

DYNAMICS IN THE CORES OF SELF-INTERACTING DARK MATTER HALOS: REDUCED STALLING AND ACCELERATED CORE COLLAPSE

FRANK C. VAN DEN BOSCH[✉] AND SHASHANK DATTATHRI[✉]

Department of Astronomy, Yale University, PO. Box 208101, New Haven, CT 06520-8101

Version November 20, 2025

ABSTRACT

Self-interacting dark matter (SIDM) is an intriguing alternative to the standard cold dark matter (CDM) paradigm, which predicts that dark matter halos typically have large, isothermal cores. Numerical simulations have shown that dynamical friction ceases to operate in cores of (roughly) constant density, a phenomenon known as core stalling. In addition, such cores often are unstable to a dipole instability that gives rise to dynamical buoyancy, causing massive central objects to move outward. Thus far, these manifestations of core dynamics have only been demonstrated in collisionless systems. In this paper, we use idealized N -body simulations to study whether core stalling and dynamical buoyancy operate in SIDM halos. We find that if the self-interactions are sufficiently strong, neither core stalling nor buoyancy are present, and a massive perturber will quickly sink all the way to the center of its host. In doing so, it gravitationally contracts the core, catalyzing a strongly accelerated core collapse. The reason why core dynamics are so different in SIDM halos is that self-interactions drive the halo’s distribution function to a featureless exponential, removing any inflections or plateaus that are responsible for the dipole instability and core stalling in the case of CDM. We discuss implications of our finding for constraining the nature of dark matter by using observations of massive objects such as supermassive black holes (SMBHs), globular clusters, and nuclear star clusters in the central regions of galaxies.

Keywords: Dark Matter – Dynamical Friction – Gravitational Instability – Galaxy Dynamics – Galaxy Nuclei – Galaxy Dark Matter Halos – N -body simulations

1. INTRODUCTION

Cores, which are regions of roughly uniform density, seem to be ubiquitous in the central regions of a wide variety of systems, including the dark matter (DM) halos of dwarf galaxies (Flores & Primack 1994; Oman et al. 2015; Agnello & Evans 2012; Oh et al. 2015), low-surface-brightness galaxies (de Blok et al. 2001; de Blok & Bosma 2002; Kuzio de Naray et al. 2008), and massive star-forming disks at redshifts $z \sim 1 - 3$ (Genzel et al. 2020; Price et al. 2021).

The presence of DM cores has often been used as an argument against Cold Dark Matter (CDM), which predicts cuspy halos (e.g., Navarro et al. 1997), and in favor of alternative dark matter models that predict halos to be cored, such as Self-Interacting Dark Matter (SIDM, e.g., Spergel & Steinhardt 2000; Kochanek & White 2000; Rocha et al. 2013; Tulin & Yu 2018). However, over the years, it has become clear that there are a variety of astrophysical processes that can transform a steep central cusp into a constant density core and/or remove dark matter from the central regions. Chief among these is explosive gas outflows driven by feedback from supernova or active galactic nuclei (AGN), which cause dark matter and stars to expand in response to the fast shallowing of the central potential well. Supernova-driven outflows, especially if they are episodic (Pontzen & Governato 2012; Teyssier et al. 2013; El-Zant et al. 2016; Li et al. 2023), are expected to be effective in low-mass galaxies (Dekel & Silk 1986; Di Cintio et al. 2014; Tollet et al. 2016; Freundlich et al. 2020), while AGN-driven outflows may be relevant for creating cores in massive galaxies (Martizzi et al. 2013; Dekel et al. 2021). Alternative methods for core creation are dynamical heating due

to dynamical friction acting on a massive perturber (El-Zant et al. 2001; Goerdt et al. 2010; Cole et al. 2011; Ogiya & Nagai 2022) or a bar (Weinberg & Katz 2002), tidal stripping of systems that are initially radially anisotropic (Chiang et al. 2025), and core scouring due to a binary supermassive black hole (Begelman et al. 1980; Quinlan 1996). The latter process is often invoked to explain the stellar cores observed in massive ellipticals (Ferrarese et al. 1994; Byun et al. 1996; Gebhardt et al. 1996). Hence, cores are a fairly natural by-product of structure formation, and their mere presence gives little to no insight regarding the nature of dark matter.

Numerical simulations of systems with central cores have revealed some intriguing dynamical phenomena. In particular, Read et al. (2006) noticed that a massive perturber that sinks due to dynamical friction toward the center of a system with a large constant density core stalls its inspiral close to the core radius. This ‘core stalling’, which has since been reproduced in numerous independent studies (Goerdt et al. 2006; Inoue 2009; Goerdt et al. 2010; Inoue 2011; Kaur & Sridhar 2018; Dutta Chowdhury et al. 2019), is not predicted by the standard Chandrasekhar (1943) formulation of dynamical friction. Another dynamical phenomenon in cores, arguably even more confounding than core stalling, is ‘dynamical buoyancy’, which was first noticed by Cole et al. (2012). Using N -body simulations to study what happens to a massive object that is initially located inside a constant density core, they found that the object starts moving outwards until it reaches the core radius, where its outward motion stalls. Finally, Dattathri et al. (2025b) showed that cored systems can be unstable to the development of a rotating dipole mode. This mode saturates into a long-lived solitonic mode that can dislodge the entire central cusp, causing it to slosh back and forth through the central

region of the galaxy without ever settling back down.

All these core-specific dynamical processes can have profound implications for various aspects of galaxy evolution. They can conspire to prevent massive objects such as nuclear star clusters and/or SMBHs from reaching, or remaining at, the center of their potential well, and may thus explain why in dwarf galaxies, largely believed to be cored, many nuclear star clusters and AGN are found to be offset from their center of mass (Binggeli et al. 2000; Côté et al. 2006; Shen et al. 2019; Reines et al. 2020; Sargent et al. 2022; Mezcua & Domínguez Sánchez 2024). They can also prevent SMBHs from forming a hard binary, thereby inhibiting the SMBHs from merging together. Interestingly, mergers among SMBHs in dwarf galaxies are expected to dominate the gravitational wave signal to be measured with the Laser Interferometer Space Antenna (LISA) (Barausse et al. 2020; Volonteri et al. 2020). However, these estimates have not taken into account core stalling and might therefore be severely overestimated. Core stalling has also been invoked to explain the presence of multiple nuclei or globular clusters in the centers of galaxies, which would otherwise coalesce in a short time (Goerdt et al. 2006; Bonfini & Graham 2016; Boldrini et al. 2019; Sánchez-Salcedo & Lora 2022).

In a recent paper (Dattathri et al. 2025a, hereafter D25), we presented a unified picture for core stalling, buoyancy, and the dipole instability operating in the cores of isotropic and spherically symmetric systems. Using idealized simulations and arguments based on kinetic theory, we showed that all these dynamical phenomena are related to specific features in the phase-space distribution function (DF) $f = f(E)$ of the host system. More specifically, core stalling occurs whenever the gradient df/dE at the orbital energy of the perturber vanishes (i.e., when the perturber reaches a plateau in the DF), while the dipole instability is triggered whenever the DF has an inflection, where $df/dE > 0$. This mode exerts a torque on a central cusp or central object, such as a supermassive black hole or a nuclear star cluster, which can set it in motion, causing it to move outward, thus giving rise to dynamical buoyancy.

Bumps and plateaus in DFs may arise through various dynamical processes related to structure formation and evolution. In particular, D25 show that inflections in the DF are common in spherical isotropic systems having a rapid transition from a steep outer density profile to a shallow inner profile (see also Weinberg 2023). Hence, processes that transform a central cusp in a core will affect the DF and might introduce local inflections that trigger the dipole instability. Moreover, dynamical friction acting on a massive perturber transfers energy and angular momentum from the perturber to the host system, thereby modifying its DF. Hence, even if the system initially has $df/dE < 0$ everywhere, the infall of a perturber can create a plateau or inflection.

So far, core stalling, buoyancy and the dipole instability have only been demonstrated in collisionless N -body systems, such as CDM halos, that obey the collisionless Boltzmann equation (CBE). However, cores are a natural prediction of several alternative dark matter models. In particular, if dark matter is self-interacting, the resulting heat conduction gives rise to dark matter halos with large isothermal cores (Spergel & Steinhardt 2000; Burkert 2000). Hence, it seems natural to postulate that massive perturbers sinking inside SIDM halos should stall near the core radius (see for example Adhikari et al. 2022). However, as shown in D25, not all cores are equal, in that the mere presence of a core does not guarantee core stalling and/or buoyancy. In this paper, we use high-

resolution, idealized numerical experiments to examine how self-interactions impact core dynamics. Interestingly, we show that self-interactions suppress core stalling, buoyancy, and the dipole instability. If sufficiently strong, the self-interactions can completely suppress these phenomena such that, contrary to the case of CDM, massive objects rapidly sink to the center of the cored host system where they remain at rest. We also demonstrate that as the massive objects sink towards the center of a SIDM halo, they catalyze a drastically accelerated core collapse. These results offer a novel pathway to probing the nature of dark matter.

This paper is organized as follows. In Section 2 we first give a brief overview of the dynamics underlying core stalling, buoyancy, and the dipole instability. Section 3 presents the initial conditions for the different dark matter halos used throughout this study, the N -body code that we use, and its validation. Section 4 studies core stalling in two different halos, comparing the results for CDM to those of SIDM. Dynamical buoyancy and the dipole instability are the topics of Section 5, again comparing how these processes manifest in CDM vs. SIDM. Section 6 presents a discussion of why core dynamics in SIDM differ drastically from that in CDM. We end in Section 7 with a summary and discussion of the implications of our findings.

2. KINETIC THEORY OF CORE DYNAMICS

Before assessing how core stalling, buoyancy, and the dipole instability are impacted by self-interactions, we first briefly review the unified kinetic picture for core dynamics advanced in D25. The key principle underlying all these core-specific dynamical processes is the secular transport of energy and angular momentum between the host particles and a perturbation. This perturbation might be a massive black hole that is sinking in due to dynamical friction, or it can be a point mode (also known as Landau mode) that is inherent to the system. The dipole instability is an example of the latter case¹. According to kinetic theory (see Binney & Tremaine 2008; Fouvry 2017; Hamilton & Fouvry 2024, for some excellent reviews), this secular transport occurs at resonances between the frequencies of the host particles and that of the perturber (i.e., the circular frequency of the inspiraling black hole, or the eigen-frequency of the dipole mode). In particular, while particles on one side of the perturber gain energy from interacting with the perturber, those on the other side lose energy by transferring it to the perturber. Hence, whether the perturber or perturbation gains or loses energy depends on the net balance between these gainers and losers, which in turn is reflected by the gradient in the distribution function across the resonance.

2.1. Dynamical Friction and the LBK Torque

Dynamical friction arises from a response induced in the DF of the host halo by the motion of a massive perturber. This response exerts a torque on the perturber that causes a transfer of angular momentum (and energy) between the perturber and the particles that make up the host halo. Applying linear perturbation theory to the CBE-Poisson system that describes a collisionless gravitational system, Tremaine & Weinberg (1984) showed that (in the time-asymptotic limit, $t \rightarrow \infty$),

¹ As shown in Weinberg (2023) and Dattathri et al. (2025b), the dipole mode is a discrete point mode, also known as a Landau mode. It is a zero of the system's dispersion relation inferred from perturbing the coupled CBE-Poisson equation.

this torque is given by

$$\mathcal{T}_{\text{LBK}} = 16\pi^4 \sum_{\ell} \ell_3 \int d\mathbf{J} \delta(\ell \cdot \boldsymbol{\Omega} - \ell_3 \Omega_p) \ell \cdot \frac{\partial f}{\partial \mathbf{J}} |\hat{\Phi}_{\ell}(\mathbf{J})|^2. \quad (1)$$

which is known as the LBK torque after [Lynden-Bell & Kalnajs \(1972\)](#) who first derived it in their treatment of angular momentum transport due to spiral structure in disk galaxies. Here, $\mathbf{J} = (J_1, J_2, J_3)$ are the three actions, $\boldsymbol{\Omega} = (\Omega_1, \Omega_2, \Omega_3)$ are the corresponding frequencies, $\ell = (\ell_1, \ell_2, \ell_3)$ is an integer vector, Ω_p is the frequency of the perturber, $|\hat{\Phi}_{\ell}(\mathbf{J})|$ are the Fourier modes of the perturber potential, $\delta(x)$ is the Dirac delta function, and $f = f(\mathbf{J})$ is the unperturbed DF of the halo. The summation runs from $-\infty$ to ∞ for ℓ_1 and ℓ_2 , and from 0 to ∞ for ℓ_3 . In the case of a spherical isotropic system, the DF is a one-dimensional function of energy, so the LBK torque reduces to

$$\mathcal{T}_{\text{LBK}} = 16\pi^4 \Omega_p \sum_{\ell} \ell_3^2 \int d\mathbf{J} \delta(\ell \cdot \boldsymbol{\Omega} - \ell_3 \Omega_p) \frac{df}{dE} |\hat{\Phi}_{\ell}(\mathbf{J})|^2. \quad (2)$$

This expression clearly exposes that the torque on the perturber arises from orbits that are on resonance with the perturber (i.e., that satisfy the commensurability condition $\ell \cdot \boldsymbol{\Omega} - \ell_3 \Omega_p = 0$) and that it is the gradient of the DF in action space at those resonances that determines the sign of the torque and thus the net direction of flux transfer.

Spherical isotropic systems that have $df/dE < 0$ at all energies are stable to both radial and non-radial modes ([Antonov 1962](#); [Doremus et al. 1973](#)). As is apparent from equation (2), the LBK torque in such systems is always negative (that is, retarding) and the perturber thus experiences regular dynamical friction. What makes kinetic theory of gravitational systems complicated is that there are typically an infinite number of resonances (represented by the summations in equations [1]–[2]), and at each resonance there are typically three gradients of the DF with respect to three actions. However, [Kaur & Sridhar \(2018\)](#) have shown that within the core region of a spherical isotropic system, the LBK torque is dominated by the contribution from the corotation resonance (CR), where the azimuthal frequency of the particles matches that of the perturber (see also [Kaur & Stone 2022](#); [Banik & van den Bosch 2022](#)). Therefore, the sign of the net torque is determined, to a good approximation, by the gradient df/dE at the energy associated with co-rotation, which we refer to as $(\nabla f)_{\text{CR}}$ in what follows. If $(\nabla f)_{\text{CR}} < 0$, then the LBK torque is retarding and the perturber experiences regular dynamical friction. However, if $(\nabla f)_{\text{CR}} > 0$, the torque is enhancing, causing the perturber to gain energy and move outward, which is a manifestation of dynamical buoyancy. Finally, if $(\nabla f)_{\text{CR}} = 0$ the torque vanishes and the perturber is expected to undergo core stalling. Hence, the expression for the LBK torque exemplifies that core stalling and buoyancy have a natural explanation in kinetic theory and should be linked to the detailed shape of the DF of the host system. Using a series of N -body experiments, [D25](#) established that dynamical friction, core stalling and buoyancy can indeed be linked to the sign of $(\nabla f)_{\text{CR}}$, at least for a massive perturber in a spherical background halo with an (initially) isotropic DF.

Although this simple picture based on the LBK torque offers a natural explanation for core stalling that is consistent with the results of numerical simulations, it only captures part of the full dynamics. Here, we summarize some of the compli-

cations and refer the reader to [D25](#) for a more detailed discussion. Firstly, the LBK torque is based on linear perturbation theory and therefore does not account for non-linear effects such as orbit trapping. Therefore, it is only valid in what [Tremaine & Weinberg \(1984\)](#) refer to as the “fast regime”, when the perturber is sinking fast enough such that it sweeps through the resonances without any of them building up to non-linear amplitudes. Hence, the LBK torque is only of limited use in describing core-stalling, which transpires in the “slow regime”, in which particles can get trapped onto librating orbits that continuously and periodically exchange energy and angular momentum with the perturber ([Tremaine & Weinberg 1984](#); [Chiba & Schönrich 2021](#); [Chiba 2023](#); [Banik & van den Bosch 2022](#)). Over time, phase-mixing among these librating orbits will rapidly drive their net action flux to zero. Hence, if the libration zone is evolving sufficiently slowly over time, the librating orbits contribute a negligible torque and the perturber will stall. Secondly, the LBK torque is calculated based on the *unperturbed* DF and does not account for the evolution in the DF due to secular flux transfer as the perturber sinks in. Quasi-linear theory (e.g., [Weinberg 2001](#); [Chavanis 2024](#)) dictates that secular flux transfer causes a diffusion in phase space that, given enough time, generates a plateau in the DF ($df/dE = 0$) at the location of each resonance ([Hamilton 2024](#); [Banik & Bhattacharjee 2025](#)). In cored systems, the co-rotation resonance dominates; therefore a slowly sinking perturber can carve out a strong plateau as it enters the core, suppressing the LBK torque and halting further inspiral. Hence, dynamical friction can be self-limiting in that the sinking in of the perturber can modify the DF to the point that it induces stalling. Indeed, such behavior has been seen in N -body simulations ([Goerdt et al. 2010](#); [Dattathri et al. 2025a](#), see also Section 6.2 below). Finally, the LBK torque is only valid in the asymptotic limit ($t \rightarrow \infty$) of adiabatic growth of the perturber. However, if the system happens to be unstable, then the developing instabilities are likely to become important before the LBK torque has a chance to develop. As discussed in Section 5, this is likely the case with dynamical buoyancy. Despite these details, [D25](#) show that both in the linear and weakly non-linear regimes, the gradient in the DF is indeed the key determinant in the magnitude and direction of frictional force, and we adhere to this framework in this paper.

2.2. Dipole Instability and Dynamical Buoyancy

Gravitational N -body systems have a spectrum of modes at discrete frequencies (referred to as point modes or Landau modes), which describe perturbations to the system that result in a response that has the same functional form as the perturbation. These modes, of which the dipole mode is a typical example, are continuously excited, either by external perturbations or by discreteness noise inherent to the system. Whether they grow or decay depends on the gradients in the DF at the resonances between the eigen-frequency of the mode and the orbital frequencies of the particles of the system in question.

Although the dipole mode is common, it is typically (weakly) damped ([Weinberg 1994b](#); [Heggie et al. 2020](#)). The first to report an example of a system with an unstable (growing) dipole mode was [Weinberg \(2023\)](#). He noticed the instability in a spherical isotropic NFW halo, commonly used to model cold dark matter halos, but only if he truncated the density profile at large radii. This truncation induces an inflection in the DF (i.e., a region in energy where $df/dE > 0$), thus violating Antonov’s stability criterion. Recently, [Dattathri et al. \(2025b\)](#) showed that spherical isotropic systems with a double

power-law density profile that have a rapid transition from a steep outer density gradient to a shallow inner density gradient have a similar inflection in their DFs. Similarly to the system studied by Weinberg (2023), these systems are unstable to the development of a rotating dipole mode, whose growth is facilitated by the positive gradient in the DF.

As the dipole instability grows in amplitude, it ultimately saturates developing into a long-lived solitonic mode, which is the dipolar equivalent of a quadrupolar bar mode in a disk galaxy. Most relevant for this paper, Dattathri et al. (2025b) showed that this mode exerts a torque that can dislodge a central cusp, setting it in motion about the system center. Over time, a long-lived dynamical equilibrium is established in which the cusp continues to slosh back and forth through the central region of the galaxy. In D25, we have shown that the same mechanism can also dislodge a massive perturber from the center. Hence, dynamical buoyancy is a manifestation of the dipole instability, which, in turn, has its origin in an inflection in the DF.

3. METHODOLOGY

The main goal of this paper is to investigate whether core stalling and/or dynamical buoyancy occur in the cores of SIDM halos. Using idealized N -body simulations of isolated dark matter halos in which we introduce a massive perturber, we study how the perturber spirals inwards due to dynamical friction, and how the time scale thereof depends on the collisional cross section of the dark matter particles. Crucially, we contrast the findings with those in standard CDM (i.e., zero cross-section for self-interactions).

3.1. Host halos

In this paper, we consider three different host halos. All are assumed to be spherical and to have an isotropic distribution function (DF) $f = f(E)$, but they have different density profiles. The first is a cored, truncated NFW (Navarro et al. 1997) halo, in which the core is created due to self-interactions. This therefore represents a typical halo one might expect if dark matter is indeed self-interacting. The second halo has a double-power law density profile with an inner core that has a shallow density gradient ($\rho \propto r^{-0.1}$). This halo is chosen because it is known, in the case of CDM, to exhibit strong core stalling at a relatively large radius. Hence, this halo is an ideal benchmark for testing how core stalling depends on the self-interaction cross section. The third and final halo has a similar double-power law density profile as halo two, but with a central $r^{-0.5}$ cusp. We use this halo because Dattathri et al. (2025b) has shown that it is highly unstable to the development of a dipole instability, at least in the case of CDM. Hence, we will use this halo to study how the dipole instability and the associated dynamical buoyancy behave in the presence of self-interactions. In what follows, we refer to these three halos as the cNFW, ABG[0.1] and ABG[0.5] halos, respectively. Here ‘ABG’ refers to the parameters α , β and γ that characterize the double power-law density profile (see below).

The initial conditions for the cNFW system are created as follows. We first initialize a halo with a NFW density profile,

$$\rho(r) = \frac{\rho_s}{(r/r_s)(1 + r/r_s)^2}, \quad (3)$$

with scale radius r_s and characteristic density ρ_s , which we truncate using the energy-truncation method of Drakos et al.

(2017)². This is similar to the method used to create King models by modifying the DF of an isothermal sphere (Michie 1963; King 1966; Binney & Tremaine 2008). Specifically, we use an energy value for the truncation such that the total truncated mass $M_{\text{tot}} = 1.2703 M_{\text{vir}}$. This is chosen such that the density profile inside the virial radius, r_{vir} , is almost identical to that of the untruncated profile. Here M_{vir} is the virial mass of the system, defined as the mass enclosed by the untruncated NFW halo inside r_{vir} , defined as the radius inside which the average density is Δ_{vir} times the critical density for closure. Throughout, we adopt $\Delta_{\text{vir}} = 97$, in accordance with the Planck cosmology (Bryan & Norman 1998; Planck Collaboration et al. 2018), and a halo concentration parameter $c = r_{\text{vir}}/r_s = 20$. The red curves in Fig. 1 show the density and 1D velocity dispersion profiles of this truncated NFW (hereafter tNFW) profile, normalized to ρ_s and $v_s \equiv \sqrt{GM_{\text{vir}}/r_s}$, respectively. For comparison, the gray-dashed lines show the corresponding untruncated profiles. Note that the tNFW system looks very similar to the untruncated NFW profile out to the virial radius, with $M(< R_{\text{vir}}) \sim 0.95 M_{\text{vir}}$.

As described in detail in Section 3.4, the tNFW halo is subsequently evolved in isolation assuming self-interacting dark matter with a cross section per unit mass equal to $\sigma_m = 25 \text{ cm}^2 \text{ g}^{-1}$, until it reaches the epoch of maximum core (i.e., minimal central density, shortly before the onset of core collapse). This cored-out version of the truncated NFW halo is representative of a typical evolved halo if dark matter has non-negligible self-interactions and serves as the initial conditions for our subsequent numerical experiments described below. The orange curves in Fig. 1 show the corresponding density and velocity dispersion profiles of this cored-out NFW profile, to which we refer as cNFW hereafter. Note how the cNFW system is indistinguishable from the tNFW halo for $r \gtrsim r_s$. Inside of the scale-radius, however, the self-interactions have created a pronounced isothermal core of constant density.

The ABG[0.1] and ABG[0.5] halos have a double power-law density profile characterized by the general (α, β, γ) -profile (Zhao 1996),

$$\rho(r) = \rho_s \left(\frac{r}{r_s} \right)^{-\gamma} \left[1 + \left(\frac{r}{r_s} \right)^\alpha \right]^{\frac{\gamma-\beta}{\alpha}}. \quad (4)$$

This profile transitions from an outer power-law, $\rho \propto r^{-\beta}$, to an inner power-law, $\rho \propto r^{-\gamma}$, with α controlling the steepness of the transition. Several commonly encountered density profiles in astrophysics are examples of this (α, β, γ) profile, including the (untruncated) NFW profile, which has $(\alpha, \beta, \gamma) = (1, 3, 1)$, and the Hernquist (1990), Jaffe (1983) and Plummer (1911) profiles, which have $(\alpha, \beta, \gamma) = (1, 4, 1)$, $(1, 4, 2)$ and $(2, 5, 0)$, respectively. For the ABG[0.1] halo we adopt $(\alpha, \beta, \gamma) = (4, 4, 0.1)$, which transitions from an outer power-law with $\rho \propto r^{-4}$ to an inner core-like region with $\rho \propto r^{-0.1}$. The corresponding density and 1D velocity dispersion profile, normalized by ρ_s and $v_s \equiv \sqrt{GM_{\text{tot}}/r_s}$, respectively, are shown as the blue curves in Fig. 1. Note that the core region of this halo is not isothermal, unlike the cNFW halo. As shown in D25 a CDM halo with this density profile causes massive perturbers to experience core stalling at a relatively large radius (see also Section 4.2 below).

² This is required since an untruncated NFW profile has infinite mass, and since a hard-truncation at the virial radius, or any other radius, results in instability (Weinberg 2023).

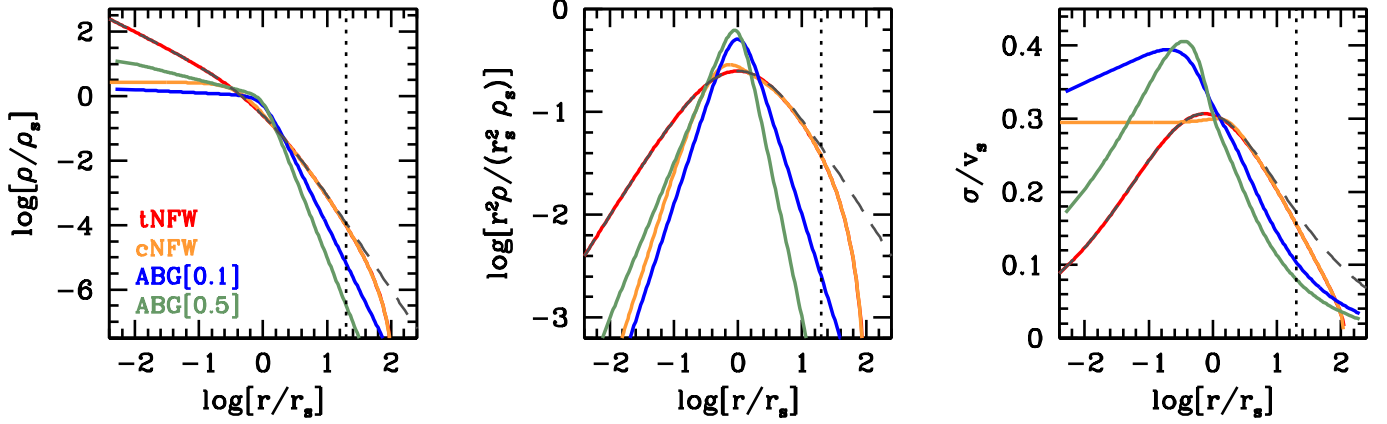


Figure 1. Initial density and velocity dispersion profiles for the various halos discussed in this paper, as indicated. The gray dashed profiles correspond to the untruncated NFW profile and is shown for comparison. Note that the density profile of the truncated NFW is almost indistinguishable from the untruncated NFW profile inside the virial radius, indicated as the vertical, dotted line.

Finally, the density profile of the ABG[0.5] halo also follows the (α, β, γ) -profile of equation (4), but with $(\alpha, \beta, \gamma) = (6, 5, 0.5)$. Hence, its density profile rapidly transitions from $\rho \propto r^{-5}$ in the outskirts to a central cusp with $\rho \propto r^{-0.5}$. Its density and velocity dispersion profile are shown as the green curves in Fig. 1. Dattathri et al. (2025b) have shown that a CDM halo with this density profile is unstable to a dipole mode that dislodges the central cusp from the center.

Throughout, we adopt model units for which the gravitational constant, G , the total mass of the system, M_{tot} , and the scale radius, r_s , are all unity. For the tNFW halo, we instead set $G = M_{\text{vir}} = r_s = 1$, which implies a total mass of $M_{\text{tot}} = 1.2703$. When converting to physical units, we adopt the following fiducial scales: For the tNFW halo, we set $M_{\text{vir}} = 3 \times 10^9 h^{-1} M_\odot$, and we adopt a concentration parameter $c = 20$. For the ABG[0.1] and ABG[0.5] halos, we set $M_{\text{tot}} = 3 \times 10^9 h^{-1} M_\odot$ while their characteristic densities ρ_s are set to be equal to that of the tNFW halo. This implies that their scale radii, r_s , are 1.23 and 1.39 times larger, respectively, than that of the tNFW halo. This ensures that all three halos have the same characteristic core-crossing time $t_{\text{core}} \equiv r_s/v_s = 32$ Myr.

In the case of SIDM, where relaxation is driven by short-range interactions resulting in large-angle deflections, another relevant time scale is the characteristic collision time between individual collisions, which is given by

$$t_0 = (a \rho_s \sigma_m v_s)^{-1}, \quad (5)$$

(Balberg et al. 2002), where $a = 4/\sqrt{\pi} \approx 2.26$ for elastic hard-sphere-like interactions as assumed here. For a cross section of $\sigma_m = 1 \text{ cm}^2 \text{ g}^{-1}$, our fiducial halos have characteristic collision times of $t_0 = 1.91$ Gyr (tNFW), 1.55 Gyr (ABG[0.1]) and 1.37 Gyr (ABG[0.5]), respectively, which are fairly comparable to the halo crossing times, $t_{\text{cross}} = r_{\text{vir}}/V_{\text{vir}}$. Hence, for a cross section of $1 \text{ cm}^2 \text{ g}^{-1}$, a particle has of order one collision per crossing time.

3.2. Initial Conditions

Each host halo is sampled using $N = 10^6$ particles, for which we draw phase-space coordinates using the corresponding isotropic DF $f = f(E)$, which is computed using the standard Eddington inversion method (Eddington 1916; Binney & Tremaine 2008).

Physical systems must have a positive phase-space density,

i.e. $f(E) > 0$ at all E . In addition, according to Antonov's stability theorem, the system is stable if $df/dE < 0$ at all E (Antonov 1962; Henon 1973; Binney & Tremaine 2008). These conditions are satisfied for the tNFW and cNFW halos, but not for the ABG[0.1] and ABG[0.5] halos. In fact, as shown by Dattathri et al. (2025b) both have an inflection in their DF where $df/dE > 0$ (see also Section 6). Using both linear modal analysis and numerical simulations (assuming collisionless particles), Dattathri et al. (2025b) showed that such systems are unstable to a dipole mode in the center of the halo. This dipole mode saturates, becoming a long-lived solitonic mode, which is the dipolar equivalent of a quadrupolar bar mode in a disk galaxy. As part of our study, we investigate whether the same instability is present in the presence of self-interactions (see Section 5 below).

3.3. N-Body Simulations

All simulations are run using a modified version of the hierarchical N -body code `treecode`, written by Joshua Barnes. `treecode` uses a Barnes & Hut (1986) octree to compute accelerations based on a multipole expansion up to quadrupole order, and a straightforward second order leap-frog integration scheme to solve the equations of motion. The forces between the particles are softened using a simple Plummer softening.

To account for DM self-interactions, we closely follow Kochanek & White (2000) and Loeb & Weiner (2011). We adopt a constant velocity-independent cross section per unit mass equal to σ_m , and assume that a particle i can only scatter with one of its $k = 32$ nearest neighbours. We compute the expectation value for the number of scattering events of particle i in a time step Δt according to

$$\langle N_{\text{scat},i} \rangle = \frac{1}{2} \sum_j \sigma_m m_p W(r_{ij}, h_i) |\mathbf{v}_j - \mathbf{v}_i| \Delta t, \quad (6)$$

where the summation is over the k nearest neighbours, m_p is the particle mass, $W(r, h)$ is a kernel function that describes the 'extent' of the simulation particle in configuration space, and h_i is the smoothing length enclosing the k nearest neighbours of particle i . The factor $1/2$ is needed to avoid double counting of the scattering events. The actual number of interactions for particle i is drawn from a Poisson distribution with an expectation value equal to $\langle N_{\text{scat},i} \rangle$. Ideally, Δt is adjusted such that the probability for $N_{\text{scat}} > 1$ is sufficiently small,

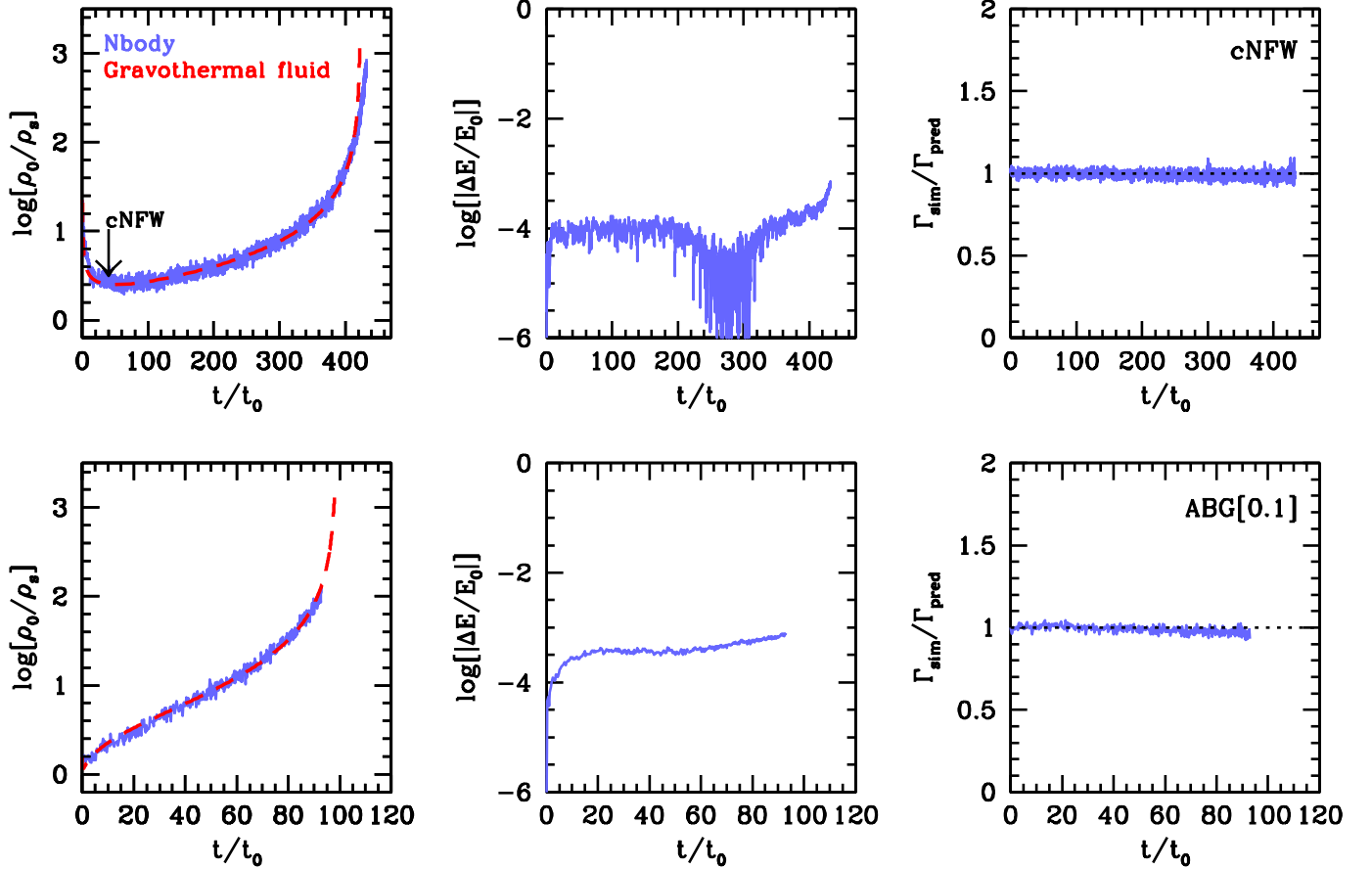


Figure 2. Evolution of two SIDM halos ($\sigma_m = 25 \text{ cm}^2 \text{ g}^{-1}$) as a function of time, expressed in units of the collision time t_0 . From left to right the panels show the evolution of the central density, the change in the total energy of the system normalized by the initial energy, and the ratio between the scattering rate in the N -body simulation, Γ_{sim} , and the predicted rate, Γ_{pred} , based on the instantaneous density and velocity dispersion (equation [14]). Top and bottom panels are for the tNFW and ABG[0.1] halos, respectively. The red dashed lines in the left-hand panels are the predictions based on the gravothermal fluid equations (see Appendix A). The downward arrow in the top-left panel marks the time of maximum core, which is used as the initial conditions for the cNFW simulations.

which requires that

$$\Delta t < (\Delta t)_{\text{SI}} \equiv \frac{\kappa}{\max(\rho_i \sigma_m \langle v_{\text{rel}} \rangle_i)} \quad (7)$$

Here κ is a free parameter, $\rho_i = m_p \sum_j W(r_{ij}, h_i)$, $\langle v_{\text{rel}} \rangle_i = \frac{1}{k} \sum_j |\mathbf{v}_j - \mathbf{v}_i|$ where the sum is over the k nearest neighbours of particle i , and the maximum in the denominator is taken over all N simulation particles i . It is common to pick a value for κ such that the probability of a particle scattering more than once per time step is negligible. For example, [Loeb & Weiner \(2011\)](#) use $\kappa \sim 0.01$. However, smaller values for κ make the simulation run slower, especially if, as in our case, the same time step is adopted for all particles. We have experimented extensively with different values of κ and find that even adopting values as large as $\kappa = 0.5$ yields results that are indistinguishable.

Once it has been determined that a target particle i has $N_{\text{scat}} \geq 1$ scattering events in time step Δt , we proceed as follows. For each scattering event, particle i is scattered against one of its 32 nearest neighbours, where the probability to scatter with particle j is given by

$$P_{ij} = \frac{W(r_{ij}, h_i) |\mathbf{v}_j - \mathbf{v}_i|}{\sum_j W(r_{ij}, h_i) |\mathbf{v}_j - \mathbf{v}_i|}, \quad (8)$$

Throughout, we assume elastic scattering. Hence, once a particle j is selected for a collision with i , we update their

velocities according to³

$$\begin{aligned} \mathbf{v}_i &= \mathbf{v}_{\text{cm}} + (\mathbf{v}_{ij}/2) \hat{\mathbf{e}} \\ \mathbf{v}_j &= \mathbf{v}_{\text{cm}} - (\mathbf{v}_{ij}/2) \hat{\mathbf{e}}, \end{aligned} \quad (9)$$

where $\mathbf{v}_{ij} = \mathbf{v}_j - \mathbf{v}_i$, $\mathbf{v}_{\text{cm}}$ is the center-of-mass velocity of the pair, and $\hat{\mathbf{e}}$ is a unit directional vector that we draw randomly from the unit sphere. Note that this procedure conserves both energy and linear momentum.

For our fiducial method, we adopt the standard (normalized) cubic spline function

$$W(r, h) = \frac{8}{\pi h^3} \begin{cases} 1 - 6q^2 - 6q^3 & 0 \leq q \leq \frac{1}{2} \\ 2(1 - q)^2 & \frac{1}{2} < q \leq 1 \\ 0 & q > 1 \end{cases}, \quad (10)$$

where $q = r/h$. However, we have also experimented with a top-hat kernel, setting

$$W(r, h) = \frac{3}{4\pi h^3} \begin{cases} 1 & q \leq 1 \\ 0 & q > 1 \end{cases}, \quad (11)$$

which is basically the approach used by [Kochanek & White \(2000\)](#), and find results that are indistinguishable.

³ This is only valid if all particles have equal mass, which is the case in all our experiments.

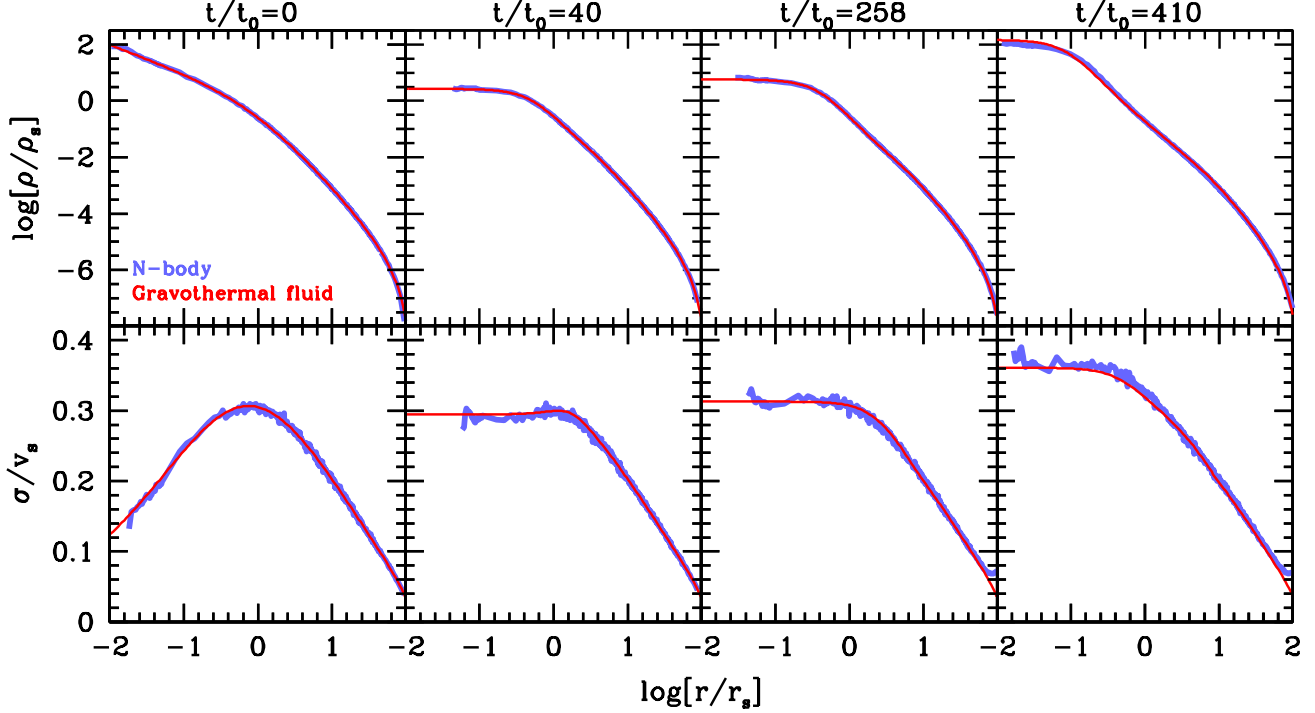


Figure 3. Evolution of the tNFW halo using SIDM with $\sigma_m = 25 \text{ cm}^2 \text{ g}^{-1}$. Top and bottom panels show the density profile and the 1D velocity dispersion profile, respectively, both normalized to their characteristic values. Different columns correspond to different times, as indicated at the top. In each panel, blue curves show the results from the N -body simulation, while the red lines show the predictions based on the gravothermal fluid code. Note the excellent agreement between the two even for $t/t_0 = 410$, which is well into the core collapse regime (cf. Fig. 2). The results at $t/t_0 = 40$, shown in the second column, serve as the initial conditions for the cNFW halo used in Section 4.1.

All particles advance with the same time step Δt , whose value is adjusted each time step so that

$$\Delta t = \min [(\Delta t)_{\text{SI}}, (\Delta t)_{\text{grav}}] . \quad (12)$$

Here $(\Delta t)_{\text{SI}}$ is the time step for the self-interactions given by equation (7) and

$$(\Delta t)_{\text{grav}} = \sqrt{2\eta \frac{\epsilon}{\max|\mathbf{a}_i|}} \quad (13)$$

is the time step needed to adequately resolve the orbits due to the gravitational force from all other particles. Here $\mathbf{a}_i$ is the gravitational acceleration of particle i , ϵ is the Plummer softening length, and η is a free parameter. Throughout, we use $\epsilon = 0.0232$ and $\eta = 0.025$, and we have verified that our results are stable to modest changes in these parameters. Although adjusting Δt each time implies a loss of symplecticity of our integration scheme, we find that the total energy in our simulations is conserved to better than one part in 10^3 up to the final stages of core collapse (see Section 3.4 below).

A useful measure to quantify self-interactions is the (local) Knudsen number, Kn , defined as the ratio of the mean free path, $\lambda_{\text{mfp}} = (\rho\sigma_m)^{-1}$, and the gravitational scale height (or Jeans length), $H = \sigma_v / \sqrt{4\pi G\rho}$, where σ_v is the local velocity dispersion. We have verified that all our simulations are at all time in the LMFP regime, for which $\text{Kn} \equiv \lambda_{\text{mfp}}/H > 1$. In this regime, the evolution of the system is self-similar when expressed as a function of t/t_0 (Balberg et al. 2002; Koda & Shapiro 2011; Yang et al. 2023).

3.4. Validation of self-interactions

In order to ensure that our implementation of the self-interactions is adequate, we evolve the systems in isolation

(i.e., no perturber) and (i) compare the scatter rate in the simulation to theoretical predictions and (ii) compare the evolution of the density and velocity dispersion profiles with predictions computed using a gravothermal fluid model.

The scattering rate of a halo is predicted to be

$$\begin{aligned} \Gamma_{\text{pred}} &= \frac{1}{2} \int n(\mathbf{x}) \rho(\mathbf{x}) \sigma_m \langle v_{\text{rel}} \rangle(\mathbf{x}) d^3\mathbf{x} \\ &= 8\sqrt{\pi} N_p \sigma_m \frac{1}{M_{\text{tot}}} \int_0^\infty \rho^2(r) \sigma_{1D}(r) r^2 dr , \end{aligned} \quad (14)$$

where $n(\mathbf{x}) = \rho(\mathbf{x})/m_p$ is the number density of particles, $\sigma_{1D}(\mathbf{x})$ is the local one-dimensional velocity dispersion, and in the second step we have assumed spherical symmetry and that the particles follow a Maxwell-Boltzmann distribution. We evaluate Γ_{pred} every few time steps using the instantaneous density and velocity dispersion profiles of the halo computed on the fly in the simulation code, and compare its value against the actual scattering rate in the simulation. The latter is evaluated as $\Gamma_{\text{sim}} = (\sum_i N_{\text{scat},i})/\Delta t$, where $N_{\text{scat},i}$ is the number of scattering events that particle i experienced during the time period Δt . As we demonstrate below, our simulations consistently have $\Gamma_{\text{sim}}/\Gamma_{\text{pred}} = 1$ within the shot noise.

As shown by numerous authors, the evolution of a spherical SIDM halo can be described using the conducting gravothermal fluid model first developed by Lynden-Bell & Eggleton (1980) to describe the collisional evolution of globular clusters (Balberg et al. 2002; Koda & Shapiro 2011; Pollack et al. 2015; Essig et al. 2019; Nishikawa et al. 2020). Appendix A describes our method of solving these gravothermal fluid equations, which closely follows the treatment of Nishikawa et al. (2020). Fig. 2 shows results for SIDM simulations with $\sigma_m = 25 \text{ cm}^2 \text{ g}^{-1}$ for both the tNFW halo (upper panels) and

the ABG[0.1] halo (lower panels). The left-hand panels show the evolution of the central density as a function of time, t , normalized by the characteristic collision time, t_0 , given by equation (5). For comparison, the red dashed lines show the predictions from the gravothermal fluid model, which are in excellent agreement with the simulation results. Note that the tNFW and ABG[0.1] halo evolve rather differently. Initially, the central density of the tNFW halo declines rapidly, reflecting the phase of core formation during which the r^{-1} cusp of the initial NFW profile is eroded away due to the self-interactions that cause a net flux of energy from the halo outskirts towards the center. The direction of this flux is dictated by the radial gradient of the velocity dispersion, which, for an NFW halo, is positive for $r \lesssim r_s$ (see right-hand panel of Fig. 1). Once the central gradient in the velocity dispersion has been removed (at $t/t_0 \sim 30$), and an isothermal core has been established, the energy flux due to self-interactions reverses sign, as now the central region is dynamically hotter than the halo outskirts. This triggers core collapse through the gravothermal catastrophe, resulting in an increase in central core density at an ever-increasing rate (e.g., Lynden-Bell & Wood 1968; Balberg et al. 2002; Koda & Shapiro 2011). During core collapse, both the central density and the central velocity dispersion increase with time, resulting in a drastic reduction of the simulation time steps causing the simulation to crawl to a halt. Therefore, we typically terminate the simulation once the central core density exceeds $\sim 500\rho_s$.

In the case of the tNFW halo, this happens when $t/t_0 \sim 420$. The vertical arrow at $t/t_0 = 40$ indicates the time of maximum core and marks the snapshot that we use as the initial conditions for the cNFW halo in the numerical experiments described in Section 4.1 below. In the case of the ABG[0.1] halo, the halo almost immediately starts to collapse after a very brief period in which it thermalizes its core. The core collapse progresses much faster than for the tNFW halo, reaching core collapse by $t/t_0 \sim 95$. This is a consequence of the much steeper gradient in velocity dispersion (cf. Fig. 1), which results in a larger conductive flux outward.

The middle panels of Fig. 2 show the evolution of the total energy of the system. As is apparent, energy conservation in our code remains at the level of 10^{-4} ($10^{-3.5}$) for the tNFW (ABG[0.1]) halo. There is some indication that $|\Delta E/E_0|$ increases during the final stages of core collapse, but at all times it remains below 10^{-3} . We emphasize that our energy conservation is better than that found in several other SIDM simulation codes. In particular, both Palubski et al. (2024) and Mace et al. (2024) have pointed to energy conservation issues unless they adopt a very small κ (typically $\kappa \lesssim 0.002$, compared to $\kappa = 0.5$ adopted here)⁴. We suspect that this is due to the fact that both Palubski et al. (2024) and Mace et al. (2024) use simulation codes with adaptive time steps (different particles have different Δt).

Finally, the right-hand panels of Fig. 2 plot the ratio $\Gamma_{\text{sim}}/\Gamma_{\text{pred}}$ as a function of time, showing that the scattering rate in the simulation is, at all times, in excellent agreement with the predictions. Together with excellent agreement with the results from the gravothermal fluid code, we conclude that our implementation of self-interactions is accurate. This is further corroborated by Fig. 3, which shows snapshots of the density and velocity dispersion profiles of the evolving tNFW halo at four epochs, as indicated. For comparison, the

red curves show the predictions from the gravothermal fluid equations and are once again in excellent agreement with the simulation results. Note that the results at $t/t_0 = 410$ shown in the right-most column correspond to a time well into core collapse.

4. CORE STALLING

In order to study whether and how core stalling operates in SIDM halos, we now inject a massive perturber of mass M_p into the cNFW and ABG[0.1] halos on a circular orbit at an initial radius r_i . We have verified (see also D25) that the value of this starting radius has no impact on any of our results, indicating that our results are not affected by potential transients due to the instantaneous introduction of the perturber. We describe the results for the cNFW and ABG[0.1] halos in turn, starting with the former.

4.1. Stalling in the cNFW halo

Fig. 4 shows the evolution of the distance and velocity with respect to the halo's center-of-mass of perturbers of mass $M_p = 0.01$ (top panels) and $M_p = 0.03$ (bottom panels), in model units. Both perturbers are started from a circular orbit at $r_i = 7.5r_s$ and are modelled as point particles but with a Plummer softening length of $\epsilon_p = 0.03r_s$ and $\epsilon_p = 0.05r_s$, respectively⁵. In the case without self-interactions, indicated by the black lines, the perturber rapidly sinks to $r \sim 0.25r_s$ (indicated by the horizontal dotted line), after which the rate of infall suddenly slows down dramatically. For comparison, the red dashed lines indicate the predictions based on Chandrasekhar's formula (Chandrasekhar 1943), obtained by integrating the orbit of the perturber in the (initial) density distribution of the host halo accounting for a dynamical friction

$$\frac{d\mathbf{v}_p}{dt} = -4\pi \frac{G^2 M_p}{v_p^2} \ln \Lambda \rho(< v_p) \frac{\mathbf{v}_p}{v_p}. \quad (15)$$

Here, $\mathbf{v}_p$ is the velocity of the perturber and $\rho(< v_p)$ is the density of host halo particles with speeds less than $v_p = |\mathbf{v}_p|$, which is computed assuming that the halo particles locally follow a Maxwellian velocity distribution (see Binney & Tremaine 2008). We tune the Coulomb logarithm, $\ln \Lambda$, which is rather uncertain (see discussion in Mo et al. 2010), to fit the simulation results at early times, which yields values in the range $2.5 \lesssim \ln \Lambda \lesssim 4$. The Chandrasekhar-based predictions nicely agree with the simulation results, but only up to the point where the perturber crosses the horizontal dotted line. This is a manifestation of core stalling and in what follows we refer to the radius where the simulation results suddenly and drastically deviate from the Chandrasekhar prediction as the stalling radius, r_{stall} . Note that the stalling is not permanent; after ~ 2 Gyr the perturber continues to sink towards the center of the halo, but at a drastically reduced infall rate.

The colored curves show the same results, but for SIDM simulations with different values for σ_m , as indicated. In the case of $M_p = 0.01$, the infall of the perturber is identical to that in the CDM case and independent of the value of σ_m , up to the point where the perturber experiences core stalling in the CDM case, at $t \sim 13$ Gyr. Beyond that point in time, though, there is a pronounced dependence on the interaction cross section: for $\sigma_m \lesssim 3 \text{ cm}^2 \text{ g}^{-1}$, the results are virtually indistinguishable

⁴ κ is called `DtimeFac` in Mace et al. (2024)

⁵ Changing the softening length of the perturber has a weak effect on the infall rate, but qualitatively the results are unaffected by changes in ϵ_p

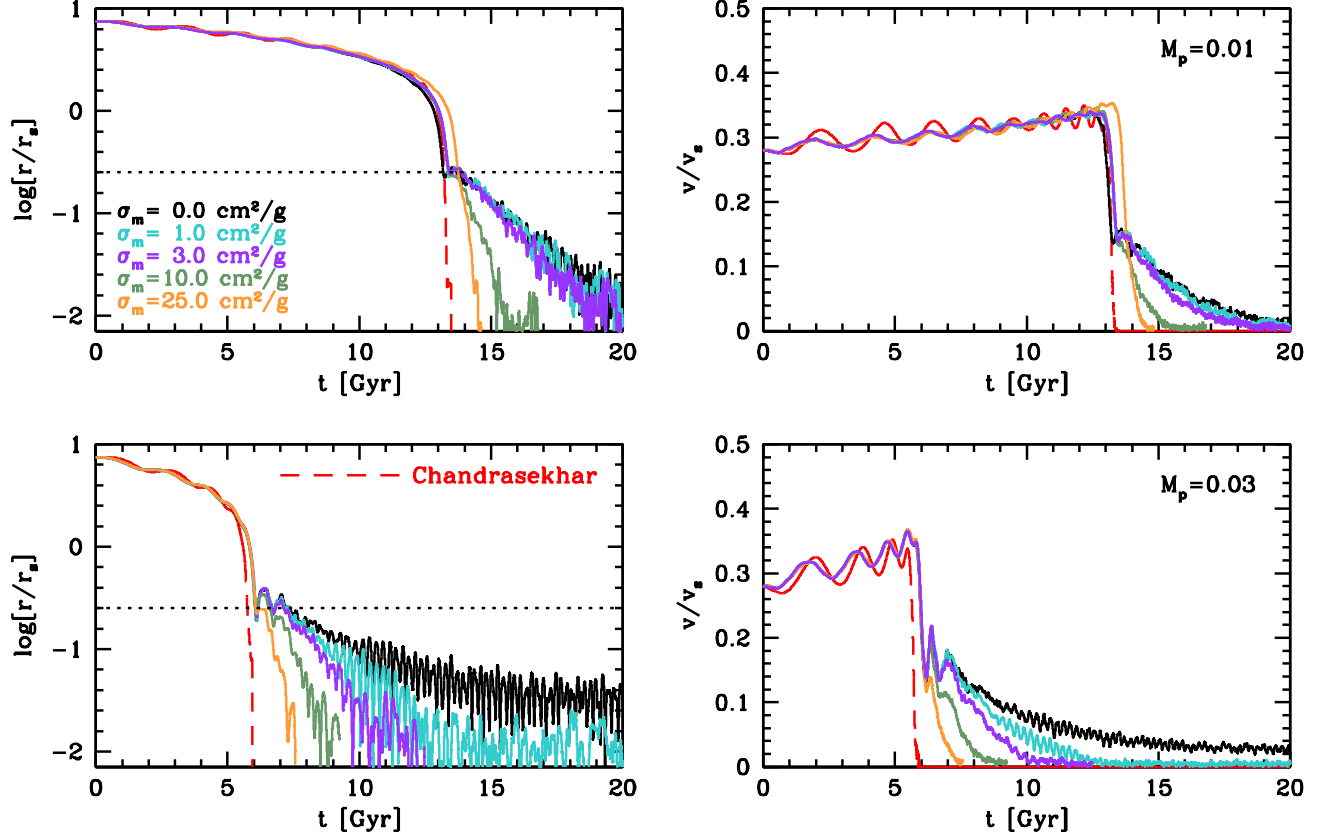


Figure 4. Dynamical friction acting on perturbors of mass $M_p = 0.01$ (top panels) and $M_p = 0.03$ (bottom panels) in the cNFW halo with different levels of self-interactions, as indicated. Left and right-hand panels show the evolution of the distance and velocity of the perturbor with respect to the center-of-mass of the host halo. The red dashed lines indicates the predictions based on Chandrasekhar’s dynamical friction formula and the initial density profile of the host halo. The horizontal dotted line marks the stalling radius, r_{stall} . Radii and velocities are in N -body units.

from those in the case of CDM. However, for $\sigma_m = 10 \text{ cm}^2 \text{ g}^{-1}$ the post-stalling infall rate is substantially faster. For $\sigma_m = 25 \text{ cm}^2 \text{ g}^{-1}$, there is no longer any sign of core stalling; while it now takes a little longer for the perturbor to reach r_{stall} , the perturbor sinks past this radius without delay. In the case of $M_p = 0.03$, the trends are similar; the main difference being that it now takes roughly half as long to reach the core stalling radius⁶. Note how once again a larger interaction cross section causes a weaker stalling and/or a faster infall inside of the stalling radius. In the case where $\sigma_m = 25 \text{ cm}^2 \text{ g}^{-1}$, there is barely any core stalling at all; except for a minor hiccup in the infall rate around 6 Gyr, the perturbor continues to sink in rapidly.

4.1.1. Impact on the host system

Fig. 5 shows how the infalling perturbor impacts the density profile of the host halo. In each panel, the colored lines that range from blue to violet to red indicate the evolution of the halo mass fractions enclosed within the fixed radii of $r/r_s = 0.1, 0.25, 0.5, 1.0, 2.0, 5.0$ and 10.0 . The thicker and thinner lines of the same color correspond to simulations with and without a perturbor, respectively. Different rows correspond to different collisional cross sections, as indicated (in units of $\text{cm}^2 \text{ g}^{-1}$), while the left- and right-hand panels show the results for $M_p = 0.01$ and 0.03 , respectively. Finally, the thick green line indicates the mass enclosed by the instantaneous

radius of the perturbor, which declines as the perturbor sinks towards the center.

Let us first focus on the results for $M_p = 0.03$. In the CDM case (top right panel), once the perturbor reaches its stalling radius at $t \sim 6$ Gyr, there is a sudden change in the enclosed mass profile of the host halo. More specifically, the enclosed mass drops and the effect is more pronounced at smaller radii. As discussed below, this corresponds to the perturbor tidally perturbing the central core. The energy required to displace the dark matter from the central core outwards is taken from the kinetic energy of the perturbor, which explains why, at the same time, the velocity of the perturbor drops precipitously (cf. right-hand panel of Fig. 4). The perturbor briefly ‘stalls’ for a period of ~ 2 Gyr, but then continues its inward trajectory, slowly sinking towards the center of its host. During this period, the enclosed density within the innermost radius first increases slightly as a result of adiabatic contraction of the core in response to sinking in of the perturbor. At late times, once the perturbor has sunk inside of $0.1 r_s$, its continued motion through the central core slightly heats the core, causing the enclosed density within the innermost radius to decrease again slightly.

In the case of SIDM, the behavior is very similar up to $t \approx 6$ Gyr when the perturbor reaches the core stalling radius. As in the case of CDM, the perturbor initially reduces the centrally enclosed mass, but subsequently the mass enclosed inside $\sim 0.25 r_s$ starts to *increase* at a rate that is a strong function of σ_m . We emphasize that this enclosed mass does not include the mass of the perturbor. For $\sigma_m = 25 \text{ cm}^2 \text{ g}^{-1}$

⁶ Note that this time scale does not scale linearly with the inverse of M_p due to the fact that the perturbors have different softening lengths.

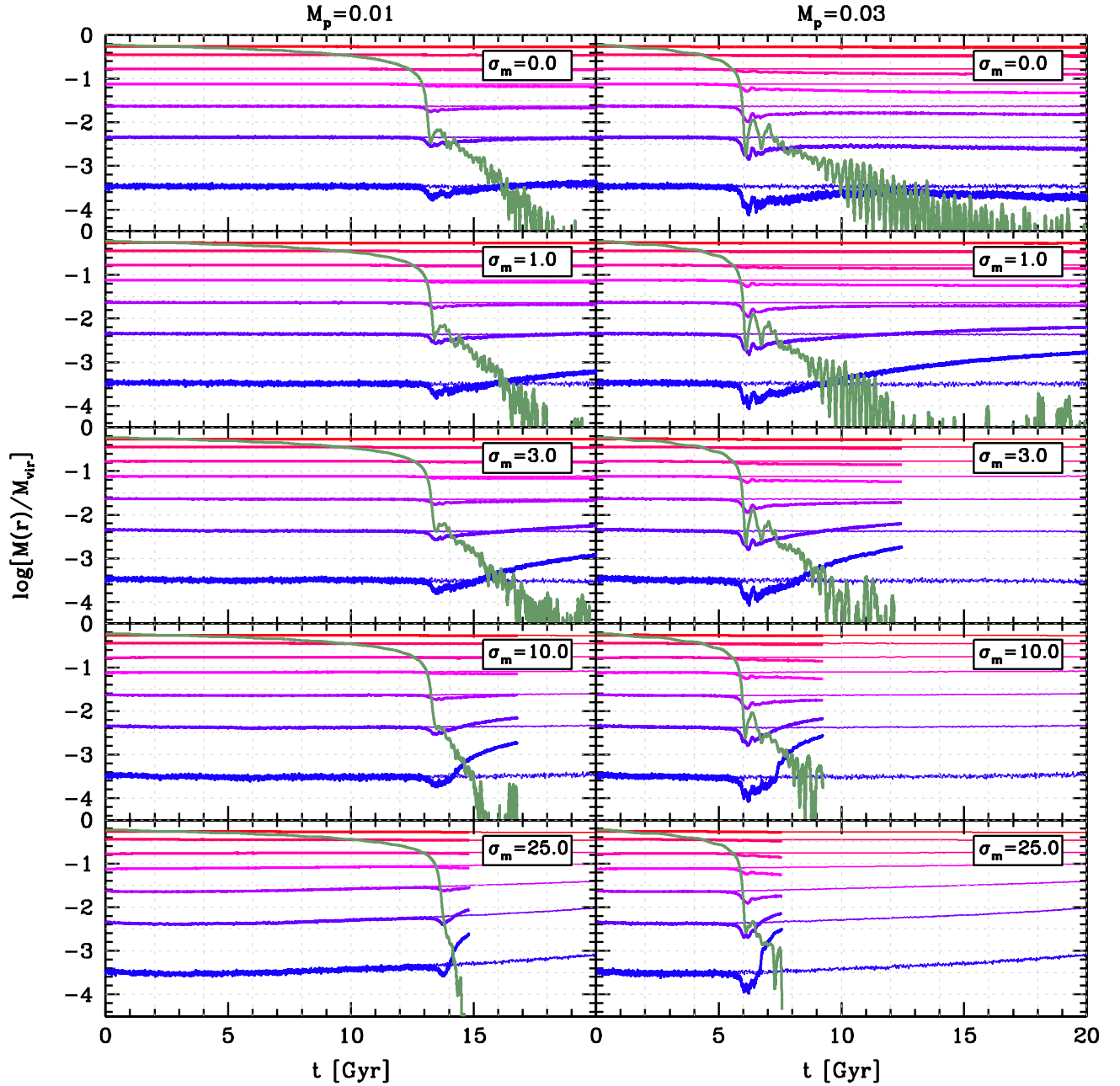


Figure 5. Enclosed masses as a function of time. The colored lines, ranging from blue to red, indicate the evolution of the enclosed mass fractions at different radii: $r/r_s = 0.1, 0.25, 0.5, 1.0, 2.0, 5.0$ and 10.0 . Thicker and thinner lines of the same color correspond to the simulations with and without a perturber, respectively. The thick, green line indicates the mass enclosed by the instantaneous radius of the perturber, which declines as the perturber sinks towards the center. Different columns and rows correspond to different perturber masses and collisional cross sections, respectively, as indicated. Note how the perturber affects the mass distribution of the host, and catalyzes core collapse in the case of SIDM. The latter is evident from the fact that when the perturber crosses a particular radius, the enclosed mass within that radius (which is indicated by the thick lines and which does not include the perturber itself) increases rapidly compared to the case without a perturber (thin lines).

the mass enclosed with $0.1r_s$ increases by more than an order of magnitude in less than 1 Gyr: the perturber, aided by the self-interactions, triggers a sudden and pronounced collapse of the central core. Note how this core collapse goes hand in hand with (i.e., occurs on the same time scale as) the perturber sinking towards the center.

For the less massive perturber with $M_p = 0.01$ shown in the left-hand panels, the behavior is qualitatively similar. Once the perturber reaches the core stalling radius, which takes significantly longer than in the case of the more massive perturber, it distorts the central density, after which it triggers a rapid

collapse of the core at a rate dictated by σ_m . The abruptness of the core collapse is most apparent when one compares the results with those in the absence of a perturber shown as the thin lines.

Fig. 6 is another illustration of the same. It shows the density (top panels) and velocity dispersion (bottom panels) profiles of the cNFW halo with the $M_p = 0.03$ perturber at four different epochs, as indicated. Different columns correspond to different values of the collisional cross section, while the halo-centric radius of the perturber at each epoch is indicated by a downward-pointing arrow. The left-hand column

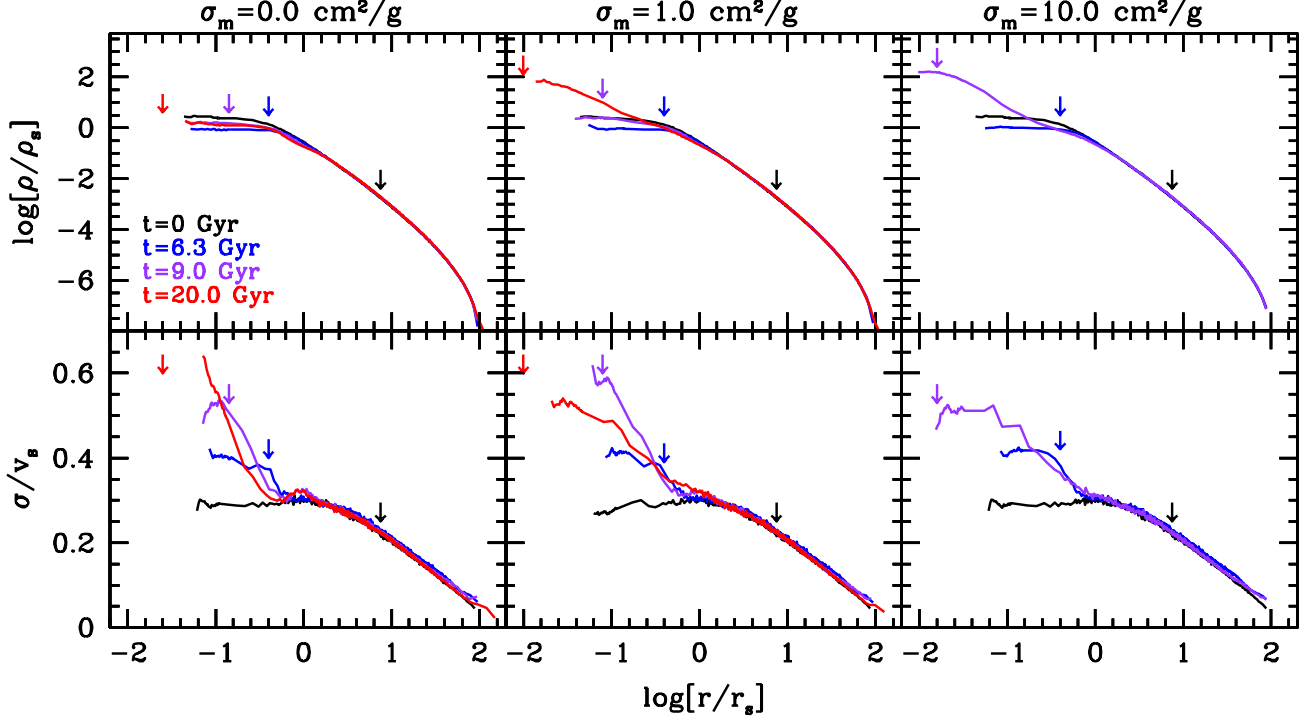


Figure 6. The density (top) and radial velocity dispersion (bottom) profiles for the cNFW halo with a perturber of mass $M_p = 0.03$ starting at $t = 0$ on a circular orbit at $r = 7.5r_s$ (cf. Fig. 4). Results are shown at four different epochs (different colors, as indicated), and for three different values of the collisional cross section (different columns, as indicated at the top). The arrows mark the halo-centric distance of the perturber at the time corresponding to its color. In the case of $\sigma_m = 10 \text{ cm}^2 \text{ g}^{-1}$, the simulation is terminated shortly after $t = 9$ Gyr, when the halo has undergone core collapse, which explains why there are no results shown for $t = 20$ Gyr. Note that the innermost data point always corresponds to the radius that encloses 30 particles, which thus is smaller when the central density is higher. See text for discussion.

shows the results for CDM. Note how the perturber initially reduces the central density, enlarging the core, after which the central density slowly creeps up as the perturber sinks inward. The transition of the central density profile around $t = 6$ Gyr is extremely rapid. For the first ~ 5.8 Gyr, as the perturber sinks from its initial radius $r_i = 7.5$ to $r \sim 0.25$, no significant change in the density profile of the host halo is noticeable. Then suddenly the perturber’s velocity plummets from $v \approx 0.34$ to 0.15 in less than 0.3 Gyr. The momentum lost by the perturber is transferred to the central core, which causes a reduction of the central density by roughly a factor of four. This sudden and drastic transfer of momentum is associated with the tidal shredding of the core by the perturber (Goerdt et al. 2010). Subsequently, the perturber temporarily stalls its infall for a period of ~ 2 Gyr before it continues to slowly sink towards the center. During this second inspiral phase, the central density slowly increases again. By $t = 20$ Gyr, the central density has roughly doubled compared to what it was during the stalling phase, while the central velocity dispersion has increased dramatically. The evolution during this phase is a complicated competition between two effects: adiabatic contraction and dynamical heating. The sinking of the perturber deepens the central potential, and since the inspiral time scale is long compared to the local dynamical time, this process conserves adiabatic invariants (Blumenthal et al. 1986; Weinberg 1994a). This tends to increase both the central density and the velocity dispersion. At the same time, though, the perturber heats the core by transferring its momentum to the dark matter particles. Such dynamical heating typically causes a decrease in the density and, because of the negative heat capacity, a decrease in the central velocity dispersion as well. As is apparent from Fig. 6, adiabatic contraction is clearly winning out, with

both the density and central velocity dispersion increasing as the perturber slowly makes its way to the center.

In the case with $\sigma_m = 1.0 \text{ cm}^2 \text{ g}^{-1}$ shown in the middle column, the density and velocity dispersion profiles are similar to those in the case of CDM up to $t \sim 9$ Gyr. However, from 9 to 20 Gyrs the evolution of the host halo is very different. The perturber catalyzes core collapse, causing an increase in the central density. As we demonstrate in Section 6.3, this catalyzation is driven by the adiabatic contraction of the core due to the sinking perturber⁷. This contraction boosts the central density and velocity dispersion, which in turn increases the outward conduction of heat, thereby further promoting the contraction of the core. And since heat conduction is more efficient for larger cross sections, while stalling is less efficient, the entire process of infall and core collapse proceeds faster for larger σ_m . Indeed, in the case where $\sigma_m = 10 \text{ cm}^2 \text{ g}^{-1}$, shown in the right-hand column, the system has already completed core collapse by $t = 9$ Gyr (which is why no results are shown for $t = 20$ Gyr).

4.2. Stalling in the ABG[0.1] Halo

Next, we turn our attention to the ABG[0.1] halo. Fig. 7 shows the evolution of the halo-centric distance of perturbers of different mass (different colors, as indicated) that at $t = 0$ are injected into the ABG[0.1] halo along a circular orbit at $r = 2.5r_s$. As before, the perturber masses are expressed in model units, that is, in units of the total mass of the ABG halo. Each perturber is modeled as a Plummer sphere with softening length $\epsilon_p = 0.0232$, which is the same value as used for dark

⁷ The fact that adiabatic contraction can trigger accelerated core collapse has also been pointed out in Elbert et al. (2018).

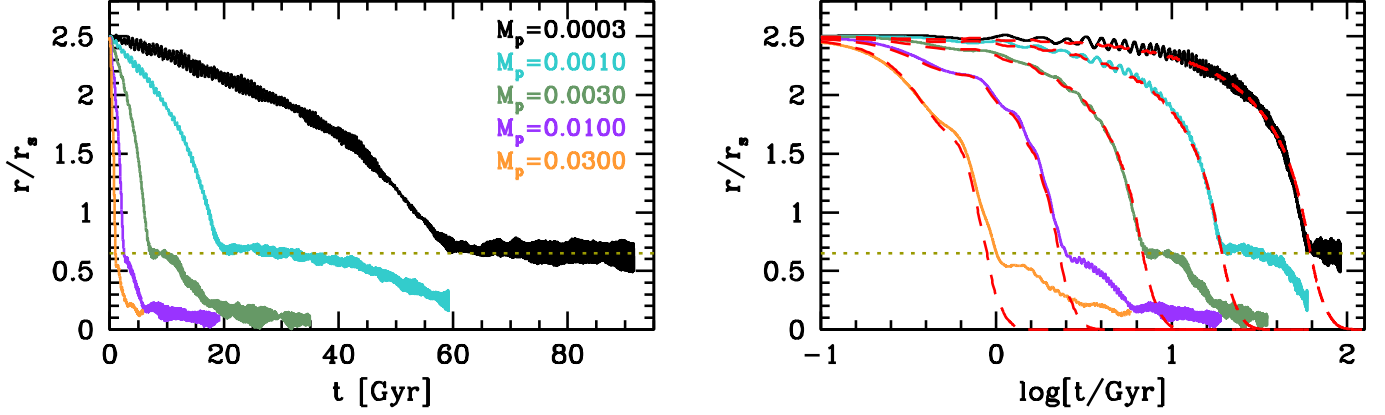


Figure 7. Evolution of the orbital radius of a massive perturber of mass M_p sinking inside the ABG[0.1] halo assuming CDM. Different colors correspond to different M_p , as indicated. Left and right hand panels show results as function of linear and logarithmic time, to better highlight features at late and early times, respectively. Low mass perturbers experience clear stalling, while more massive halos seem to ‘crunch’ through the stalling radius. Note, though, that even the most massive perturber experiences a small ‘hick-up’ sinking towards the center close to the stalling radius, r_{stall} , indicated by the horizontal dotted line. The red dashed lines in the right hand panel indicate the predictions based on Chandrasekhar’s dynamical friction formula and the initial density profile of the host halo. These fit the data perfectly, but only up to the point of core stalling.

matter particles. We have verified that changing the softening length only impacts the time scale at which the perturber sinks in without qualitatively changing the outcome. Note how the low-mass perturbers experience a pronounced core stalling at $r = r_{\text{stall}} \simeq 0.65$ (indicated by the horizontal red dotted line), significantly further out than in the case of the cNFW halo. More massive perturbers appear to stall for a shorter period before they continue their infall to the center. In fact, the most massive perturber, which has a mass that is 3 percent of that of the host halo, barely seems to stall at all. These results are in excellent agreement⁸ with D25, who showed that this phenomenology is related to evolution in the DF of the host halo. We address this in detail in Section 6 below.

Fig. 8 compares the results for different values of the collision cross section, σ_m , as indicated. Once more, the perturbers start from a circular orbit at $r = 2.5$ and are modeled as Plummer spheres of size $\epsilon_p = 0.0232$. The upper panels show the results for a perturber of mass $M_p = 0.001$. In the case of CDM ($\sigma_m = 0$), the perturber experiences pronounced and prolonged stalling at $r_{\text{stall}} \simeq 0.65$, indicated by the red horizontal dotted line. Similarly to what we found in the case of the cNFW halo, larger values of σ_m result in weaker stalling. Even for a modest $\sigma_m = 0.3 \text{ cm}^2 \text{ g}^{-1}$, the pronounced stalling that is apparent in the case of CDM is gone, and the perturber sinks past r_{stall} without delay. For $\sigma_m = 1.0 \text{ cm}^2 \text{ g}^{-1}$, the perturber initially sinks past r_{stall} at a rate that is significantly faster than in the case with $\sigma_m = 0.3 \text{ cm}^2 \text{ g}^{-1}$, but then seems to temporally stall at $r \simeq 0.2$, before continuing its inward trajectory. For $\sigma_m = 3.0 \text{ cm}^2 \text{ g}^{-1}$, the initial inspiral rate is slower than in the case of CDM. However, there is no indication of any stalling; the perturber passes r_{stall} without interruption and comes to rest at the center of the halo after ~ 29 Gyr. Finally, for $\sigma_m = 10 \text{ cm}^2 \text{ g}^{-1}$, results are only shown up to $t \simeq 12.9$ Gyr. The reason is that by this time the ABG halo has already undergone core collapse. Interestingly, by $t = 12.9$ Gyr, the perturber has clearly sunk less far than in the other cases, suggesting that it experienced significantly reduced friction. As we will see in Section 6.2, this is due to the fact that the self-interactions have significantly modified

the DF.

The bottom panels show the results for a perturber that is ten time more massive. With CDM, the perturber first stalls at $r = r_{\text{stall}}$, but only for ~ 1 Gyr. Thereafter, it continues to sink inwards, only to stall once more at $t \sim 6.5$ Gyr at a radius $r \simeq 0.2$. When $\sigma_m \gtrsim 1.0 \text{ cm}^2 \text{ g}^{-1}$ this second stalling disappears and for $\sigma_m \gtrsim 10 \text{ cm}^2 \text{ g}^{-1}$ there is no longer any sign of stalling. In fact, for $\sigma_m = 10 \text{ cm}^2 \text{ g}^{-1}$, the perturber comes to rest at the center of the halo after only 4 Gyr. For comparison, with CDM the perturber is still orbiting through the central region of the core at an appreciable speed even after 10 Gyr of evolution.

Fig. 9 is similar to Fig. 5, but shows the results of the simulations with the ABG[0.1] halo. In the case with $M_p = 0.001$, shown in the left-hand column, the perturber has barely any impact on the host halo. This is evident from the fact that the thin colored lines that show the results without a perturber are barely if at all visible; they lie below the thick colored lines. There are two clear trends with increasing cross section for self-interactions: (i) the perturber experiences less stalling (i.e., it sinks faster after it has crossed the stalling radius), and (ii) the halo experiences more rapid core collapse. In the case with $\sigma_m = 10 \text{ cm}^2 \text{ g}^{-1}$, shown in the bottom panel, the core of the host halo has collapsed by $t \sim 15$ Gyr, when we are forced to terminate the simulation. Up to that point, the perturber has barely sunk in at all and has had no noticeable impact on the host system (i.e., without perturber the host halo experiences core collapse in the same way and in the same amount of time). For $M_p = 0.01$, shown in the right-hand column, the sinking rate of the perturber is significantly enhanced, as is the impact that the perturber has on its host halo. In particular, for $\sigma_m \gtrsim 1.0 \text{ cm}^2 \text{ g}^{-1}$ it is evident that the perturber causes an acceleration of the core collapse of the host halo, albeit less pronounced as in the case of the cNFW halo. All in all, though, the trends are qualitatively similar for the two different host halos.

5. DYNAMICAL BUOYANCY AND THE DIPOLE INSTABILITY

We now turn our attention to dynamical buoyancy and the associated dipole instability. Here, to study the impact of self-interactions, we consider the ABG[0.5] halo, which in the case

⁸ This is noteworthy because the simulations were run with independent N-body codes.

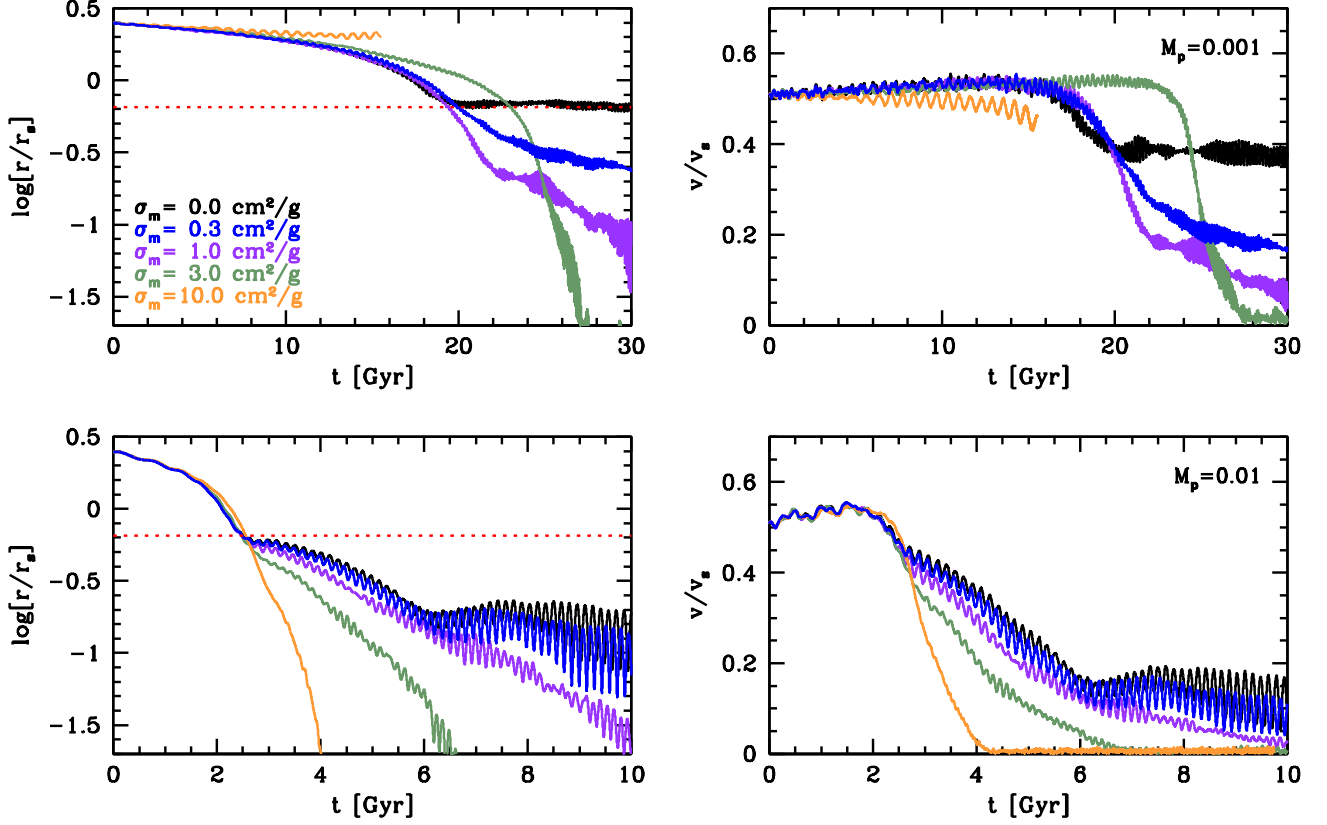


Figure 8. Same as Fig. 4 but for the ABG[0.1] host halo and a perturber of mass $M_p = 0.001$ (top panels) and $M_p = 0.01$ (bottom panels). Note how increasing the cross section for self-interactions suppresses core stalling at late times, while delaying the infall at early times.

of CDM is known to manifest a pronounced dipole instability (Dattathri et al. 2025b). We evolve the halo in isolation and evolve it for roughly a Hubble time for four different values of σ_m : 0 (CDM), 0.1, 0.3, and $1.0 \text{ cm}^2 \text{ g}^{-1}$. In order to be able to probe the impact of dynamical buoyancy, we run two simulations for each value of σ_m ; one without a perturber and one in which, at $t = 0$, we place a perturber of mass $M_p = 0.01$ at rest at the center of the system. In the latter case, the DF that is used to setup the ICs is adjusted accordingly⁹. The perturber particle has the same softening length as the regular particles but is not susceptible to self-interactions. It represents a massive object such as a SMBH or a nuclear star cluster that, in the absence of buoyancy and stalling, will have a tendency to sink towards the center by dynamical friction.

For completeness, in order to demonstrate the presence of the dipole mode, Fig. 10 shows the projected over/underdensity, $\Delta\Sigma = \Sigma - \Sigma_0$, where Σ_0 is the unperturbed analytic projected density, for four snapshots of the CDM simulation without perturber. While the snapshot at $t = 0$ mainly reveals shot noise due to the discreteness of the N -body system, at later times a clear dipole is evident. Importantly, the dipole is not stationary but rather rotating in a plane. We emphasize that these results are in excellent agreement, both qualitatively and quantitatively, with those of Dattathri et al. (2025b, cf. their Fig. 3), despite significant differences in the simulations¹⁰. As discussed in detail in Dattathri et al.

(2025b), the strong dipole apparent in the center is not actually the dipole mode; rather, it is the central $r^{-0.5}$ cusp that has been dislodged and set in motion by the mode. The cusp continues to slosh through the central region of the system, along a roughly elliptical orbit, without any sign of ever coming back to rest.

The left-hand panels of Fig. 11 show the evolution of the density at four different radii with respect to the system’s center of mass (different colors). The dashed lines of the corresponding color indicate the densities at $t = 0$ and are shown to highlight the extent to which the densities evolve over time. Different panels correspond to different values of the interaction cross section, σ_m , as indicated. First, let us focus on the case of CDM (top panel). Clearly, the central densities at $r \lesssim 0.1 r_s$ (blue and violet curves) reveal strong evolution, consisting of a large oscillation with small oscillations superimposed. This reflects the onset and development of the dipole mode, with the small oscillations reflecting the motion of the cusp along its elliptical orbit. Note that most of the action is confined to the central region of the system; the density at $\log[r/r_s] = -0.25$ (red curve) remains constant with time within the noise.

As we dial up the self-interactions, the amplitude of the oscillations becomes less pronounced, and for $\sigma_m = 1.0 \text{ cm}^2 \text{ g}^{-1}$ (bottom panel), the small-scale oscillations have completely disappeared. Rather, after an initial decline, the density in the entire central region starts to increase, which reflects the onset of core collapse.

The right-hand panels of Fig. 11 show how all this is related

⁹ We find that simply placing the perturber at rest in the simulation that is initialized without perturber yields results that are virtually indistinguishable.

¹⁰ The simulation analyzed in Dattathri et al. (2025b) used $N_p = 5 \times 10^6$ particles with a mass spectrum that boosts mass resolution near the center, and was as run using the code Gyrfalcon (Dehnen 2002) with hierarchical time

stepping. For comparison, the simulation presented here only uses $N_p = 10^6$ particles, all of identical mass, and is run using treecode without hierarchical time stepping.

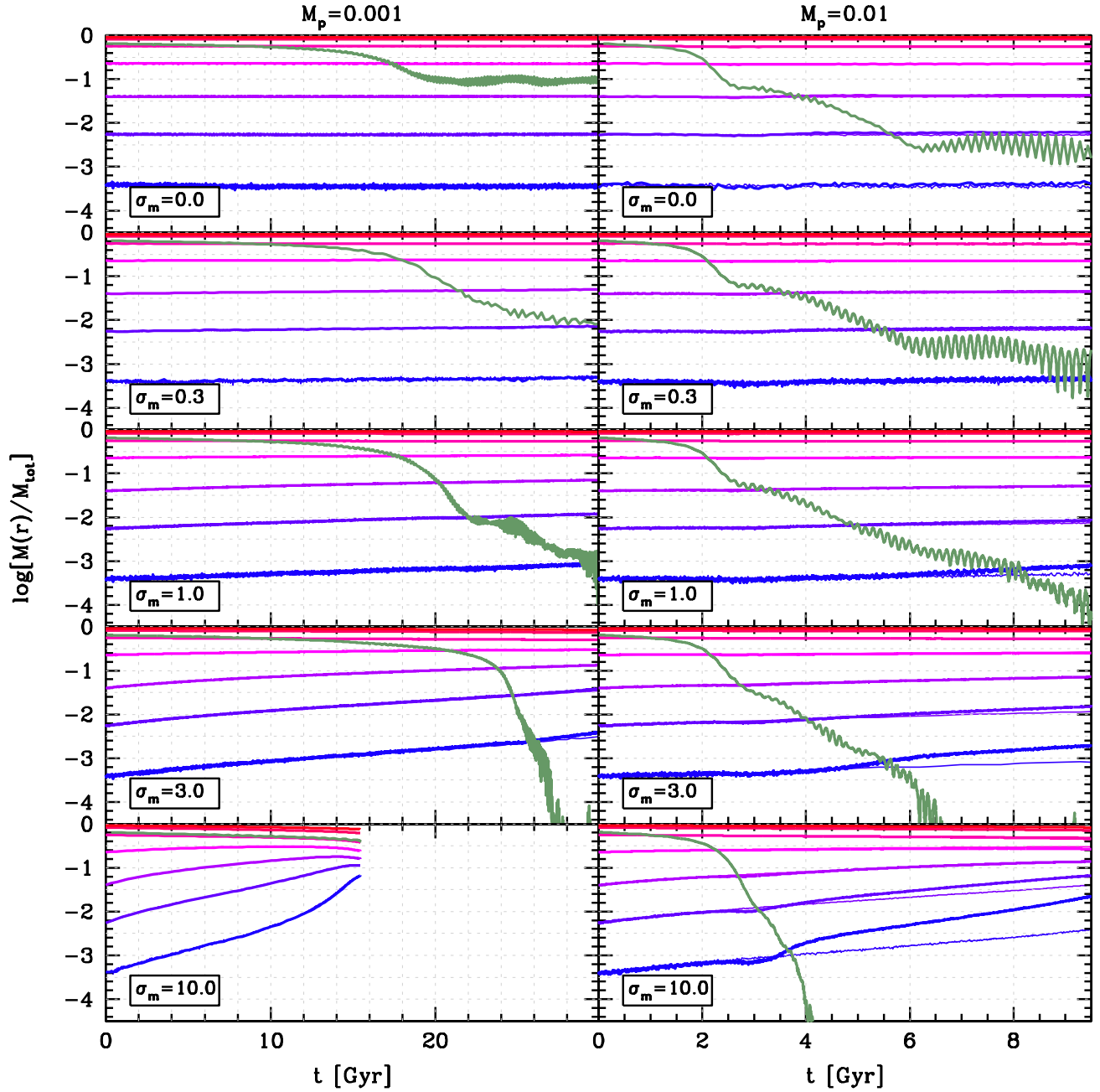


Figure 9. Same as Fig. 5, but for the ABG[0.1] halo. Note how with increasing self-interaction cross section the perturber sinks more rapidly towards the center, catalyzing core collapse along the way. In the lower left-hand panel the simulation is terminated at $t = 15.3$ Gyr when the host halo has undergone core collapse. By that time the perturber has only experienced a small amount of dynamical friction (cf. orange curves in top panels of Fig. 8).

to buoyancy. Once again, different panels correspond to different values for the self-interaction cross section, as indicated. In each panel, the black curve shows the temporal evolution of the center of mass of the massive perturber that at $t = 0$ starts out at rest at the center of the system. Note how, in the case of CDM, the perturber moves outward during the first 2 Gyr, reaching as far as $0.3r_s$; this is a manifestation of dynamical buoyancy. Later, it sinks back somewhat before stalling at $\sim 0.1r_s$ along a highly eccentric orbit. We also computed the center of mass of the 500 densest particles, which accurately traces the motion of the dislodged cusp, and find that it accurately follows the trajectory of the massive perturber (not shown). This indicates that the buoyancy experienced by the massive perturber is a direct manifestation of the dipole insta-

bility; the massive perturber simply ‘rides along’, and remains centered on, the dislodged cusp (see also D25).

The azure curve in the top-right panel of Fig. 11 represents the trajectory of the center of mass of the 500 densest particles (which basically tracks the location of the original cusp) in the simulation *without* perturber. As in the simulation with perturber, the dipole instability dislodges the cusp and sets it in motion along a trajectory that is very similar to that of the massive perturber (cf. Fig. 10), albeit with a time-lag of ~ 1.5 Gyr. This time lag reflects the fact that the massive perturber has slightly modified the DF of the ICs, which in turn impacts the growth of the dipole mode.

With increasing cross section, the amplitudes of the radial excursions of both the massive perturber (black curves)

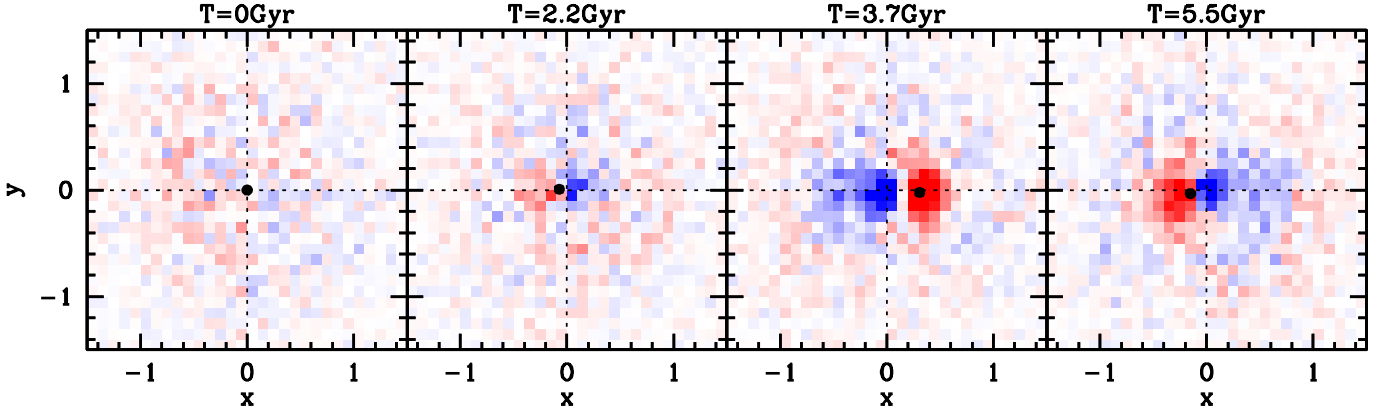


Figure 10. Four snapshots of the excess surface density $\Delta\Sigma = \Sigma(x, y) - \Sigma_0(x, y)$ of the ABG[0.5] system evolved assuming CDM (no self-interactions), where $\Sigma_0(x, y)$ is the analytical surface density at $t = 0$. Red and blue pixels indicate over- and under-densities respectively (cf. Fig. 3 in Dattathri et al. 2025b). At $t = 0$ the system starts out symmetrically, with only some discreteness noise. By $t = 1.8$ Gyr, a small dipole appears above the noise, which grows in amplitude as it rotates and sloshes through the central region. At each snapshot the solid dot marks the center-of-mass of the 500 densest particles, the trajectory of which is shown by the blue curve in the top-right panel of Fig. 11.

and the dislodged cusp in the simulations without perturber (azure curves) become weaker and weaker. In fact, for $\sigma_m = 1.0 \text{ cm}^2 \text{ g}^{-1}$ (bottom-left), the massive perturber initially moves slightly outward (never exceeding $0.1r_s$), but then rapidly sinks back and settles at the center of the system. The small remaining ‘jitter’ in the radial position of the massive perturber is due to Brownian motion. Hence, increasing σ_m inhibits the growth of the dipole instability and thus also the amplitude of dynamical buoyancy.

6. DISCUSSION

We have shown that self-interactions among dark matter particles have a profound impact on core dynamics. In particular, self-interactions suppress core stalling, dynamical buoyancy, and the onset of the dipole instability. In fact, for sufficiently large σ_m core stalling can be completely avoided, and the perturber sinks to the center much faster than in the case of CDM. Interestingly, in doing so it causes a drastically accelerated core collapse. Here, we discuss the physical origin of these effects.

As discussed in Section 2, both core stalling and the dipole instability are related to features in the DF of the host system. In particular, stalling happens when the distribution function of the host system has a plateau (i.e., $df/dE \approx 0$) at the orbital energy of the perturber, while systems with an inflection in their DF, that is, a region in energy-space where $df/dE > 0$, are unstable to the dipole mode¹¹. Self-interactions modify the DF, and this might thus explain why core dynamics in SIDM halos can be substantially different from that in CDM halo. In particular, self-interactions cause a diffusion in phase space, which tends to erase (small-scale) gradients, and drive the system toward an isothermal form $f(E) \propto \exp(-E/\sigma^2)$ (Spitzer & Shapiro 1972), at least in the central core region. Since the corresponding df/dE is always negative, we ought to expect that in systems in which the self-interactions have established isothermality, core stalling and/or buoyancy are absent. This is qualitatively consistent with the trends seen in Section 4. In what follows, we test this hypothesis in detail by computing the evolution in the DF as the perturber sinks towards the center. For each simulation output, we calculate

¹¹ We emphasize that it remains unproven whether this is always the case. However, in all such systems studied by the authors we have found the system to be unstable.

the phase space distribution of the simulation particles as a function of their instantaneous energy, $f(E)$, using the method of Dattathri et al. (2025b), summarized in Appendix B.

6.1. Suppression of the dipole instability & buoyancy

We first focus on the cNFW halo, the ABG[0.1] halo, and the ABG[0.5] halo in isolation (i.e., without perturber). Fig. 12 shows the DFs of the three halos (different columns) at four different epochs (different colors), which have been slightly offset from each other in the vertical direction for the sake of clarity. Different columns correspond to different values for the cross section for self-interactions, as indicated, with the top row corresponding to CDM (i.e., $\sigma_m = 0$). The DF of the cNFW halo, shown in the left-hand panels, is close to a pure exponential and shows little to no evolution over the period of 10 Gyr shown here. This is consistent with expectations. After all, the cNFW halo was created by evolving a (truncated) NFW halo with strong self-interactions ($\sigma_m = 25 \text{ cm}^2 \text{ g}^{-1}$) until the original cusp was transformed into an isothermal core, which has a DF of the form $f(E) \propto \exp(-E/\sigma^2)$. As mentioned in Section 3.1, the characteristic collision time for the cNFW halo is $t_0 = 1.91 \text{ Gyr} (\sigma_m / \text{cm}^2 \text{ g}^{-1})^{-1}$, and core collapse takes on the order of 400 collision times (see top-left panel of Fig. 2). Hence, for the largest cross section considered here, $\sigma_m = 3 \text{ cm}^2 \text{ g}^{-1}$, core collapse takes over 250 Gyr. Hence, it is not surprising that the DFs show no sign of evolution over the duration of 10 Gyr shown here.

Both the ABG[0.1] and ABG[0.5] halos have initial DFs that do not decrease monotonically with E , but rather reveal a pronounced dip near $E \sim -0.7$, which results in an inflection and thus a region in phase space where $df/dE > 0$, in violation of Antonov’s stability criterion. Consequently, as previously demonstrated in Dattathri et al. (2025b, see also Section 5 above), both systems are unstable to the dipole instability. When the dipole mode develops in strength, non-linear effects such as orbit trapping become important. As particles undergo separatrix crossing and become trapped into librating orbits, their actions undergo permanent changes, causing irreversible evolution of the DF. This evolution is evident in the top-right panel, which shows how, in the case of CDM, the dipole mode slowly flattens the dip in the DF of the ABG[0.5] halo. By $t = 10 \text{ Gyr}$, most of the dip in the original DF has disappeared (see Dattathri et al. 2025b, for a detailed discussion). In the case of the ABG[0.1] halo (top middle panel), the dipole mode

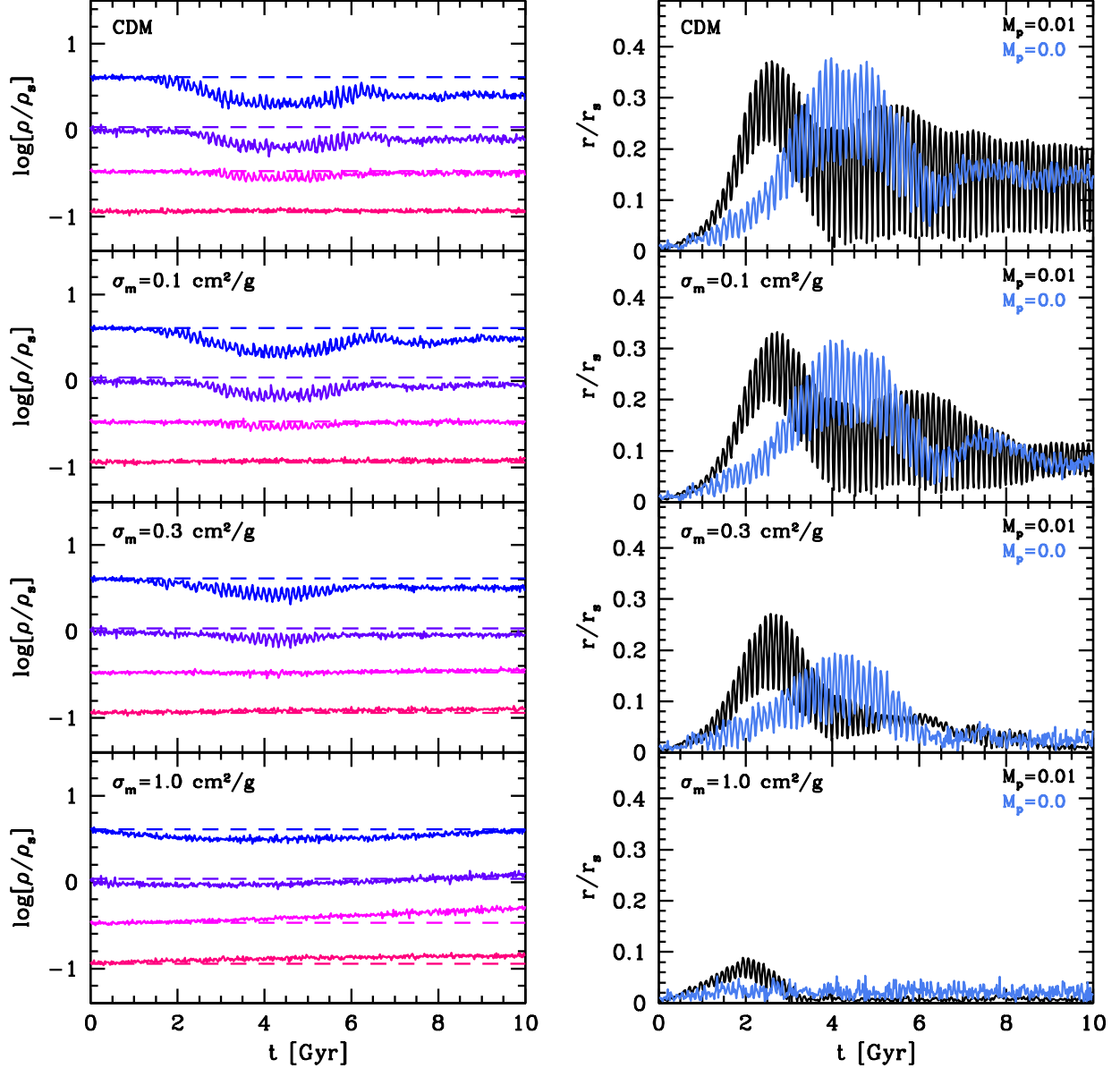


Figure 11. *Left-hand panels:* The evolution of the density of the ABG[0.5] halo in simulations without any perturber but with different values for σ_m (different rows), as indicated. In each panel, the different colors correspond to the densities measured at different radii; $\log[r/r_s] = -1.25$ (blue), -1.0 (violet), -0.75 (magenta), and -0.25 (red). The dashed, horizontal lines of corresponding color indicate the values of the density at $t = 0$ and are shown to highlight by how much the densities evolve. The strong evolution in the central densities in the case of CDM (top panel) is due to the dipole instability. Note how the instability becomes weaker with increasing σ_m ; for $\sigma_m = 1.0 \text{ cm}^2 \text{ g}^{-1}$ (red curves), the halo starts to undergo core collapse before the dipole instability can develop. *Right-hand panels:* the colored lines indicate the temporal evolution of the cusp in the four simulations shown in the left-hand panels. Note how, in the case of CDM (top panel), the cusp starts moving outward along a highly eccentric orbit, after which it settles at $r \sim 0.15 r_s$. As discussed in the text, this reflects the cusp being dislodged, and set in motion, by the dipole mode. This effect becomes weaker with increasing σ_m (going from top to bottom). The black lines in each panel indicate the halo-centric radii of a massive perturber ($M_p = 0.01$) which at $t = 0$ is positioned at rest at the CoM of the system. Note how, for small enough σ_m , the perturber moves outward (dynamical buoyancy). The motion is similar to that of the cusp in the simulations *without* perturber, indicating that dynamical buoyancy is a direct manifestation of the dipole instability. See text for a detailed discussion.

is significantly weaker and takes longer to become nonlinear. This explains why after 10 Gyr the DF has evolved very little, with only a slight reduction in the strength of the inflection.

In the case of SIDM, the self-interactions cause diffusion in phase space that eliminates the local suppression in the DF on a time scale that, as expected, decreases with increasing σ_m . In both the ABG[0.1] and ABG[0.5] halos, a cross section of $1 \text{ cm}^2 \text{ g}^{-1}$ suffices to completely remove the inflection in under 10 Gyr. Note that the interactions drive the distribution function towards isothermality, which explains why $f(E)$ takes on an exponential form at the most negative E (that is,

in the strongly bound central region).

6.2. Reduced core stalling in SIDM

In order to demonstrate how core stalling is impacted by the detailed shape of the host halo's DF, we focus on two representative simulations that follow the evolution of a perturber of mass $M_p = 0.001$, starting on a circular orbit at $r = 2.5 r_s$ in the ABG[0.1] halo. One simulation assumes CDM ($\sigma_m = 0$), while the other assumes SIDM with $\sigma_m = 1.0 \text{ cm}^2 \text{ g}^{-1}$. The radial evolution of the perturber in both cases is shown in the top left panel of Fig. 8 as the black and violet curves, respectively.

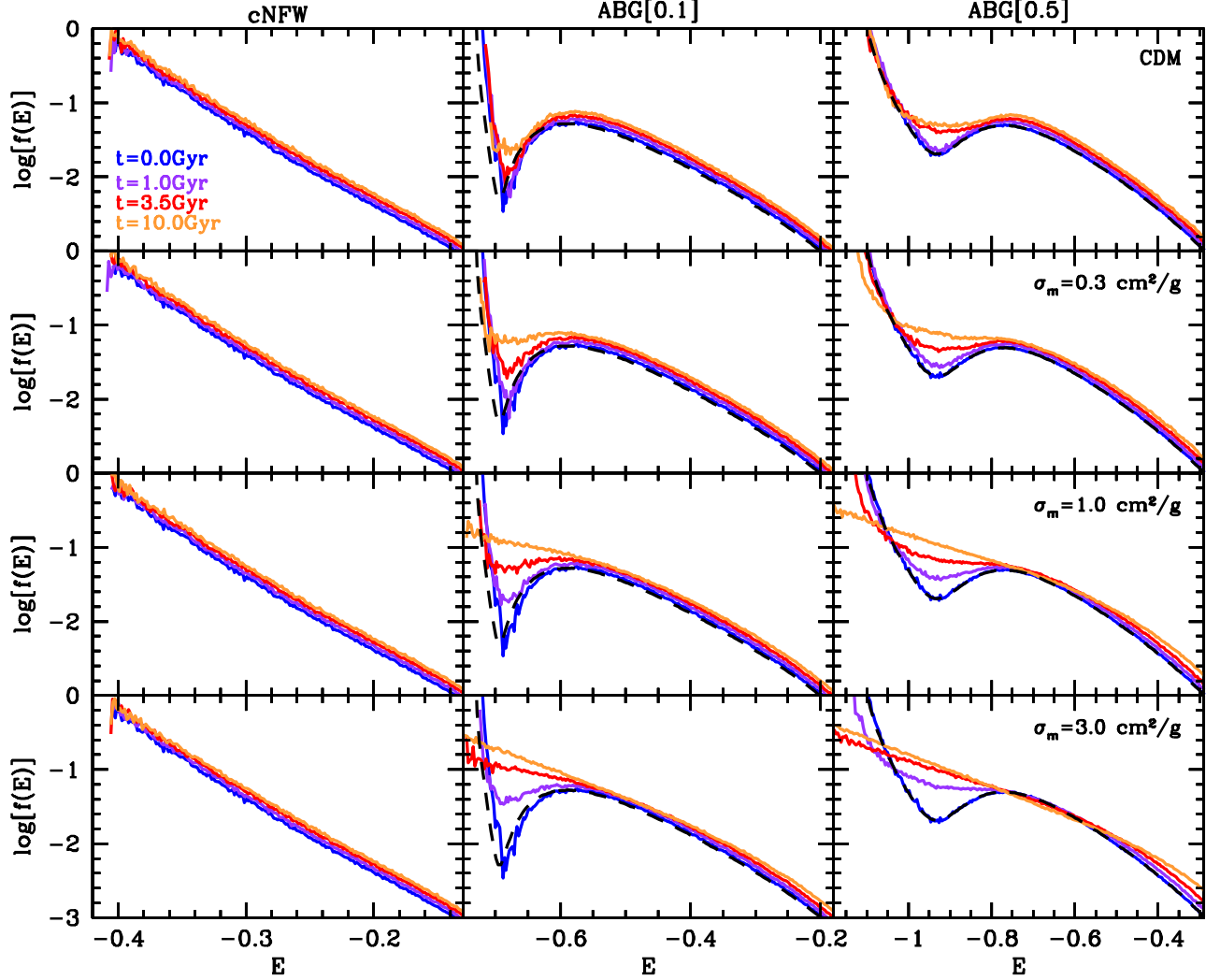


Figure 12. Evolution in the DFs for the three different cored systems studied here; from left to right, these are the cNFW halo, the ABG[0.1] halo, and the ABG[0.5] halo. Different colored lines correspond to different times, as indicated. Different rows correspond to different collisional cross sections, as indicated in the right-hand panels. Top panels correspond to CDM (i.e., $\sigma_m = 0$). The black dashed lines in the two right-most columns indicate the analytical DFs at $t = 0$ used to set up the ICs of the halos, and are shown for comparison.

In the case of CDM, the perturber sinks to $r = r_{\text{stall}} = 0.65r_s$ in ~ 20 Gyr, at which point the radial infall comes to an abrupt end, and the perturber stalls out at r_{stall} indefinitely. In the case of SIDM, the evolution is virtually identical for the first 20 Gyr; however, rather than stalling at r_{stall} , the perturber continues to sink inward, reaching $r \approx 0.2$ by $t = 26$ Gyr.

Fig. 13 shows the DFs of the ABG[0.1] halo in these two simulations at four different epochs. The top and bottom panels correspond to the CDM and SIDM simulations, respectively. The initial DF, which is identical for the CDM and SIDM cases, is shown in blue in both panels. As already shown in Fig. 12, this system has a pronounced inflection in its DF at $E \approx -0.7$, rendering it unstable to the dipole mode. However, as shown above, the instability takes a very long time to develop and therefore only has a marginal impact on the system’s evolution over the period considered here. The violet, red, and orange curves show the system’s DFs at $t = 16$ Gyr, $t = 20$ Gyr, and $t = 26$ Gyr, respectively, while the downward arrows of the same color indicate the orbital energy of the perturber at those corresponding epochs. Note how in the case of CDM, the red arrow indicates that at $t = 20$ Gyr, when stalling starts, the DF at the orbital energy of the perturber has a vanishing gradient,

that is $df/dE = 0$. This explains why the system experiences core stalling (see Section 2). Note that the orbital energy of the perturber at $t = 26$ Gyr, indicated by the orange arrow, is almost identical to that of 6 Gyr earlier (red arrow), indicating that the perturber has indeed stalled its inward motion. Note also that from $t = 0$ to $t = 26$ Gyr the depth of the dip in the initial DF has reduced significantly. As discussed in D25, this is the combined effect of the dipole instability and the perturber modifying the DF of the host system.

In the case of SIDM with $\sigma_m = 1 \text{ cm}^2 \text{ g}^{-1}$, shown in the lower panel, the inflection in the DF has been completely erased, already at $t = 16$ Gyr, as a consequence of the self-interactions. By the time the perturber reaches r_{stall} at $t = 20$ Gyr, the gradient in the DF at the orbital energy of the perturber has become negative, as opposed to zero in the case of CDM, which explains why the perturber continues to sink inwards without stalling. By $t = 26$ Gyr, the perturber has sunk almost entirely to the center of the system, which has started to undergo core collapse; this is evident from the fact that the DF toward the center (at small E) has become a featureless exponential consistent with an isothermal core. Note also that the DF at large E becomes somewhat shallower with time.

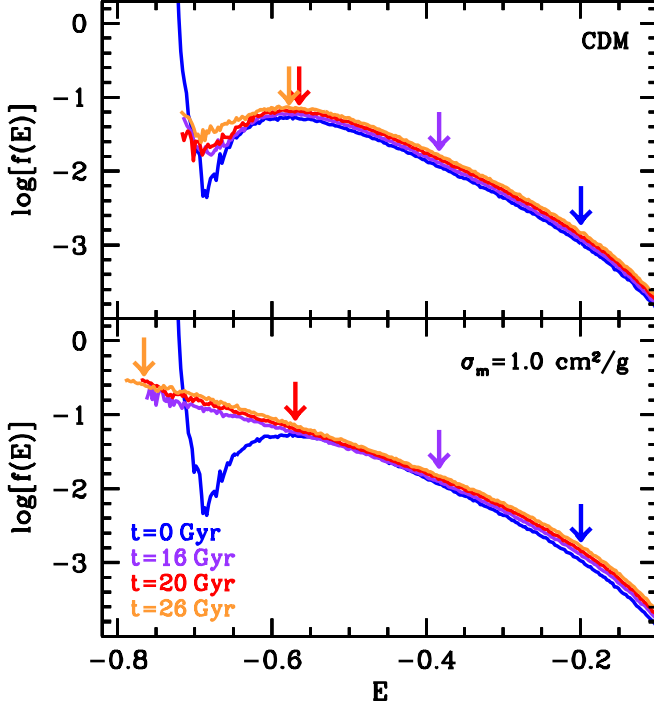


Figure 13. Evolution in the DFs of the ABG[0.1] halo in the presence of a massive perturber with $M_p = 0.001$ in the case of CDM (top panel) and SIDM with $\sigma_m = 1.0 \text{ cm}^2 \text{ g}^{-1}$ (bottom panel). Results are shown at four different epochs: $t = 0$ (blue), $t = 16$ Gyr (violet), $t = 20$ Gyr (red) and $t = 26$ Gyr (orange). Downward arrows in matching color indicate the orbital energy of the perturber at the corresponding times. For comparison, the radial evolution of the perturber in both cases is shown by the black and violet lines in the top-left panel of Fig. 8.

Since the LBK torque is directly proportional to the local gradient in the DF, this explains why, in the case of larger cross sections, the perturber takes longer to sink in (see top-left panel of Fig. 8).

Finally, to really drive home the point, Fig. 14 shows the results for a perturber of mass $M_p = 0.01$ sinking inside the cNFW halo. Recall, from Fig. 12, that this halo starts out with a DF that is almost a perfect power law, without any bumps or inflection points. Hence, naive predictions based on the DF gradient are that the perturber should sink to the center without interruption. However, as shown in Fig. 4, this is not the case. In particular, in the case of CDM the perturber stalls at a stalling radius $r_{\text{stall}} \approx 0.25 r_s$. However, in the case of SIDM, as long as the cross section is large enough, the perturber sinks through r_{stall} and sinks rapidly toward the center of the system. The top panels of Fig. 14 show the evolution of the halocentric distance of the perturber in the case of CDM (upper left) and SIDM with $\sigma_m = 25 \text{ cm}^2 \text{ g}^{-1}$ (upper right). These are identical to the corresponding curves in the upper panels of Fig. 4, except that here we zoom in on the period between 10 and 16.5 Gyr. The solid colored dots indicate specific epochs at which the distribution functions $f(E)$ are shown in the 3×3 grid of panels at the bottom of Fig. 14 (in corresponding color). Note that we have offset the DF for the CDM case by -1.2 dex for the sake of clarity. The downward pointing arrows indicate the instantaneous orbital energy of the perturber. Initially, up until $t \sim 13$ Gyr the DF of the host remains largely unchanged with respect to the initial power law, indicated by the black dashed line). However, once the perturber comes close to r_{stall} , it causes a strong distortion of the host (cf. left-hand

panels of Fig. 6) that creates a suppression in the DF at $E \lesssim -0.42$. This results in a bump where $df/dE = 0$, and this in turn causes the perturber to stall its inward motion. In the case of CDM, the suppression remains, although the DF continues to undergo small changes due to interactions with the perturber. This in turn causes the perturber to slowly drift inwards over the period from 13 to 16.5 Gyr. In the case of SIDM, the self-interactions oppose the creation of an inflection in the DF by continuously driving the system towards isothermality. With a large cross section of $25 \text{ cm}^2 \text{ g}^{-1}$, the self-interactions are able to re-establish a negative gradient in the DF shortly after the perturber introduced a suppression, which is why the perturber continues to sink inward with only a minor ‘hick-up’ in its sinking rate around $t \sim 13.2$ Gyr. These results make the important point that dynamical friction on a perturber acts to modify the DF of the host (see discussion in Section 2.1), which can introduce features in the DF that cause stalling and/or buoyancy. Hence, predicting a priori whether a perturber will stall or not is difficult without some knowledge of how its inward journey will affect the DF.

6.3. Dynamical friction as a catalyst of core collapse

As shown in Section 4, when a massive perturber sinks toward the center of its SIDM host halo, it causes drastically accelerated core collapse. For example, the core collapse time of the cNFW halo in the absence of a massive perturber is $t_{\text{cc}} \approx 400 t_0$ (cf. Fig. 2) which amounts to ~ 255 Gyr (30 Gyr) for a cross section of $\sigma_m = 3 \text{ cm}^2 \text{ g}^{-1}$ ($25 \text{ cm}^2 \text{ g}^{-1}$). However, as soon as a perturber of mass $M_p = 0.01$ (that is, one percent of the halo virial mass) reaches the core radius, the core collapses in only ~ 6 Gyr (1 Gyr).

The main cause of this accelerated collapse is adiabatic contraction. As the perturber sinks inward, it pulls in dark matter due to its gravity. This boosts the density and velocity dispersion of the core, which in turn boosts the interaction rate $\Gamma = \rho \sigma_m v_{\text{rel}}$, with v_{rel} the relative velocity among the dark matter particles. This, in turn, increases the outward heat conduction, which accelerates the gravothermal instability. This is counteracted, though, by the fact that the perturber is also injecting heat into the core as a consequence of the dynamical friction it experiences. Heat injection causes the core to expand and cool due to its negative heat capacity, which thus acts in the opposite direction. Hence, the net outcome is a tug-of-war between adiabatic contraction, accelerating core collapse, and dynamical heating, delaying core collapse.

The bottom panel of Fig. 15 compares the evolution of the central density of the SIDM ABG[0.1] halo with $\sigma_m = 10 \text{ cm}^2 \text{ g}^{-1}$ in the absence of a perturber (violet curve) and in the case with $M_p = 0.01$ (blue curve). The top panel shows the evolution of the halo-centric distance of the perturber as it sinks towards the center. The gray shaded region indicates the period between the perturber crossing the stalling radius (where stalling occurs in the case of CDM) and the perturber reaching the center. Note how the perturber accelerates core collapse during this period. Once it has reached the center, core collapse continues, but with a rate that is reminiscent of core collapse in the absence of a perturber. The red curve shows what happens when the perturber is instantaneously removed from the simulation at $t = 3.7$ Gyr, when it has reached $r \approx 0.06 r_s$. As is apparent, at that point in time it has already started to accelerate the collapse of the core. However, upon its instantaneous removal, the system reacts by expanding to the impulsive change in the potential, and the central density

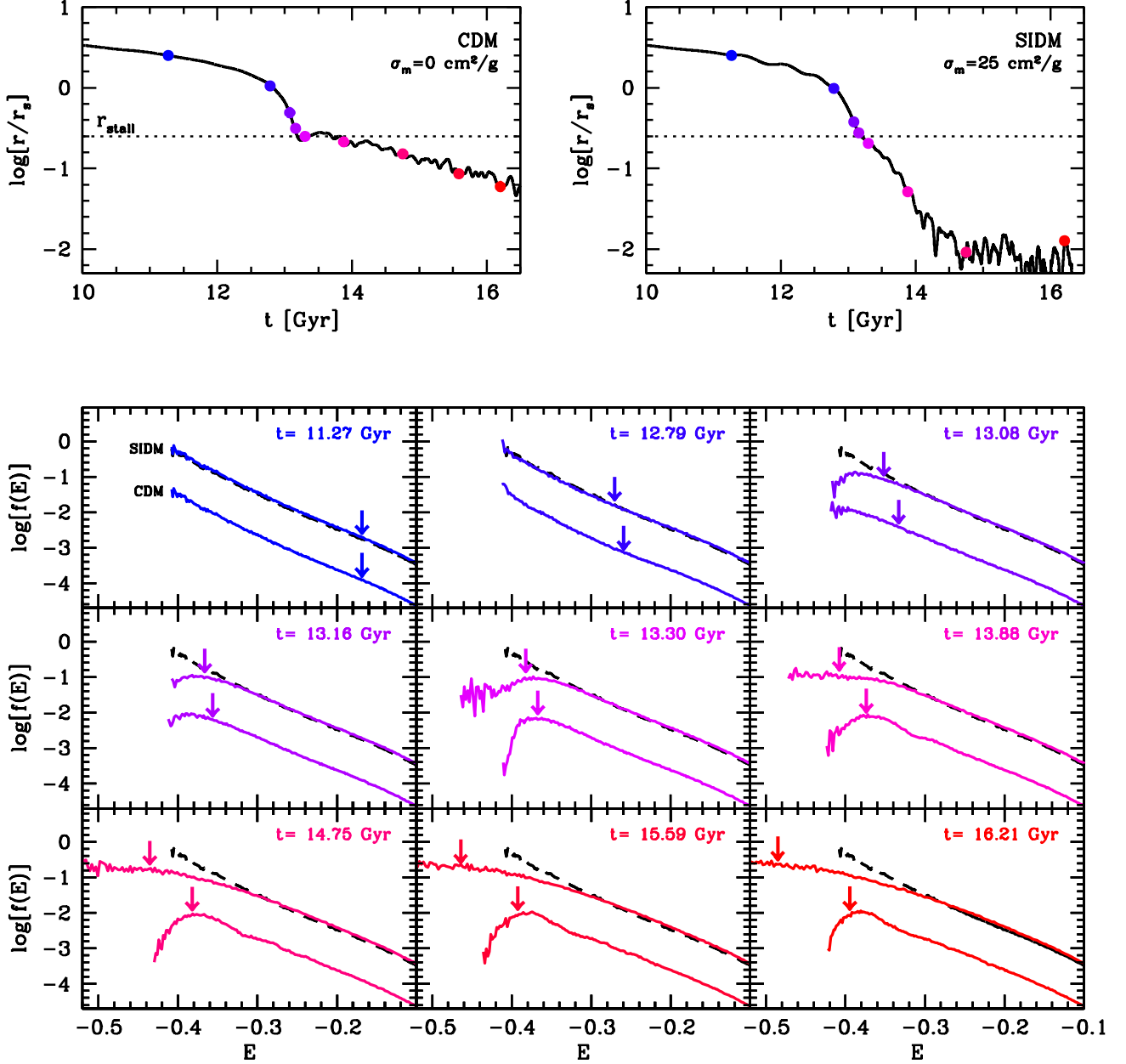


Figure 14. The impact of a perturber of mass $M_p = 0.01$ sinking inside the cNFW halo. Upper panels show the halocentric radius of the perturber as a function of time in the case of CDM (right) and SIDM (left). The lower 3×3 panels show the corresponding DFs at 9 different epochs, with the downward arrow indicating the energy of the perturber at the corresponding time. The DFs for the CDM case are offset by -1.2 dex for clarity. As in Fig. 13, the downward pointing arrows indicate the instantaneous orbital energy of the perturber. The black dashed lines indicate the initial DF at $t = 0$ and is shown for comparison. Note how in the case of CDM, the perturber introduces an inflection in the DF, resulting in a bump where df/dE vanishes, which in turn causes the perturber to stall. In the case of SIDM, the self-interactions oppose this by driving the DF towards an isothermal; consequently df/dE remains negative and the perturber sinks all the way to the center.

drops down again. Subsequently, the system continues its core collapse, which is now somewhat *delayed* compared to the case without a perturber. These results show that *core collapse is only accelerated while the perturber sinks through it*. During this period, the core collapse is aided by the gravitational contraction induced by the sinking perturber. When the perturber is removed, the acceleration ceases immediately, indicating that the acceleration is catalyzed, rather than triggered, by the sinking-in of the perturber.

7. CONCLUSIONS

Over the years, it has become clear that gravitational N -body systems with a core, roughly defined as a central region with a logarithmic density gradient $-0.5 \lesssim d \log \rho / d \log r \leq 0$, are subject to intriguing dynamical phenomena. As first shown by Read et al. (2006), using numerical simulations, dynamical friction can cease to operate inside cores, giving rise to core stalling. In later work, Cole et al. (2012) showed, again using simulations, that a massive object positioned within such a core experiences a dynamical buoyancy that causes it to move outward. Finally, Dattathri et al. (2025b) demonstrated that cores can be unstable to a rotating dipole mode.

Recently, D25 presented a unified picture of these unex-

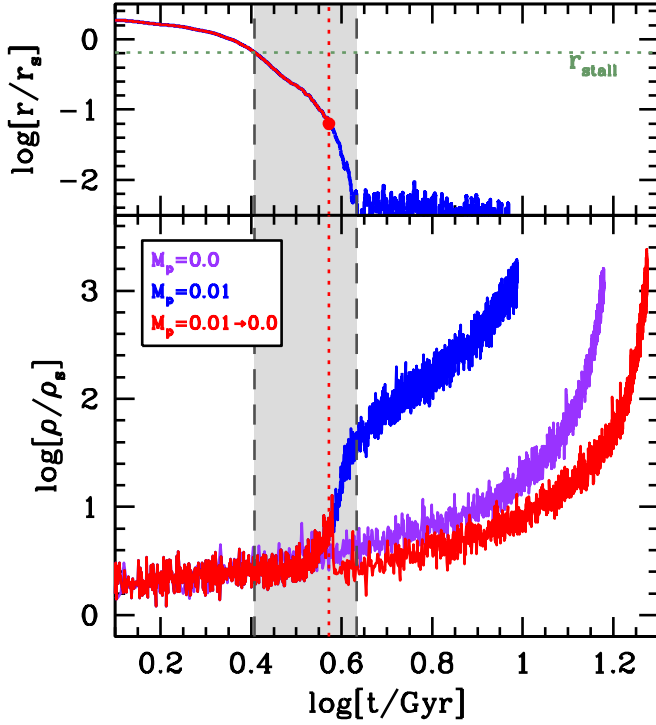


Figure 15. The bottom panel shows the evolution of the central density of the ABG[0.1] halo in SIDM simulations with $\sigma_m = 10 \text{ cm}^2 \text{ g}^{-1}$. The violet curve shows how the system undergoes core collapse in ~ 15 Gyr in the absence of any perturber. The blue curve shows the results in the presence of a massive perturber of mass $M_p = 0.01$, which at $t = 0$ starts out on a circular orbit at $r = 2.5r_s$. The top panel shows how the perturber sinks in due to dynamical friction. The gray-shaded region marks the period between the time when the perturber crosses the stalling radius r_{stall} , where the perturber *would* have stalled in the case of CDM (see Fig. 8), and the time when it reaches the center of the halo. Note that during this period the core collapse is accelerated compared to the case without perturber. Finally, the red curves show what happens if the perturber is instantaneously removed from the system at $t = 3.7$ Gyr. Its removal causes the core to expand, after which core collapse continues. This illustrates that core collapse is only accelerated while the perturber is sinking through it.

pected dynamical phenomena grounded in kinetic theory. In particular, core stalling, dynamical buoyancy and the dipole instability are all shown to be related to specific features (plateaus and inflections) in the phase-space DF. Using high resolution numerical simulations they showed that the sign of the torque on the perturber is set by the sign of df/dE at the orbital energy of the perturber, here called $(\nabla f)_{\text{CR}}$. Dynamical friction and buoyancy correspond to $(\nabla f)_{\text{CR}} < 0$ and > 0 , respectively, while a vanishing $(\nabla f)_{\text{CR}}$ gives rise to core stalling. This is exactly what is predicted based on the expression for the LBK torque of equation (1) as long as corotation is the dominant resonance.

An important corollary is that, since the DF evolves as the perturber sinks inward, it is difficult to predict a priori whether or not a perturber will stall in a certain system. For example, a massive perturber may start to sink inside a system that has $df/dE < 0$ everywhere, but modify the DF as it sinks in to such an extent that it develops an inflection, causing subsequent stalling and or buoyancy. Similarly, a perturber may stall at a plateau or bump where $(\nabla f)_{\text{CR}} \approx 0$, modify the DF by nonlinear effects such as particle trapping, and then re-ignite friction (or buoyancy) when the local gradient in the DF has become nonzero. Explicit examples of such behavior are shown in Figs. 4, 7, and 8 and in D25.

All previous studies of core dynamics have focused on colli-

sionless systems, either galaxies or CDM halos. In this paper we studied whether and how core dynamics differ in the case of SIDM, which is known to produce cored dark matter halos. Our main conclusions are:

- Self-interactions act to suppress core stalling, dynamical buoyancy, and the core dipole instability. The reason is that the self-interactions cause a diffusion in phase space that acts to erase small-scale features in the DF and drive the system’s DF toward a featureless exponential, $f(E) \propto \exp(-E/\sigma^2)$, characteristic of an isothermal system. Since the corresponding df/dE is always negative, cores in SIDM halos are not expected to be subject to any of the core dynamical processes discussed here. Obviously, the extent to which self-interactions can suppress stalling and buoyancy depends on the cross section, with larger cross sections resulting in stronger and more rapid suppression.
- As the massive perturber sinks inside the core of an SIDM halo, it catalyzes core collapse by gravitationally contracting the core. This can be extremely efficient, shortening the time to reach core collapse by over an order in magnitude.

These findings have a number of important implications. Most importantly, the fact that the dynamics in CDM and SIDM cores can be very different opens a potential avenue for constraining the nature of dark matter. Dipole instabilities, buoyancy, and stalling have a number of observational consequences. Among others, they give rise to lopsidedness in the central regions of galaxies and to nuclear star clusters and AGN that are offset from the galaxy’s center inferred from the outer isophotes. Interestingly, such offsets are not uncommon (e.g., Binggeli et al. 2000; Côté et al. 2006; Shen et al. 2019; Reines et al. 2020; Sargent et al. 2022; Mezcua & Domínguez Sánchez 2024). Taken at face value, this seems to favor CDM over SIDM. However, it is important to account for the fact that there are various other explanations for offset NSCs or AGN, including, but not restricted to, recent accretion (time since accretion is less than the dynamical friction time scale), galaxy mergers, recoil kicks following a SMBH merger (Campanelli et al. 2007), or dynamical heating by substructure (Boldrini et al. 2020). Nevertheless, we are optimistic that future studies with better statistics, based on larger data sets and simulation suites, will be able to put meaningful constraints on the self-interaction cross sections for the dark matter.

Related to this, core stalling and buoyancy can potentially delay or prevent the merger of two or more supermassive black holes, which has implications for the total gravitational wave signal to be measured by LISA and/or pulsar timing arrays, and may also impede the merging of globular clusters in dwarf galaxies. The latter is one of the leading scenarios for the formation of nuclear star clusters (Tremaine et al. 1975; Capuzzo-Dolcetta & Miocchi 2008; Arca-Sedda & Capuzzo-Dolcetta 2016). Interestingly, our results suggest that these processes will be more efficient in SIDM than in CDM, potentially opening another avenue for constraining σ_m .

Our findings also impact the conclusions of several studies that invoked or relied on core stalling. For example, Modak et al. (2023) argued that the formation of nuclear star clusters via the merging of globular clusters requires the presence of a cusp, since cores cause core stalling. In light of the results presented here, this conclusion is subject to the assumption that dark matter self-interactions are negligible. Similarly, core stalling and buoyancy have been invoked as a solution to the timing problem of globular clusters in dwarf spheroidals

(Goerdt et al. 2006; Cole et al. 2012; Boldrini et al. 2019; Sánchez-Salcedo & Lora 2022), or the presence of multiple nuclei in the centers of galaxies (Bonfini & Graham 2016; Meadows et al. 2020; Nasim et al. 2021). As shown here, these results depend on the nature of dark matter; in the case of SIDM one can only invoke these processes to solve the timing problem if the self-interaction cross sections are sufficiently small. Hence, future studies tailored to individual systems such as the Fornax galaxy, which is one of the galaxies for which the timing problem is the most severe, again have the potential to constrain the SIDM self-interaction cross section.

Finally, the fact that massive perturbers sinking inside SIDM cores catalyze core collapse may have important ramifications. For example, Slone et al. (2023) argue that the observed central densities of Milky Way satellite galaxies such as Draco and Ursa Minor can be used to constrain the self-interaction cross sections of SIDM models. In particular, they argue that one can rule out a significant part of SIDM parameter space as it predicts that the dark matter haloes of the dwarf spheroidals in question would have produced overly large cores that are inconsistent with their high observed central densities. However, this did not account for the fact that the accretion of a massive object such as a supermassive black hole during the dwarf spheroidal's hierarchical formation history might have drastically shortened the core collapse time. Hence, it is crucial to first rule out that the nuclei of the dwarf spheroidals used for this test are devoid of any central massive objects.

ACKNOWLEDGEMENTS

In loving memory of Avishai Dekel, whose kindness, encouragement, and physical insight have been a source of inspiration throughout our careers. We are grateful to Uddipan Banik, Kimberly Body, Annika Peter, Oren Slone, Charlie Mace, Fangzhou Jiang, and Martin Weinberg for useful discussions, and to Anne Lisa Vari, Jean-Baptiste Fouvy, Matthew Kunz, and Jonathan Squire for organizing the workshop “Interconnections between the Physics of Plasmas and Self-gravitating Systems” at the KITP in Santa Barbara where part of this work was performed. The KITP Santa Barbara is supported in part by the National Science Foundation (NSF) under Grant No. NSF PHY-2309135. This work has been supported by the NSF through grant AST 2407063.

REFERENCES

- Adhikari S., et al., 2022, *arXiv e-prints*, p. [arXiv:2207.10638](#)
- Agnello A., Evans N. W., 2012, *ApJL*, **754**, L39
- Antonov V. A., 1962, Solution of the problem of stability of stellar system
- Emden's density law and the spherical distribution of velocities
- Arca-Sedda M., Capuzzo-Dolcetta R., 2016, *MNRAS*, **461**, 4335
- Balberg S., Shapiro S. L., Inagaki S., 2002, *ApJ*, **568**, 475
- Banik U., Bhattacharjee A., 2025, *arXiv e-prints*, p. [arXiv:2506.02104](#)
- Banik U., van den Bosch F. C., 2022, *ApJ*, **926**, 215
- Barausse E., Dvorkin I., Tremmel M., Volonteri M., Bonetti M., 2020, *ApJ*, **904**, 16
- Barnes J., Hut P., 1986, *Nature*, **324**, 446
- Begelman M. C., Blandford R. D., Rees M. J., 1980, *Nature*, **287**, 307
- Binggeli B., Barazza F., Jerjen H., 2000, *A&A*, **359**, 447
- Binney J., Tremaine S., 2008, *Galactic Dynamics: Second Edition*. Princeton University Press
- Blumenthal G. R., Faber S. M., Flores R., Primack J. R., 1986, *ApJ*, **301**, 27
- Boldrini P., Mohayaee R., Silk J., 2019, *MNRAS*, **485**, 2546
- Boldrini P., Mohayaee R., Silk J., 2020, *MNRAS*, **495**, L12
- Bonfini P., Graham A. W., 2016, *ApJ*, **829**, 81
- Bryan G. L., Norman M. L., 1998, *ApJ*, **495**, 80
- Burkert A., 2000, *ApJL*, **534**, L143
- Byun Y. I., et al., 1996, *AJ*, **111**, 1889
- Campanelli M., Lousto C., Zlochower Y., Merritt D., 2007, *ApJL*, **659**, L5
- Capuzzo-Dolcetta R., Miocchi P., 2008, *MNRAS*, **388**, L69
- Chandrasekhar S., 1943, *ApJ*, **97**, 255
- Chavanis P.-H., 2024, *European Physical Journal Plus*, **139**, 51
- Chiang B. T., van den Bosch F. C., Schive H.-Y., 2025, *MNRAS*, **544**, 36
- Chiba R., 2023, *MNRAS*, **525**, 3576
- Chiba R., Schönrich R., 2021, *arXiv e-prints*, p. [arXiv:2109.10910](#)
- Cole D. R., Dehnen W., Wilkinson M. I., 2011, *MNRAS*, **416**, 1118
- Cole D. R., Dehnen W., Read J. I., Wilkinson M. I., 2012, *MNRAS*, **426**, 601
- Côté P., et al., 2006, *ApJS*, **165**, 57
- Dattathri S., van den Bosch F. C., Banik U., Weinberg M., Natarajan P., Li Z., Dekel A., 2025a, *arXiv e-prints*, p. [arXiv:2511.11804](#)
- Dattathri S., van den Bosch F. C., Weinberg M. D., Banik U., 2025b, *ApJ*, **993**, 150
- Dehnen W., 2002, *Journal of Computational Physics*, **179**, 27
- Dekel A., Silk J., 1986, *ApJ*, **303**, 39
- Dekel A., et al., 2021, *MNRAS*, **508**, 999
- Di Cintio A., Brook C. B., Macciò A. V., Stinson G. S., Knebe A., Dutton A. A., Wadsley J., 2014, *MNRAS*, **437**, 415
- Doremus J. P., Baumann G., Feix M. R., 1973, *A&A*, **29**, 401
- Drakos N. E., Taylor J. E., Benson A. J., 2017, *MNRAS*, **468**, 2345
- Dutta Chowdhury D., van den Bosch F. C., van Dokkum P., 2019, *ApJ*, **877**, 133
- Eddington A. S., 1916, *MNRAS*, **76**, 572
- El-Zant A., Shlosman I., Hoffman Y., 2001, *ApJ*, **560**, 636
- El-Zant A. A., Freundlich J., Combes F., 2016, *MNRAS*, **461**, 1745
- Elbert O. D., Bullock J. S., Kaplinghat M., Garrison-Kimmel S., Graus A. S., Rocha M., 2018, *ApJ*, **853**, 109
- Essig R., McDermott S. D., Yu H.-B., Zhong Y.-M., 2019, *Phys. Rev. Lett.*, **123**, 121102
- Ferrarese L., van den Bosch F. C., Ford H. C., Jaffe W., O'Connell R. W., 1994, *AJ*, **108**, 1598
- Flores R. A., Primack J. R., 1994, *ApJL*, **427**, L1
- Fouvy J.-B., 2017, *Secular Evolution of Self-Gravitating Systems Over Cosmic Age*, doi:10.1007/978-3-319-66760-7
- Freundlich J., Dekel A., Jiang F., Ishai G., Cornuault N., Lapiner S., Dutton A. A., Macciò A. V., 2020, *MNRAS*, **491**, 4523
- Gebhardt K., et al., 1996, *AJ*, **112**, 105
- Genzel R., et al., 2020, *ApJ*, **902**, 98
- Goerdt T., Moore B., Read J. I., Stadel J., Zemp M., 2006, *MNRAS*, **368**, 1073
- Goerdt T., Moore B., Read J. I., Stadel J., 2010, *ApJ*, **725**, 1707
- Hamilton C., 2024, *MNRAS*, **528**, 5286
- Hamilton C., Fouvy J.-B., 2024, *Physics of Plasmas*, **31**, 120901
- Heggie D. C., Breen P. G., Varri A. L., 2020, *MNRAS*, **492**, 6019
- Henon M., 1973, *A&A*, **24**, 229
- Hernquist L., 1990, *ApJ*, **356**, 359
- Inoue S., 2009, *MNRAS*, **397**, 709
- Inoue S., 2011, *MNRAS*, **416**, 1181
- Jaffe W., 1983, *MNRAS*, **202**, 995
- Kaur K., Sridhar S., 2018, *ApJ*, **868**, 134
- Kaur K., Stone N. C., 2022, *MNRAS*, **515**, 407
- King I. R., 1966, *AJ*, **71**, 64
- Kochanek C. S., White M., 2000, *ApJ*, **543**, 514
- Koda J., Shapiro P. R., 2011, *MNRAS*, **415**, 1125
- Kuzio de Naray R., McGaugh S. S., de Blok W. J. G., 2008, *ApJ*, **676**, 920
- Li Z., Dekel A., Mandelker N., Freundlich J., François T. L., 2023, *MNRAS*, **518**, 5356
- Lifshitz E. M., Pitaevskii L. P., 1981, *Physical Kinetics*. Pergamon Press
- Loeb A., Weiner N., 2011, *Phys. Rev. Lett.*, **106**, 171302
- Lynden-Bell D., Eggleton P. P., 1980, *MNRAS*, **191**, 483
- Lynden-Bell D., Kalnajs A. J., 1972, *MNRAS*, **157**, 1
- Lynden-Bell D., Wood R., 1968, *MNRAS*, **138**, 495
- Mace C., Zeng Z. C., Peter A. H. G., Du X., Yang S., Benson A., Vogelsberger M., 2024, *Phys. Rev. D*, **110**, 123024
- Mace C., Yang S., Carton Zeng Z., Peter A. H. G., Du X., Benson A., 2025, *arXiv e-prints*, p. [arXiv:2504.13004](#)
- Martizzi D., Teyssier R., Moore B., 2013, *MNRAS*, **432**, 1947
- Meadows N., Navarro J. F., Santos Santos I., Benítez-Llambay A., Frenk C., 2020, *MNRAS*, **491**, 3336
- Mezcua M., Domínguez Sánchez H., 2024, *MNRAS*, **528**, 5252
- Michie R. W., 1963, *MNRAS*, **125**, 127
- Mo H., van den Bosch F. C., White S., 2010, *Galaxy Formation and Evolution*. Cambridge University Press
- Modak S., Danieli S., Greene J. E., 2023, *ApJ*, **950**, 178
- Nasim I. T., Gualandris A., Read J. I., Antonini F., Dehnen W., Delorme M., 2021, *MNRAS*, **502**, 4794
- Navarro J. F., Frenk C. S., White S. D. M., 1997, *ApJ*, **490**, 493
- Nishikawa H., Boddy K. K., Kaplinghat M., 2020, *Phys. Rev. D*, **101**, 063009
- Ogiya G., Nagai G., 2022, *MNRAS*, **514**, 555
- Oh S.-H., et al., 2015, *AJ*, **149**, 180
- Oman K. A., et al., 2015, *MNRAS*, **452**, 3650
- Palubski I., Slone O., Kaplinghat M., Lisanti M., Jiang F., 2024, *J. Cosmology Astropart. Phys.*, **2024**, 074
- Planck Collaboration et al., 2018, *arXiv e-prints*, p. [arXiv:1807.06209](#)
- Plummer H. C., 1911, *MNRAS*, **71**, 460
- Pollack J., Spergel D. N., Steinhardt P. J., 2015, *ApJ*, **804**, 131
- Pontzen A., Governato F., 2012, *MNRAS*, **421**, 3464
- Price S. H., et al., 2021, *ApJ*, **922**, 143
- Quinlan G. D., 1996, *New A*, **1**, 35

Read J. I., Goerdt T., Moore B., Pontzen A. P., Stadel J., Lake G., 2006, *MNRAS*, **373**, 1451
 Reif F., 1965, *Fundamentals of Statistical and Thermal Physics*
 Reines A. E., Condon J. J., Darling J., Greene J. E., 2020, *ApJ*, **888**, 36
 Rocha M., Peter A. H. G., Bullock J. S., Kaplinghat M., Garrison-Kimmel S., Oñorbe J., Moustakas L. A., 2013, *MNRAS*, **430**, 81
 Sánchez-Salcedo F. J., Lora V., 2022, *MNRAS*, **511**, 1860
 Sargent A. J., Johnson M. C., Reines A. E., Secrest N. J., van der Horst A. J., Cigan P. J., Darling J., Greene J. E., 2022, *ApJ*, **933**, 160
 Shen Y., Hwang H.-C., Zakamska N., Liu X., 2019, *ApJL*, **885**, L4
 Slone O., Jiang F., Lisanti M., Kaplinghat M., 2023, *Phys. Rev. D*, **107**, 043014
 Spergel D. N., Steinhardt P. J., 2000, *Phys. Rev. Lett.*, **84**, 3760
 Spitzer Jr. L., Shapiro S. L., 1972, *ApJ*, **173**, 529
 Teyssier R., Pontzen A., Dubois Y., Read J. I., 2013, *MNRAS*, **429**, 3068
 Tollet E., et al., 2016, *MNRAS*, **456**, 3542
 Tremaine S., Weinberg M. D., 1984, *MNRAS*, **209**, 729

Tremaine S. D., Ostriker J. P., Spitzer L. J., 1975, *ApJ*, **196**, 407
 Tulin S., Yu H.-B., 2018, *Phys. Rep.*, **730**, 1
 Volonteri M., et al., 2020, *MNRAS*, **498**, 2219
 Weinberg M. D., 1994a, *AJ*, **108**, 1398
 Weinberg M. D., 1994b, *ApJ*, **421**, 481
 Weinberg M. D., 2001, *MNRAS*, **328**, 311
 Weinberg M. D., 2023, *MNRAS*, **525**, 4962
 Weinberg M. D., Katz N., 2002, *ApJ*, **580**, 627
 Yang S., Du X., Zeng Z. C., Benson A., Jiang F., Nadler E. O., Peter A. H. G., 2023, *ApJ*, **946**, 47
 Zhao H., 1996, *MNRAS*, **278**, 488
 de Blok W. J. G., Bosma A., 2002, *A&A*, **385**, 816
 de Blok W. J. G., McGaugh S. S., Bosma A., Rubin V. C., 2001, *ApJL*, **552**, L23

APPENDIX

A. GRAVOTHERMAL FLUID EQUATIONS

Consider a spherical SIDM halo with a density profile $\rho(r, t)$ and an enclosed mass $M(r, t)$ at radius r and time t . Assume the halo to be isotropic and in quasi-static virial equilibrium with a one-dimensional velocity dispersion $v(r, t)$. Due to self-interactions, characterized by a cross section per unit mass σ_m , heat flows in the radial direction with a luminosity $L(r, t)$ through a spherical shell at radius r . The corresponding heat flux obeys Fourier's law of thermal conduction,

$$\frac{L}{4\pi r^2} = -\kappa \frac{\partial T}{\partial r}, \quad (\text{A.1})$$

with κ conductivity and $k_B T = m v^2$ a measure of the temperature of the particles of mass m . The temporal evolution of such a halo is a time-dependent diffusion problem that can be modeled as a gravothermal fluid (Lynden-Bell & Eggleton 1980; Balberg et al. 2002). In addition to the heat conduction equation, such a fluid obeys mass conservation,

$$\frac{\partial M}{\partial r} = 4\pi r^2 \rho, \quad (\text{A.2})$$

hydrostatic equilibrium,

$$\frac{\partial (\rho v^2)}{\partial r} = -\frac{GM\rho}{r^2}, \quad (\text{A.3})$$

where we have defined the pressure $P = \rho v^2$, and the first law of thermodynamics which can be written as

$$\frac{\partial L}{\partial r} = -4\pi \rho r^2 v^2 \left(\frac{\partial}{\partial t} \right)_M \ln(v^3/\rho), \quad (\text{A.4})$$

with $s = \ln(v^3/\rho)$ an effective entropy. Each time step, heat flows in the radial direction as dictated by equation (A.1), after which the system re-establishes hydrostatic equilibrium, while satisfying mass conservation and the first law of thermodynamics.

Strictly speaking, Fourier's law of thermal conduction is valid in the short mean-free path (SMFP) regime, where the Knudsen number $\text{Kn} \ll 1$. Here $\text{Kn} \equiv \lambda/H$ with $\lambda = (\rho\sigma_m)^{-1}$ the mean-free path and $H = \sqrt{v^2/(4\pi G\rho)}$ the gravitational Jeans length, which is a measure for the size of a virialized system. In the SMFP regime, Chapman-Enskog theory (Lifshitz & Pitaevskii 1981) dictates that the conductivity is given by

$$\kappa_{\text{SMFP}} = \frac{3}{2} a^{-1} b n k_B \frac{\lambda^2}{t_r}, \quad (\text{A.5})$$

(Balberg et al. 2002; Koda & Shapiro 2011), where $n = \rho/m$ is the number density of particles, k_B is the Boltzmann constant, $b = 25\sqrt{\pi}/32$ is the effective impact parameter, and $t_r \equiv \lambda/(av)$ is the collision time (cf. equation [5] in the main text) with $a = 4/\sqrt{\pi}$ a coefficient relevant for hard-sphere scattering of particles with a Maxwell-Boltzmann distribution (Reif 1965). In the long mean-free path (LMFP) regime, which is more common for SIDM halos, no formal expression for the conductivity exists, but Lynden-Bell & Eggleton (1980) have shown that an empirical relation of the form

$$\kappa_{\text{LMFP}} = \frac{3}{2} C n k_B \frac{H^2}{t_r}, \quad (\text{A.6})$$

with C a free calibration parameter of order unity (see below), can adequately describe the gravothermal collapse of globular clusters. Following Balberg et al. (2002), we interpolate between these two regimes using $\kappa^{-1} = \kappa_{\text{SMFP}}^{-1} + \kappa_{\text{LMFP}}^{-1}$.

Given appropriate boundary conditions, equations (A.1)-(A.4) can be solved given a set of initial radial profiles. Typically, these are the profiles of a virialized, collisionless system, which is also the approach used to set up the initial conditions for the N -body simulations¹². We segment the halo into $N = 200$ concentric shells, logarithmically spaced between $r_{\min} = 0.01r_s$ and $r_{\max} = 100r_s$, and closely follow the methodology of Nishikawa et al. (2020) to evolve the system in time.

The conduction parameter C , which controls the efficiency of heat transfer in the LMFP regime, needs to be calibrated against numerical simulations. We do so by comparing the evolution of the central density of the tNFW halo in the numerical simulation to that obtained using the gravothermal fluid equations. This yields $C = 0.76$ (red-dashed line in top-left panel of Fig. 2), which is in excellent agreement with Koda & Shapiro (2011) and Nishikawa et al. (2020), who obtained a similar value for C to match the evolution of an untruncated NFW profile. Using $C = 0.76$ also accurately fits the evolution of the ABG[0.1] halo (red-dashed line in bottom-left panel of Fig. 2). For completeness, we point out that some studies have advocated different calibration values for C . In particular, Essig et al. (2019) find $C \sim 0.6$, while a recent study by Mace et al. (2025) advocates for $C \sim 1.0$ ¹³. The reason for these discrepancies is currently unclear.

¹² This approach is valid as long as the impact of the SIDM self-interactions is small during the process of halo formation via violent relaxation.

¹³ Both Essig et al. (2019) and Mace et al. (2025) refer to C as β .

B. CALCULATING THE PHASE-SPACE DISTRIBUTION FUNCTION

In order to compute the evolution of the DF in our simulations, we proceed as follows. Given the phase-space coordinates of all particles in a simulation output, we compute for each particle its energy $E = \frac{1}{2}v^2 + \Phi(\mathbf{r})$ and angular momentum $L = |\mathbf{r} \times \mathbf{v}|$. These are used to infer

$$\frac{d^2 N(E, L)}{dE dL} = f(E, L) g(E, L), \quad (\text{B.1})$$

where $f(E, L)$ is the DF, and

$$g(E, L) = 16\pi^2 L \int_{r_p}^{r_a} \frac{dr}{\sqrt{2(E - \Phi) - (L/r)^2}}, \quad (\text{B.2})$$

is the density of states. Here r_p and r_a are the pericenter and apocenter distances for an orbit with energy E and angular momentum L , which, in a spherical potential, are the roots of

$$\frac{1}{r^2} + \frac{2(\Phi - E)}{L^2} = 0. \quad (\text{B.3})$$

We assume that the system remains spherical throughout its evolution, so that the density of states follows from equation (B.2) and the instantaneous spherically averaged potential.

The instantaneous DF then follows from equation (B.1).

Note that while we start out with an isotropic DF, $f = f(E)$, the sinking of the perturber transfers momentum to the halo particles, which may cause the DF to develop a dependence on L . Hence, in computing the DF we consider the more generic form $f = f(E, L)$. In practice, though, we typically infer the DF to have only a weak L -dependence. Therefore, in what follows, we compute the one-dimensional DF $f(E)$, which we obtain by marginalizing $f(E, L)$ over L , that is,

$$f(E) = \frac{\int_0^{L_c(E)} dL f(E, L) g(E, L)}{\int_0^{L_c(E)} dL g(E, L)}, \quad (\text{B.4})$$

with $L_c(E)$ the angular momentum of a circular orbit of energy E . For isotropic systems, the $f(E)$ thus obtained is identical to that inferred from its density distribution using the Eddington inversion.

This paper was built using the Open Journal of Astrophysics L^AT_EX template. The OJA is a journal which provides fast and easy peer review for new papers in the astro-ph section of the arXiv, making the reviewing process simpler for authors and referees alike. Learn more at <http://astro.theoj.org>.