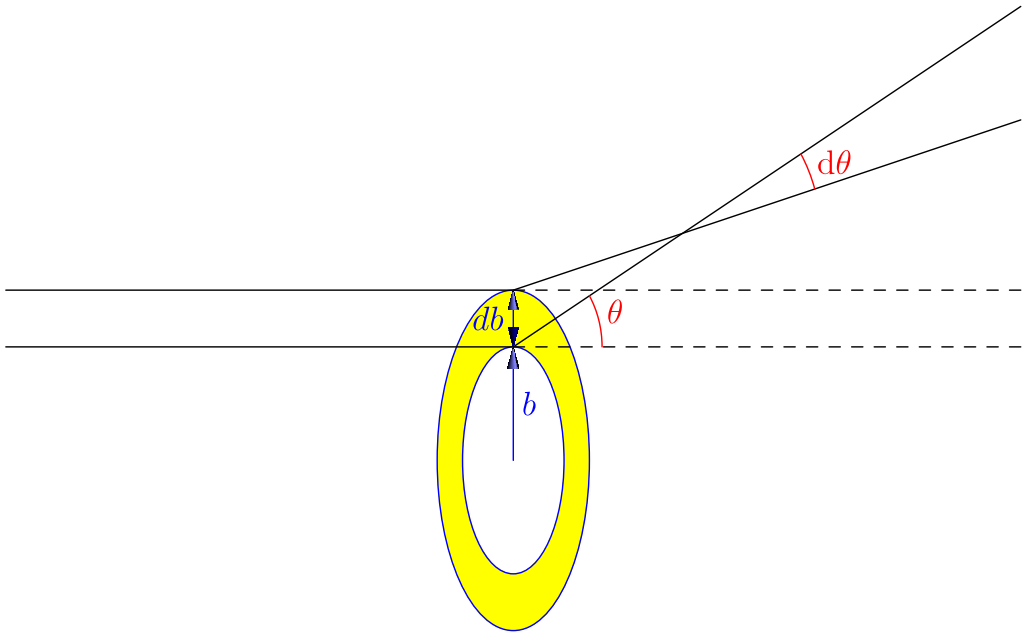


Figure B.2: A particle with impact parameter b will scatter at an angle θ, while a particle with impact parameter $b+\mathrm{d}b$ will scatter at an angle $\theta+\mathrm{d}\theta$, where here $\mathrm{d}\theta$ is negative as larger impact parameters lead to weaker scattering. The yellow annulus has an area $\mathrm{d}\sigma$, and particles with trajectories through this annulus will scatter by a polar angle in the range $[\theta, \theta + \mathrm{d}\theta]$.

transfer cross-section, described in §6.2.2. For the Coulomb interaction, the momentum-transfer cross-section is divergent. However, if we introduce a minimum angle, $\theta_{\min}$, and ignore all scattering that deflects the particle by $\theta < \theta_{\min}$, we find

$$\sigma_T(\theta_{\min}) = 2\pi \int_{-1}^{\cos\theta_{\min}} (1-\cos\theta)\left(\frac{Ze^2}{8\pi\epsilon_0 m}\right)^2 \frac{1}{v^4 \sin^4\frac{\theta}{2}}\,\mathrm{d}\cos\theta. \tag{B.9}$$

Then using the identity $(1-\cos\theta) = 2\sin^2\frac{\theta}{2}$ and integrating

$$\sigma_T(\theta_{\min}) = \frac{8\pi}{v^4}\left(\frac{Ze^2}{8\pi\epsilon_0 m}\right)^2 \ln\left(\frac{2}{1-\cos\theta_{\min}}\right), \tag{B.10}$$

which we will now use to explain the form of $\sigma_T(v)$ for scattering in a Yukawa potential.

B.2 Scattering through a Yukawa potential

The Yukawa potential has previously been used to study the screened Coulomb scattering in a plasma (Khrapak et al., 2004), for which the momentum-transfer cross section can be

approximated by

$$\frac{\sigma_T}{\sigma_T^{\max}} \approx \begin{cases} \frac{4\pi}{22.7}\,\beta^2\,\ln\left(1+\beta^{-1}\right), & \beta < 0.1 \\ \frac{8\pi}{22.7}\,\beta^2\,\left(1+1.5\beta^{1.65}\right)^{-1}, & 0.1 < \beta < 10^3 \\ \frac{\pi}{22.7}\,\left(\ln\beta + 1 - \frac{1}{2}\ln^{-1}\beta\right)^2, & \beta > 10^3, \end{cases} \tag{B.11}$$

where $\beta = \pi v_{\max}^2/v^2$ and $\sigma_T^{\max} = 22.7/m_\phi^2$ (Feng et al., 2010; Finkbeiner et al., 2011; Loeb & Weiner, 2011). Here $v_{\max} = \sqrt{2\alpha m_\phi/\pi m_\chi}$ is the velocity at which $(\sigma_T\, v)$ peaks, with $\sigma_T(v_{\max}) = \sigma_T^{\max}$.

Equation (B.11) is an analytical formula fit to the result of numerical calculations. In this section we attempt to gain insight into the form of $\sigma_T(v)$, by using a simplified model for scattering in a Yukawa potential. We assume that a particle of mass m_χ moves in the potential

$$V_Y(r) = -\alpha\frac{e^{-m_\phi r}}{r}, \tag{B.12}$$

which leads to the radial force

$$F_Y(r) = -\alpha\frac{e^{-m_\phi r}}{r^2}(1+m_\phi r). \tag{B.13}$$

Introducing the typical range of the interaction, $r_\phi \equiv 1/m_\phi$,[2] we can rewrite equation (B.13)

$$F_Y(r) = -\alpha\frac{e^{-r/r_\phi}}{r^2}(1+r/r_\phi), \tag{B.14}$$

which looks like the Coulomb force

$$F_C(r) = -\alpha\frac{1}{r^2}, \tag{B.15}$$

when $r \ll r_\phi$ and exponentially decays to zero for $r > r_\phi$.

Using the result that the Yukawa force is Coulomb-like at small distances and exponentially cut off outside of $r = r_\phi$, we can crudely approximate Yukawa scattering by treating scattering processes with impact parameter $b \leq r_\phi$ as if the scattering were from the equivalent Coulomb potential, and assuming no interaction when $b > r_\phi$.

Equation (B.10) gives the momentum-transfer cross-section for scattering in a Coulomb potential when one ignores the contribution of scattering by less than an angle $\theta_{\min}$. A

[2] We are using natural units here with $\hbar = c = 1$.

maximum impact parameter $b = r_\phi$, corresponds to a minimum scattering angle, which we calculate using the relationship (equation B.5) between impact parameter and scattering angle

$$\theta_{\rm min} = 2\tan^{-1}\left(\frac{2\alpha}{m_\chi v^2 r_\phi}\right), \tag{B.16}$$

where we have replaced $Ze^2/4\pi\epsilon_0$ from the electromagnetic case, with α, and also changed m to $m_\chi/2$ to reflect that for two equal masses the reduced mass[3] is half of the actual mass of one of the particles. Making the same substitutions in equation (B.10) we have

$$\sigma_T(v) = \frac{2\pi}{v^4}\left(\frac{2\alpha}{m_\chi}\right)^2 \ln\left(\frac{2}{1-\cos\theta_{\rm min}}\right). \tag{B.17}$$

Looking at equation (B.16) we can see that as the pairwise velocity of particles increases, the minimum scattering angle decreases. The velocity around which $\theta_{\rm min}$ transitions from being large to small is when $\alpha/m_\chi v^2 r_\phi \sim 1$, which is when the initial kinetic energy is similar to the magnitude of the potential at the typical range of the interaction, $m_\chi v^2 \sim |V_Y(r_\phi)|$. At velocities below this, particles that scatter do so with $\theta \sim \pi$, while for large velocities the majority of scatters are low momentum transfer events with small θ.

Taking $(\pi - \theta_{\rm min}) \ll 1$, we find the low-velocity limit of equation (B.17) to be

$$\sigma_T(v \to 0) = \frac{2\pi}{m_\phi^2} = 2\pi r_\phi^2. \tag{B.18}$$

This is simply the geometric cross section from the requirement that the impact parameter is less than r_ϕ, multiplied by a factor of two because when $\theta \approx \pi$ the momentum transfer is twice the average momentum transfer for an isotropic scattering process.

If instead we take the high-velocity limit of equation (B.17), we find

$$\sigma_T(v \to \infty) = \frac{4\pi}{v^4}\left(\frac{2\alpha}{m_\chi}\right)^2 \ln\left(\frac{m_\chi v^2}{2m_\phi \alpha}\right). \tag{B.19}$$

This is exactly what one gets taking the low-β limit of σ_T (equation B.11) and using the definition of β in terms of particle physics parameters.

At low velocities, this approximate treatment for Yukawa scattering agrees less well with the numerical calculation. As $v \to 0$, the approximate treatment gives a constant

[3]For two bodies of mass m_1 and m_2 separated by a distance r and with potential energy $V(r)$, the evolution of the separation is described by $\mu\ddot{r} = -\partial V/\partial r$, where $\mu = m_1 m_2/(m_1 + m_2)$ is the reduced mass.

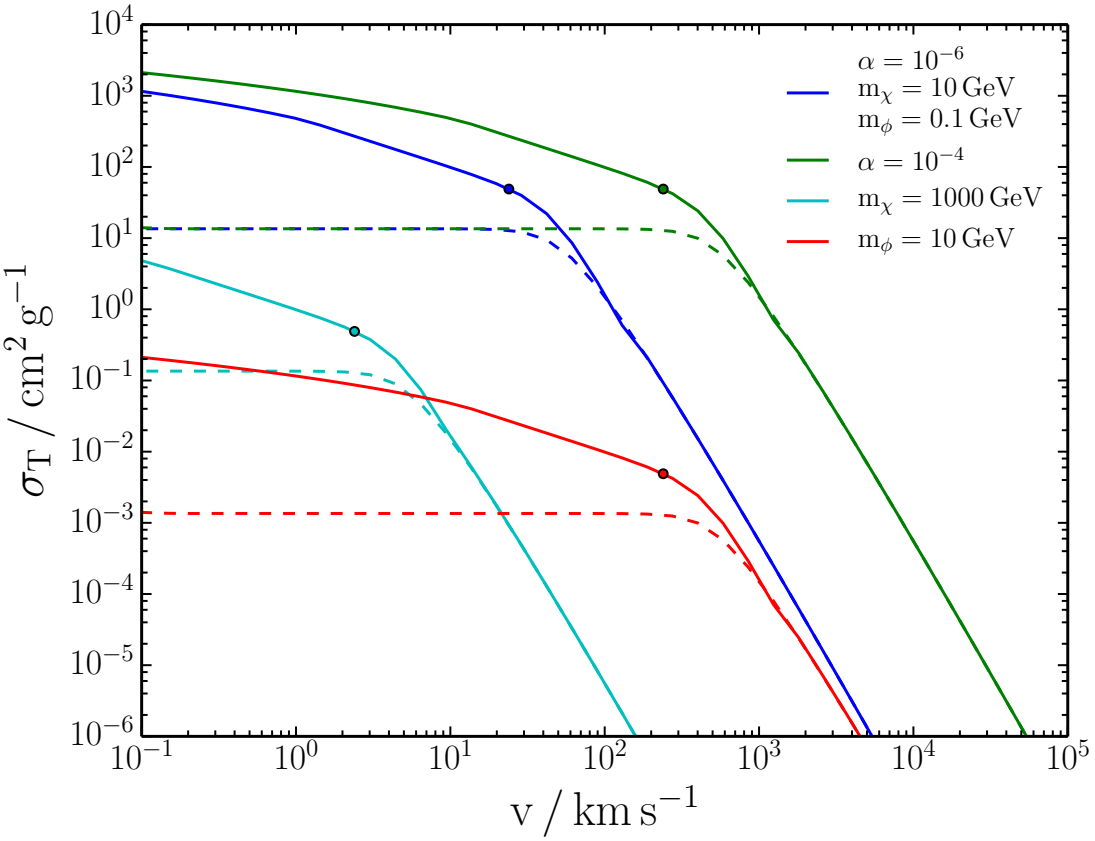


Figure B.3: The momentum transfer cross-section for a fiducial set of particle physics parameters (blue lines), as well as three models where one of the three parameters (DM mass, mediator mass and coupling strength) have been changed from their fiducial values. The numerical calculation results (equation B.11) are the solid lines, while the results from a simple model that ignores scattering with $b > r_\phi$ are drawn as dashed lines. The points along the solid lines are at $v = v_{\rm max}$. Our simple model agrees with the numerical results at velocities $\gg v_{\rm max}$, while for velocities $\lesssim v_{\rm max}$ it significantly under-predicts the true cross-section.

momentum-transfer cross section of $2\pi/m_\phi^2$, while the fitting function to the numerical result is

$$\sigma_T = \frac{\pi}{m_\phi^2}(\ln\beta + 1 - 1/2\ln\beta)^2. \tag{B.20}$$

The breakdown of our simple model (in which scattering with $b > r_\phi$ is ignored) at low velocities can be readily understood. At low velocities even the exponentially suppressed force can be sufficient to significantly scatter particles. As such, we expect our simple model to under-predict the cross-section when $v \lesssim v_{\rm max}$, as seen in Fig. B.3.

Appendix C

The contribution of mass at different radii to the projected density and shear signals

In Fig. 5.13 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 $R_{\rm min} = 200$ and $60\,{\rm kpc}$, 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. 5.12 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 ($\gamma_{\rm t}$) from a circularly symmetric mass distribution can be written in terms of the 'excess surface density'

$$\Delta\Sigma = \bar{\Sigma}(< R) - \Sigma(R) = \Sigma_{\rm crit}\gamma_{\rm t}. \tag{C.1}$$

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

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

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 $M_0 \ln(R_{\rm max}/R_0)$, where we have truncated the integration at a maximum radius $R_{\rm max}$. For $R_{\rm max}$ 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 $R_{\rm max}$, 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

$$\mathrm{SNR}_g \propto R_0\Sigma(R_0)\ln\left(\frac{R_{\rm max}}{R_0}\right). \tag{C.3}$$

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

$$\mathrm{SNR}_\Sigma \propto R_0\sqrt{\Sigma(R_0)}. \tag{C.4}$$

SNR_g and SNR_Σ are the quantities plotted in the bottom panel of Fig. 5.13, where they have been normalised by their maximum value.

Bibliography

Aalseth C. E., et al., 2011, Physical Review Letters, 106, 131301

Aalseth C. E., et al., 2014, preprint, (arXiv:1401.3295)

Ackerman L., Buckley M. R., Carroll S. M., Kamionkowski M., 2009, Phys. Rev. D, 79, 023519

Adriani O., et al., 2009, Nature, 458, 607

Agnese R., et al., 2014, Physical Review Letters, 112, 241302

Aguirre A., Schaye J., Quataert E., 2001, ApJ, 561, 550

Ahn K., Shapiro P. R., 2005, MNRAS, 363, 1092

Akerib D. S., et al., 2013, Nuclear Instruments and Methods in Physics Research A, 704, 111

Alcock C., et al., 2000, ApJ, 542, 281

Alpher R. A., Bethe H., Gamow G., 1948, Physical Review, 73, 803

Ananthaswamy A., O'Hara C., 2010, Physics Today, 63, 51

Angloher G., et al., 2012, European Physical Journal C, 72, 1971

Angulo R. E., White S. D. M., 2010, MNRAS, 401, 1796

Angus G. W., Famaey B., Zhao H. S., 2006, MNRAS, 371, 138

Aprile E., et al., 2013, Physical Review Letters, 111, 021301

Atwood W. B., et al., 2009, ApJ, 697, 1071

Balberg S., Shapiro S. L., Inagaki S., 2002, ApJ, 568, 475

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

Barrena R., Biviano A., Ramella M., Falco E. E., Seitz S., 2002, A&A, 386, 816

Bartelmann M., 2010, Classical and Quantum Gravity, 27, 233001

Bartels R., Krishnamurthy S., Weniger C., 2016, Physical Review Letters, 116, 051102

Bechtol K., et al., 2015, ApJ, 807, 50

Behroozi P. S., Wechsler R. H., Conroy C., 2013, ApJ, 762, L31

Benson A. J., Lacey C. G., Baugh C. M., Cole S., Frenk C. S., 2002, MNRAS, 333, 156

Bento M. C., Bertolami O., Rosenfeld R., Teodoro L., 2000, Phys. Rev. D, 62, 041302

Berezhiani Z. G., Dolgov A. D., Mohapatra R. N., 1996, Physics Letters B, 375, 26

Bernabei R., et al., 1998, Physics Letters B, 424, 195

Bernabei R., et al., 2003, Nuovo Cimento Rivista Serie, 26, 1

Bernabei R., et al., 2008, European Physical Journal C, 56, 333

Bertone G., 2010, Nature, 468, 389

Bertone G., Hooper D., Silk J., 2005, Phys. Rep., 405, 279

Blennow M., Fernandez Martinez E., Mena O., Redondo J., Serra e. P., 2012, J. Cosmology Astropart. Phys., 7, 022

Blinnikov S. I., Khlopov M. Y., 1983, Soviet Ast., 27, 371

Blumenthal G. R., Faber S. M., Flores R., Primack J. R., 1986, ApJ, 301, 27

Boddy K. K., Feng J. L., Kaplinghat M., Tait T. M. P., 2014, Phys. Rev. D, 89, 115017

Boddy K. K., Kaplinghat M., Kwa A., Peter A. H. G., 2016, Phys. Rev. D, 94, 123017

Bœhm C., Schewtschenko J. A., Wilkinson R. J., Baugh C. M., Pascoli S., 2014, MNRAS, 445, L31

Boesgaard A. M., Steigman G., 1985, ARA&A, 23, 319

Bond J. R., Cole S., Efstathiou G., Kaiser N., 1991, ApJ, 379, 440

Bose S., et al., 2017, MNRAS, 464, 4520

Bower R. G., 1991, MNRAS, 248, 332

Boylan-Kolchin M., Bullock J. S., Kaplinghat M., 2011, MNRAS, 415, L40

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

Brandenberger R. H., 1991, Physica Scripta Volume T, 36, 114

Bringmann T., Huang X., Ibarra A., Vogl S., Weniger C., 2012, J. Cosmology Astropart. Phys., 7, 054

Bringmann T., Ihle H. T., Kersten J., Walia P., 2016, Phys. Rev. D, 94, 103529

Brownstein J. R., Moffat J. W., 2006, ApJ, 636, 721

Bryan G. L., et al., 2014, ApJS, 211, 19

Buchert T., Räsänen S., 2012, Annual Review of Nuclear and Particle Science, 62, 57

Buckley M. R., Fox P. J., 2010, Phys. Rev. D, 81, 083522

Buckley M. R., Zavala J., Cyr-Racine F.-Y., Sigurdson K., Vogelsberger M., 2014, Phys. Rev. D, 90, 043524

Bullock J. S., Kravtsov A. V., Weinberg D. H., 2000, ApJ, 539, 517

Burkert A., 2000, ApJ, 534, L143

Carlson E. D., Machacek M. E., Hall L. J., 1992, ApJ, 398, 43

Carroll S. M., 2001, Living Reviews in Relativity, 4, 1

Ceverino D., Klypin A., 2009, ApJ, 695, 292

Ceverino D., Klypin A., Klimek E. S., Trujillo-Gomez S., Churchill C. W., Primack J., Dekel A., 2014, MNRAS, 442, 1545

Cirelli M., Iocco F., Panci P., 2009, J. Cosmology Astropart. Phys., 10, 009

Cirelli M., Panci P., Petraki K., Sala F., Taoso M., 2017, J. Cosmology Astropart. Phys., 5, 036

Cline J. M., Liu Z., Xue W., 2012, Phys. Rev. D, 85, 101302

Cline J. M., Liu Z., Moore G. D., Xue W., 2014a, Phys. Rev. D, 89, 043514

Cline J. M., Liu Z., Moore G. D., Xue W., 2014b, Phys. Rev. D, 90, 015023

Clowe D., Bradač M., Gonzalez A. H., Markevitch M., Randall S. W., Jones C., Zaritsky D., 2006, ApJ, 648, L109

Cole S., et al., 2005, MNRAS, 362, 505

Colín P., Avila-Reese V., Valenzuela O., Firmani C., 2002, ApJ, 581, 777

Correa C. A., Wyithe J. S. B., Schaye J., Duffy A. R., 2015, MNRAS, 452, 1217

Creasey P., Sameie O., Sales L. V., Yu H.-B., Vogelsberger M., Zavala J., 2017, MNRAS, 468, 2283

Crocce M., Pueblas S., Scoccimarro R., 2006, MNRAS, 373, 369

Cyr-Racine F.-Y., Sigurdson K., 2013, Phys. Rev. D, 87, 103515

Cyr-Racine F.-Y., Sigurdson K., Zavala J., Bringmann T., Vogelsberger M., Pfrommer C., 2016, Phys. Rev. D, 93, 123527

D'Onghia E., Burkert A., 2003, ApJ, 586, 12

Dahle H., 2006, ApJ, 653, 954

Dalcanton J. J., Stilp A. M., 2010, ApJ, 721, 547

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., 2001, MNRAS, 324, 273

Dehnen W., Read J. I., 2011, European Physical Journal Plus, 126, 55

Di Cintio A., Tremmel M., Governato F., Pontzen A., Zavala J., Bastidas Fry A., Brooks A., Vogelsberger M., 2017, MNRAS, 469, 2845

Diemand J., Kuhlen M., Madau P., Zemp M., Moore B., Potter D., Stadel J., 2008, Nature, 454, 735

Diemer B., Kravtsov A. V., 2015, ApJ, 799, 108

Dolag K., Bartelmann M., Perrotta F., Baccigalupi C., Moscardini L., Meneghetti M., Tormen G., 2004, A&A, 416, 853

Drlica-Wagner A., et al., 2015, ApJ, 813, 109

Duffy L. D., van Bibber K., 2009, New Journal of Physics, 11, 105008

Duffy A. R., Schaye J., Kay S. T., Dalla Vecchia C., 2008, MNRAS, 390, L64

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

Dutton A. A., Macciò A. V., Frings J., Wang L., Stinson G. S., Penzo C., Kang X., 2016, MNRAS, 457, L74

Efstathiou G., 1992, MNRAS, 256, 43P

Eisenstein D. J., Hu W., 1998, ApJ, 496, 605

Eisenstein D. J., et al., 2005, ApJ, 633, 560

Elbert O. D., Bullock J. S., Garrison-Kimmel S., Rocha M., Oñorbe J., Peter A. H. G., 2015, MNRAS, 453, 29

Elbert O. D., Bullock J. S., Kaplinghat M., Garrison-Kimmel S., Graus A. S., Rocha M., 2016, preprint, (arXiv:1609.08626)

Fan J., Katz A., Randall L., Reece M., 2013, Physics of the Dark Universe, 2, 139

Fan J., Katz A., Shelton J., 2014, J. Cosmology Astropart. Phys., 6, 059

Farrar G. R., Rosen R. A., 2007, Physical Review Letters, 98, 171302

Fattahi A., Navarro J. F., Sawala T., Frenk C. S., Sales L. V., Oman K., Schaller M., Wang J., 2016, preprint, (arXiv:1607.06479)

Feng J. L., Kaplinghat M., Tu H., Yu H.-B., 2009, J. Cosmology Astropart. Phys., 7, 004

Feng J. L., Kaplinghat M., Yu H.-B., 2010, Physical Review Letters, 104, 151301

Finkbeiner D. P., Goodenough L., Slatyer T. R., Vogelsberger M., Weiner N., 2011, J. Cosmology Astropart. Phys., 5, 002

Foot R., 2004, International Journal of Modern Physics D, 13, 2161

Foot R., Vagnozzi S., 2015, Phys. Rev. D, 91, 023512

Foot R., Vagnozzi S., 2016, J. Cosmology Astropart. Phys., 7, 013

Foreman-Mackey D., 2016, JOSS, 1

Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306

Fornasa M., et al., 2016, Phys. Rev. D, 94, 123005

Fry A. B., et al., 2015, MNRAS, 452, 1468

Fryxell B., et al., 2000, ApJS, 131, 273

Gao L., Navarro J. F., Cole S., Frenk C. S., White S. D. M., Springel V., Jenkins A., Neto A. F., 2008, MNRAS, 387, 536

Garrison-Kimmel S., Rocha M., Boylan-Kolchin M., Bullock J. S., Lally J., 2013, MNRAS, 433, 3539

Garrison-Kimmel S., Boylan-Kolchin M., Bullock J. S., Kirby E. N., 2014, MNRAS, 444, 222

Gentile G., Famaey B., de Blok W. J. G., 2011, A&A, 527, A76

George M. R., et al., 2012, ApJ, 757, 2

Gingold R. A., Monaghan J. J., 1977, MNRAS, 181, 375

Gnedin N. Y., Kravtsov A. V., 2006, ApJ, 645, 1054

Gnedin O. Y., Kravtsov A. V., Klypin A. A., Nagai D., 2004, ApJ, 616, 16

Goodman J., Hut P., 1989, Nature, 339, 40

Goodman J., Weare J., 2010, Communications in Applied Mathematics and Computational Science, 5, 65

Governato F., et al., 2010, Nature, 463, 203

Governato F., et al., 2012, MNRAS, 422, 1231

Graham P. W., Irastorza I. G., Lamoreaux S. K., Lindner A., van Bibber K. A., 2015, Annual Review of Nuclear and Particle Science, 65, 485

Guth A. H., 1981, Phys. Rev. D, 23, 347

Han H., Zheng S., 2017, Nuclear Physics B, 914, 248

Hargis J. R., Willman B., Peter A. H. G., 2014, ApJ, 795, L13

Harvey D., Massey R., Kitching T., Taylor A., Tittley E., 2015, Science, 347, 1462

Hayashi E., White S. D. M., 2006, MNRAS, 370, L38

Hayashi E., et al., 2004, MNRAS, 355, 794

Hernquist L., 1990, ApJ, 356, 359

Hernquist L., Barnes J. E., 1990, ApJ, 349, 562

Hockney R. W., Eastwood J. W., 1981, Computer Simulation Using Particles

Hoekstra H., Jain B., 2008, Annual Review of Nuclear and Particle Science, 58, 99

Hofmann S., Schwarz D. J., Stöcker H., 2001, Phys. Rev. D, 64, 083507

Hooper D., Goodenough L., 2011, Physics Letters B, 697, 412

Hopkins P. F., Kereš D., Oñorbe J., Faucher-Giguère C.-A., Quataert E., Murray N., Bullock J. S., 2014, MNRAS, 445, 581

Hu W., Barkana R., Gruzinov A., 2000, Physical Review Letters, 85, 1158

Hubble E., 1929, Proceedings of the National Academy of Science, 15, 168

Hunt J. A. S., Bovy J., Carlberg R. G., 2016, ApJ, 832, L25

Ibe M., Yu H.-B., 2010, Physics Letters B, 692, 70

Jelley N. A., 1990, Fundamentals of Nuclear Physics

Jenkins A., 2013, MNRAS, 434, 2094

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

Jesseit R., Naab T., Burkert A., 2002, ApJ, 571, L89

Jungman G., Kamionkowski M., Griest K., 1996, Phys. Rep., 267, 195

Kahlhoefer F., Schmidt-Hoberg K., Frandsen M. T., Sarkar S., 2014, MNRAS, 437, 2865

Kahlhoefer F., Schmidt-Hoberg K., Kummer J., Sarkar S., 2015, MNRAS, 452, L54

Kainulainen K., Tuominen K., Vaskonen V., 2016, Phys. Rev. D, 93, 015016

Kaplinghat M., Tulin S., Yu H.-B., 2014a, Phys. Rev. D, 89, 035009

Kaplinghat M., Keeley R. E., Linden T., Yu H.-B., 2014b, Physical Review Letters, 113, 021302

Kaplinghat M., Tulin S., Yu H.-B., 2016, Physical Review Letters, 116, 041302

Kassiola A., Kovner I., 1993, ApJ, 417, 450

Kauffmann G., White S. D. M., 1993, MNRAS, 261

Kauffmann G., White S. D. M., Guiderdoni B., 1993, MNRAS, 264, 201

Kennedy R., Frenk C., Cole S., Benson A., 2014, MNRAS, 442, 2487

Khrapak S. A., Ivlev A. V., Morfill G. E., Zhdanov S. K., Thomas H. M., 2004, IEEE Transactions on Plasma Science, 32, 555

Kibble T. W. B., 1976, Journal of Physics A Mathematical General, 9, 1387

Kim S. Y., Peter A. H. G., Wittman D., 2017, MNRAS, 469, 1414

King L. J., et al., 2016, MNRAS, 459, 517

Klypin A., Kravtsov A. V., Valenzuela O., Prada F., 1999, ApJ, 522, 82

Klypin A., Karachentsev I., Makarov D., Nasonova O., 2015, MNRAS, 454, 1798

Kochanek C. S., White M., 2000, ApJ, 543, 514

Koda J., Shapiro P. R., 2011, MNRAS, 415, 1125

Kolb E. W., Tkachev I. I., 1996, ApJ, 460, L25

Kouvaris C., Shoemaker I. M., Tuominen K., 2015, Phys. Rev. D, 91, 043519

Kravtsov A. V., 1999, PhD thesis, NEW MEXICO STATE UNIVERSITY

Kusenko A., Steinhardt P. J., 2001, Physical Review Letters, 87, 141301

Kuzio de Naray R., McGaugh S. S., de Blok W. J. G., 2008, ApJ, 676, 920

Lacey C., Cole S., 1993, MNRAS, 262, 627

Lage C., Farrar G., 2014, ApJ, 787, 144

Leauthaud A., et al., 2007, ApJS, 172, 219

Lee S. K., Lisanti M., Safdi B. R., Slatyer T. R., Xue W., 2016, Physical Review Letters, 116, 051103

Loeb A., Weiner N., 2011, Physical Review Letters, 106, 171302

Łokas E. L., 2002, MNRAS, 333, 697

Łokas E. L., Mamon G. A., 2001, MNRAS, 321, 155

Lovell M. R., Frenk C. S., Eke V. R., Jenkins A., Gao L., Theuns T., 2014, MNRAS, 439, 300

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

Lynden-Bell D., Eggleton P. P., 1980, MNRAS, 191, 483

Macciò A. V., Fontanot F., 2010, MNRAS, 404, L16

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

Markevitch M., 2006, in Wilson A., ed., ESA Special Publication Vol. 604, The X-ray Universe 2005. p. 723

Markevitch M., Gonzalez A. H., Clowe D., Vikhlinin A., Forman W., Jones C., Murray S., Tucker W., 2004, ApJ, 606, 819

Mashchenko S., Couchman H. M. P., Wadsley J., 2006, Nature, 442, 539

Massey R., Kitching T., Richard J., 2010, Reports on Progress in Physics, 73, 086901

Massey R., et al., 2015, MNRAS, 449, 3393

Mastropietro C., Burkert A., 2008, MNRAS, 389, 967

McConnachie A. W., 2012, AJ, 144, 4

McGaugh S. S., de Blok W. J. G., 1998, ApJ, 499, 66

McGaugh S. S., Lelli F., Schombert J. M., 2016, Physical Review Letters, 117, 201101

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

Milgrom M., 1983, ApJ, 270, 371

Milosavljević M., Koda J., Nagai D., Nakar E., Shapiro P. R., 2007, ApJ, 661, L131

Miralda-Escudé J., 2002, ApJ, 564, 60

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

Monaghan J. J., Lattanzio J. C., 1985, A&A, 149, 135

Moore B., Ghigna S., Governato F., Lake G., Quinn T., Stadel J., Tozzi P., 1999, ApJ, 524, L19

Moore B., Gelato S., Jenkins A., Pearce F. R., Quilis V., 2000, ApJ, 535, L21

Moster B. P., Naab T., White S. D. M., 2013, MNRAS, 428, 3121

Navarro J. F., Eke V. R., Frenk C. S., 1996, MNRAS, 283, L72

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

Newman A. B., Treu T., Ellis R. S., Sand D. J., Richard J., Marshall P. J., Capak P., Miyazaki S., 2009, ApJ, 706, 1078

Newman A. B., Treu T., Ellis R. S., Sand D. J., 2013, ApJ, 765, 25

Nusser A., Pointecouteau E., 2006, MNRAS, 366, 969

Oñorbe J., Boylan-Kolchin M., Bullock J. S., Hopkins P. F., Kereš D., Faucher-Giguère C.-A., Quataert E., Murray N., 2015, MNRAS, 454, 2092

Oh S.-H., de Blok W. J. G., Brinks E., Walter F., Kennicutt Jr. R. C., 2011, AJ, 141, 193

Okabe N., Smith G. P., Umetsu K., Takada M., Futamase T., 2013, ApJ, 769, L35

Okada N., Nefer Şenoğuz V., Shafi Q., 2014, preprint, (arXiv:1403.6403)

Oman K. A., et al., 2015, MNRAS, 452, 3650

Oman K. A., Navarro J. F., Sales L. V., Fattahi A., Frenk C. S., Sawala T., Schaller M., White S. D. M., 2016, MNRAS, 460, 3610

Pace A., Andrade K., Kaplinghat M., Tulin S., Yu H.-b., 2016, in American Astronomical Society Meeting Abstracts. p. 337.05

Papastergis E., Giovanelli R., Haynes M. P., Shankar F., 2015, A&A, 574, A113

Peñarrubia J., Pontzen A., Walker M. G., Koposov S. E., 2012, ApJ, 759, L42

Peccei R. D., Quinn H. R., 1977, Physical Review Letters, 38, 1440

Penzias A. A., Wilson R. W., 1965, ApJ, 142, 419

Perlmutter S., et al., 1999, ApJ, 517, 565

Peter A. H. G., Rocha M., Bullock J. S., Kaplinghat M., 2013, MNRAS, 430, 105

Petraki K., Volkas R. R., 2013, International Journal of Modern Physics A, 28, 1330028

Petraki K., Pearce L., Kusenko A., 2014, J. Cosmology Astropart. Phys., 7, 039

Pineda J. C. B., Hayward C. C., Springel V., Mendes de Oliveira C., 2017, MNRAS, 466, 63

Planck Collaboration et al., 2014, A&A, 571, A16

Planck Collaboration et al., 2016, A&A, 594, A13

Plummer H. C., 1911, MNRAS, 71, 460

Pointecouteau E., Silk J., 2005, MNRAS, 364, 654

Pollack J., Spergel D. N., Steinhardt P. J., 2015, ApJ, 804, 131

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

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

Quinn T., Katz N., Stadel J., Lake G., 1997, ArXiv Astrophysics e-prints,

Ragozzine B., Clowe D., Markevitch M., Gonzalez A. H., Bradač M., 2012, ApJ, 744, 94

Rajaraman A., Shepherd W., Tait T. M. P., Wijangco A. M., 2011, Phys. Rev. D, 84, 095013

Randall L., 2015, Dark Matter and the Dinosaurs: The Astounding Interconnectedness of the Universe

Randall S. W., Markevitch M., Clowe D., Gonzalez A. H., Bradač M., 2008, ApJ, 679, 1173

Read J. I., Iorio G., Agertz O., Fraternali F., 2016, MNRAS, 462, 3628

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

Richard J., Kneib J.-P., Limousin M., Edge A., Jullo E., 2010, MNRAS, 402, L44

Riess A. G., et al., 1998, AJ, 116, 1009

Rocha M., Peter A. H. G., Bullock J. S., Kaplinghat M., Garrison-Kimmel S., Oñorbe J., Moustakas L. A., 2013, MNRAS, 430, 81

Rodríguez-Torres S. A., et al., 2016, MNRAS, 460, 1173

Rubin V. C., Ford Jr. W. K., Thonnard N., 1980, ApJ, 238, 471

Rutherford E., 1911, The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, 21, 669

Sand D. J., Treu T., Smith G. P., Ellis R. S., 2004, ApJ, 604, 88

Sand D. J., Treu T., Ellis R. S., Smith G. P., Kneib J.-P., 2008, ApJ, 674, 711

Sawala T., et al., 2016, MNRAS, 457, 1931

Schaller M., et al., 2015a, MNRAS, 451, 1247

Schaller M., et al., 2015b, MNRAS, 452, 343

Schaller M., Robertson A., Massey R., Bower R. G., Eke V. R., 2015c, MNRAS, 453, L58

Schaye J., et al., 2015, MNRAS, 446, 521

Schutz K., Slatyer T. R., 2015, J. Cosmology Astropart. Phys., 1, 021

Scott D., White M., Cohn J. D., Pierpaoli E., 2001, ArXiv Astrophysics e-prints,

Seljak U., Zaldarriaga M., 1999, CMBFAST: A microwave anisotropy code, Astrophysics Source Code Library (ascl:9909.004)

Sheth R. K., Tormen G., 1999, MNRAS, 308, 119

Sheth R. K., Mo H. J., Tormen G., 2001, MNRAS, 323, 1

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

Smoot G. F., et al., 1992, ApJ, 396, L1

Somerville R. S., 2002, ApJ, 572, L23

Soni A., Zhang Y., 2016, Phys. Rev. D, 93, 115025

Soussa M. E., Woodard R. P., 2004, Physics Letters B, 578, 253

Spergel D. N., Steinhardt P. J., 2000, Physical Review Letters, 84, 3760

Spergel D. N., et al., 2003, ApJS, 148, 175

Springel V., 2005, MNRAS, 364, 1105

Springel V., Farrar G. R., 2007, MNRAS, 380, 911

Springel V., Di Matteo T., Hernquist L., 2005a, MNRAS, 361, 776

Springel V., et al., 2005b, Nature, 435, 629

Springel V., Frenk C. S., White S. D. M., 2006, Nature, 440, 1137

Springel V., et al., 2008, MNRAS, 391, 1685

Steigman G., Dasgupta B., Beacom J. F., 2012, Phys. Rev. D, 86, 023506

Swaters R. A., Verheijen M. A. W., Bershady M. A., Andersen D. R., 2003, ApJ, 587, L19

Takahashi R., Chiba T., 2007, ApJ, 671, 45

Tegmark M., 2005, J. Cosmology Astropart. Phys., 4, 001

Teyssier R., 2002, A&A, 385, 337

Teyssier R., Pontzen A., Dubois Y., Read J. I., 2013, MNRAS, 429, 3068

Trujillo-Gomez S., Klypin A., Colín P., Ceverino D., Arraki K. S., Primack J., 2015, MNRAS, 446, 1140

Tulin S., Yu H.-B., Zurek K. M., 2013a, Phys. Rev. D, 87, 115007

Tulin S., Yu H.-B., Zurek K. M., 2013b, Physical Review Letters, 110, 111301

Vachaspati T., 1986, Physical Review Letters, 57, 1655

Viel M., Lesgourgues J., Haehnelt M. G., Matarrese S., Riotto A., 2005, Phys. Rev. D, 71, 063534

Vogelsberger M., Zavala J., 2013, MNRAS, 430, 1722

Vogelsberger M., et al., 2009, MNRAS, 395, 797

Vogelsberger M., Zavala J., Loeb A., 2012, MNRAS, 423, 3740

Vogelsberger M., et al., 2014a, MNRAS, 444, 1518

Vogelsberger M., Zavala J., Simpson C., Jenkins A., 2014b, MNRAS, 444, 3684

Vogelsberger M., Zavala J., Cyr-Racine F.-Y., Pfrommer C., Bringmann T., Sigurdson K., 2016, MNRAS, 460, 1399

Walker M. G., Mateo M., Olszewski E. W., Peñarrubia J., Wyn Evans N., Gilmore G., 2009, ApJ, 704, 1274

Wang W., Zhang M., Zhao J., 2016, preprint, (arXiv:1604.00123)

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

Weinberg D. H., Bullock J. S., Governato F., Kuzio de Naray R., Peter A. H. G., 2015, Proceedings of the National Academy of Science, 112, 12249

White S. D. M., Frenk C. S., Davis M., 1983, ApJ, 274, L1

Williams L. L. R., Saha P., 2011, MNRAS, 415, 448

Willman B., 2010, Advances in Astronomy, 2010, 285454

Wolf J., Martinez G. D., Bullock J. S., Kaplinghat M., Geha M., Muñoz R. R., Simon J. D., Avedo F. F., 2010, MNRAS, 406, 1220

Wu K. K. S., Lahav O., Rees M. J., 1999, Nature, 397, 225

Yoshida N., Springel V., White S. D. M., Tormen G., 2000a, ApJ, 535, L103

Yoshida N., Springel V., White S. D. M., Tormen G., 2000b, ApJ, 544, L87

Zavala J., Vogelsberger M., Walker M. G., 2013, MNRAS, 431, L20

Zhao D. H., Mo H. J., Jing Y. P., Börner G., 2003, MNRAS, 339, 12

Zolotov A., et al., 2012, ApJ, 761, 71

Zwicky F., 1937, ApJ, 86, 217

de Blok W. J. G., Bosma A., McGaugh S., 2003, MNRAS, 340, 657

van den Aarssen L. G., Bringmann T., Pfrommer C., 2012, Physical Review Letters, 109, 231301