Dark Matter and the Final Parsec Problem

Caitlyn Dewar

Department of Physics
McGill University, Montreal
Quebec, Canada

December, 2024

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Master of Science

Abstract

We show that the dynamical friction from a dark matter spike is sufficient to solve the bottleneck in modeling supermassive black hole binary coalescence from galactic mergers. This issue is often referred to as the "final parsec problem," since stellar loss-cone deficit from gravitational slingshot interactions causes the binary to take over a Hubble time to reach a parsec of separation. As a result, the black holes would be unable to reach the gravitational wave driven phase of inspiralling. We then demonstrate that self-interacting dark matter, which has been shown to solve small-scale structure problems like the core-cusp problem, is also better suited than cold dark matter to solve the final parsec problem. This is due to the presence of an isothermal core that could replenish the spike through scattering interactions. Finally, we compare our models to the recently observed characteristic strain from the pulsar timing array data and find that models with dark matter are generally preferred over a model without.

Abrégé

Nous démonstrons que la friction dynamique exercer par une surdensité de matière sombre est suffisant pour résoudre “le problème du dernier parsec” des trous noirs supermassives. Ce problème arrive quand nous modélisons les fusions des trous noirs supermassives qui viennent des fusions galactique. Lorsque la quantité d’étoiles est réduite par les interactions gravitationelles à trois corps, les trous noirs prennent plus long qu’un temps de Hubble pour s’approcher d’un parsec de séparation. Par consequence, les trous noirs sont trop éloingnés pour l’émission d’ondes gravitationelles. Ensuite, nous démonstrons que la matière sombre auto-interaggisant résoudre le problème du dernier parsec, autant plus que la matière sombre froide. Ceci est parce que le noyau isotherme reconstitute la surdensité par les diffusions. Finalement, nous comparons nos modèles avec les observations récentes des collaborations Pulsar Timing Array et trouvons que les modèles avec le matière sombre sont généralement préférés.

Statement of Authorship

This thesis is based off a paper that I wrote in collaboration with James Cline and Gonzalo Alonso-Álvarez [1]. I contributed to the development, discussion, and validation of the calculations. The code used here is either my own or adapted from their code. The data digitization of the Pulsar Timing Array (PTA) featured in table 5.2 was done by Gonzalo.

Acknowledgements

I would like to sincerely thank my supervisor James Cline for his guidance and support during my studies, as well as our collaborator Gonzalo Alonso-Álvarez for his insightful discussions. I would like to thank the Natural Sciences and Engineering Research Council of Canada (NSERC), the McGill Department of Physics, and the Trottier Space Institute (TSI) for their financial support which made this work possible. I would also like to thank Erny Au Yang, Srobona Basak, and Nirmalya Brahma for their friendship during my time here at McGill. Most importantly, I would like to thank my mother and father whose love means the world to me.

Contents

List of Tables

List of Figures

List of Abbreviations

Abbreviation	**Meaning**
AGN	Active Galactic Nuclei
BH	Black Hole
CDM	Cold Dark Matter
CPTA	Chinese Pulsar Timing Array
DF	Dynamical Friction
DM	Dark Matter
EPTA	European Pulsar Timing Array
FPP	Final Parsec Problem
GW	Gravitational Wave
HD	Hellings-Downs
InPTA	Indian Pulsar Timing Array
MM	Massive Mediator
NANOGrav	North American Nanohertz Observatory for Gravitational Waves
NFW	Navarro-Frenk-White
PPTA	Parkes Pulsar Timing Array
PTA	Pulsar Timing Array
SIDM	Self-Interacting Dark Matter
SMBH	Supermassive Black Hole
TOA	Time-of-Arrival

Chapter 1

Introduction

In June of 2023, the North American Nanohertz Observatory for Gravitational Waves (NANOGrav) announced that they had observed the Hellings-Downs (HD) correlation to $\sim 3\sigma$ from their analysis of 67 millisecond pulsars over 15 years [3]. This HD correlation of the timing residuals signifies the existence of a stochastic gravitational wave (GW) background [4]. Alongside NANOGrav, other pulsar timing array (PTA) collaborations like the Parkes Pulsar Timing Array (PPTA) [5], the European Pulsar Timing Array (EPTA) and the Indian Pulsar Timing Array (InPTA) [6], and the Chinese Pulsar Timing Array (CPTA) [7], have released complementary results.

The origin of the nanohertz GW background is likely due to the inspiralling of supermassive black hole (SMBH) binaries given the consistency of the characteristic strain spectrum with the predicted $h_c \propto f^{-2/3}$ power law [8, 9, 2, 10]. This result has brought back into focus the "final parsec problem" (FPP). The FPP is an issue that arises from modelling SMBH binary coalescence from galactic mergers [11]. Before reaching a stage where the SMBHs are sufficiently close for GW emission to be effective, they must undergo three-body interactions with stars in order to inspiral [12]. This depletes the amount of stars in their vicinity and results in the binaries taking over a Hubble time to reach $\mathcal{O}(1 \text{ pc})$ of separation [13, 14]. If, however, the PTA observations are consistent with SMBH inspiralling, then this would suggest that they are somehow able to bridge their final parsec.

A preliminary feature of the pulsar data suggests that there may be a low frequency turnover in the characteristic strain spectrum. This turnover is typically attributed to environmental effects on the binary at large separations [15, 16, 17]. Shortly after the release of the 2023 PTA dataset, it was suggested by Shen et. al that the turnover could be modeled by the dynamical friction from a cold dark matter (CDM) spike [18]. With the growing observational evidence for GWs produced by stellar mass black hole (BH) binaries [19, 20], many have begun to investigate whether the presence of dark matter (DM) around inspiralling

BHs could produce detectable signatures in GW experiments. This has been considered for a diverse range of black hole populations and dark matter models ranging from CDM ([21, 22, 23]) to ultra light dark matter ([24, 25, 26, 27]).

What few people have investigated is whether or not the presence of a DM spike around an SMBH binary could solve the FPP bottleneck. A DM spike is an overdensity that exists in the immediate area surrounding the black hole [28]. Assuming that this spike is centered around the binary, then it would exist within the region where interactions with stars become inefficient at driving inspiralling. The drag force that the DM exerts on the binary as it orbits (also called dynamical friction [29]) will cause the binary to lose orbital energy, thereby reducing the separation. We would therefore like to know whether this effect is enough to bring the binary to the GW dominated regime within a Hubble time.

CDM is viewed as the paradigm of DM models for its ability to solve large scale problems such as hierarchical structure formation [30]. However, in the last several decades it has been shown that CDM is less capable at modeling structure at smaller (sub Mpc) scales. This is particularly relevant for modeling DM halos in galaxies since CDM predicts overly "cuspy" inner regions [31], an issue which is referred to as the core-cusp problem [32, 33, 34]. Self-interacting dark matter (SIDM) was proposed as a solution to this problem since the self-scattering interactions would cause the halo to relax to the desired "cored" isothermal profile [35]. Since the features of the halo (and hence the spike) can vary heavily depending on whether we have CDM or SIDM, we are interested in studying the implications of both for the FPP.

Both dark matter and black holes have been challenging to study because they cannot be observed directly with traditional telescopes. Instead, their presence and feature must by deduced by their influence on the surrounding observable matter. Using GW signatures to study the role that DM plays in SMBH binary dynamics will allow us to probe previously inaccessible aspects of galactic environments.

The outline of this thesis are as follows. Chapter 2 discusses SMBHs, SMBH binaries, and the FPP. Several mass scaling relations are shown and used to convert from a SMBH mass to its surrounding DM halo mass. In chapter 3, we discuss the differences between CDM and SIDM and then construct several corresponding density profiles. In chapter 4, we study the dynamics of the binary from dynamical friction and gravitational wave emission. We then determine the conditions that are needed to drive the binaries to coalescence within a Hubble time. In chapter 5, we compare our results to the recent PTA data to determine which DM models are best suited. Finally, in chapter 6 we provide a summary of the results as well as a discussion of future directions.

Chapter 2

Supermassive Black Holes and Galaxies

2.1 Supermassive Black Hole Binaries

Supermassive black holes (SMBHs) are massive compact objects on the order of $10^6 - 10^{11} M_\odot$ that are found at the centres of most galaxies [36, 37]. Their origin and evolution are inextricably tied to the development of their host. This mutual relationship is reflected in the tight correlations that exist between SMBH mass and the properties of the stellar bulge [38, 39, 40].

SMBH binaries are thought to form as a result of galactic mergers. Such mergers are frequent and are well understood to occur in hierarchical structure formation scenarios from ΛCDM cosmology [41, 42]. Given that most galaxies contain SMBHs in their centres, it is reasonable to assume that galactic mergers would eventually bring the black holes (BHs) together. Among the first to propose forming SMBH binaries in this manner were Begelman et. al, who were interested in modelling the apparent bending and precession of radio jets [11]. They suggested that if the embedded SMBHs could get sufficiently close, then they would be able to emit gravitational waves (GWs).

Understanding the dynamics of SMBH binary formation from galactic mergers may provide insight into the development of SMBHs. While the dominant mode of growth occurs by gas accretion from the galactic core during the optically bright quasar phase [43], active galactic nuclei (AGN) feedback would eventually redistribute this gas away from the galactic core [44]. However, the influx of gas from a merger could help reignite SMBH growth in gas-poor environments [45, 46].

Galactic mergers are also a possible channel for SMBH seed formation. The influx of gas into the core of the merged galaxy could provide a sufficient overdensity leading to the direct collapse into an SMBH [47, 48, 49]. It should be noted that this is far from the only possible scenario for seed formation. Given the diverse range of SMBHs that exist from large blazars at high redshift to more moderate mass SMBHs at low redshift, it is likely that there are several paths to SMBH formation [50]. Other possible mechanisms for SMBH seeding include the direct collapse of population III stars [51], runaway collisions in young stellar clusters [52, 53], and the accretion of objects in the early universe such as primordial black holes or cosmic strings [54, 55].

Aside from the recent PTA observations, evidence for SMBH binaries is limited. Multi-wavelength searches have revealed that there are several candidate systems containing dual AGN [56]. Dual AGN contain SMBHs that are not yet close enough to be gravitationally bound. Since hardened binaries typically form at separations ≤ 10 pc [48, 49], it is difficult to resolve binary AGN into separate spatial components. In addition, it is not clear if systems containing precessing radio jets harbour SMBH binaries since this effect could be attributed to tilted accretion disks from single AGN [57]. In spite of these difficulties there are still several proposed SMBH binary candidates, most notably OJ 287 [58].

2.2 The Final Parsec Problem

The path to SMBH binary coalescence can be separated into three distinct stages [16]. In the first stage, the merging of the host galaxies brings the SMBHs towards the centre of the system via dynamical friction. When the binding energy of the SMBHs surpass that of the material inside their orbits, the binary will become hard [59]. The hardened binary then enters the second stage of the merger where they undergo three-body interactions with stars in their stellar loss-cone.[1] This is known as the gravitational slingshot mechanism, as the binary is brought closer together by the ejection of the stars [12]. In the final stage, the SMBHs are close enough that their inspiralling will be driven by GW emission.

There is a significant issue that presents itself in the second stage of the model. The ejection of stars from the stellar loss-cone results in fewer stars for the binary to be able to eject. The time it takes for the binary to reach the GW phase is therefore dependent on the time it takes for the loss-cone to be replenished [13]. This process can take longer than a Hubble time, leading the binary to stall at about a parsec of separation. Hence, this issue is often referred to as the "final parsec problem" (FPP).

[1]The stellar loss-cone is given by the set of trajectories that bring nearby orbiting stars within the capture/disruption region around the SMBHs [60] The trajectories fall within an opening angle that is determined by the capture/disruption radius.

In order to solve the FPP, one must either look to an alternative geometric configuration or outline a new mechanism that could drive the binary towards the GW dominated regime. Some have suggested that the FPP may come from assuming a spherically symmetric gravitational potential, and that stellar loss-cone deficit is not an issue for axisymmetric models [61] (though this has been disputed [62]). Other potential solutions involve interactions with a gaseous circumbinary disk [63, 64, 65] (which has also been disputed [66]), or the inclusion of a third black hole brought in by a subsequent merger [67, 68].

We would like to determine if the dynamical friction from a dark matter spike is efficient at bringing the binary to the GW driven phase of inspiralling within a Hubble time. To do this, we will assume that we have a circular binary in a spherically symmetric distribution of dark matter. We will build these profiles in the following section and in chapter 3, and then study their implications in chapters 4 and 5.

2.3 Black Hole to Halo Mass Relations

The overarching goal of this work is to investigate the effect of dynamical friction from a dark matter spike on SMBH binary orbits. This will depend heavily on the choice of density profile whose scale will be set by the total enclosed halo mass. Therefore, the first challenge is to determine this halo mass for a given SMBH mass $M_\bullet$. We follow the approach that is shown in Shen et al. [18], who use a series of mass relations to convert from the SMBH mass to the halo mass at a given redshift.

The first of these relations is the stellar bulge to black hole mass relation from Kormendy and Ho [40], given by

$$\log_{10}\left(\frac{M_\bullet}{M_\odot}\right) = 8.7 + 1.1\log_{10}\left(\frac{M_{bul}}{10^{11}M_\odot}\right) \tag{2.1}$$

where M_{bul} is the mass of the stellar bulge. It is straightforward to determine M_{bul} for a given choice of $M_\bullet$ by algebraic manipulation.

Next we have the stellar to stellar bulge mass relation from Chen et. al [69]. This is given by

$$M_{bul} = f_\star(M_\star)M_\star \tag{2.2}$$

where $M_\star$ is the total stellar mass and $f_\star(M_\star)$ is a factor of proportionality given by

$$f_\star(M_\star) = \begin{cases} 0.615 & M_\star \leq 10^{10}M_\odot \\ 0.615 + \frac{\sqrt{6.9}}{(\log_{10}M_\star - 10)^{1.5}}\exp\left[\frac{-3.45}{\log_{10}M_\star - 10}\right] & M_\star > 10^{10}M_\odot \end{cases}. \tag{2.3}$$

Since the factor of proportionality is dependent upon the stellar mass of the galaxy, then $M_\star$ must be determined numerically via interpolation. This is done using the `interp` function in the `Python` library `NumPy`.

Finally, the halo mass is calculated using the redshift dependent halo to stellar mass relation

$$\frac{M_\star}{M_{200}}(z) = 2A(z)\left[\left(\frac{M_{200}}{M_A(z)}\right)^{-\beta(z)} + \left(\frac{M_{200}}{M_A(z)}\right)^{\gamma(z)}\right]^{-1}, \tag{2.4}$$

where M_{200} is the virial mass (see chapter 3 for a definition) and the z-dependent parameters are shown in equations (7-10) of Girelli et. al [70]. Similar to the stellar to stellar bulge mass relation, the halo mass must also be determined via interpolation.

The end result of combining equations (2.1-2.4) is a black hole to halo mass relation, which is shown in figure 2.1. We will assume that the SMBH mass in our relation is given by the total binary mass $M_\bullet = M_1 + M_2 = M_1(1+q)$, where q is the mass fraction defined as $q \equiv M_2/M_1$. Therefore, a given halo mass is characterized by the parameters (M_1, q, z). This serves as the starting point for creating a dark matter profile from the outside-in.

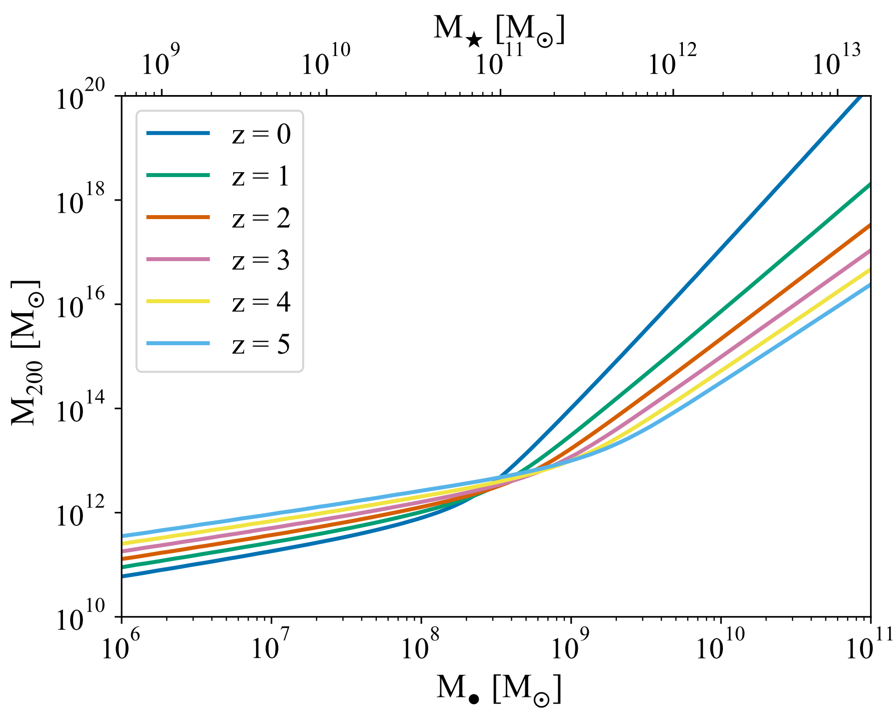

Figure 2.1: The black hole to halo mass relation using equations (2.1-2.4) for redshifts $z = 0, 1, ..., 5$. The corresponding stellar mass is shown on the upper x-axis.

Chapter 3

Dark Matter Density Profiles

3.1 The CDM Paradigm and SIDM

Most of the matter in our universe is composed of a substance that cannot be observed directly by traditional telescopes. However, the effect of this so-called "dark" matter (DM) can be seen in a variety of phenomena including (but not limited to) the rotation curves of galaxies [71, 72], the anisotropies in the cosmic microwave background [73], and large scale structure formation simulations [30]. Typically when modelling these phenomena the cold dark matter (CDM) model is invoked. This is a blanket term for a species of dark matter that decoupled from the early universe plasma at non-relativistic speeds and interacts minimally with Standard Model particles.

In spite of its success, there has been some concern as to whether or not the CDM paradigm translates to small (sub Mpc) scales. One relevant issue for this study is the core-cusp problem. Simulations of CDM tend to predict "cuspy" density profiles towards the centre [31]. However, observations of galactic halos suggest the realistic halos tend to become "cored" and flatten out in the centre [32, 33]. A comprehensive review of the core-cusp problem and other small scale structure problems is given by Tulin and Yu in reference [74].

One proposed solution to the small-scale structure problem is self-interacting dark matter (SIDM). SIDM has the basic feature that it can undergo $2 \to 2$ scattering interactions unlike its inert counterpart CDM [35]. This means that in regions where the DM is sufficiently sparse it will retain the same features as CDM, while in denser regions the scatterings will cause the halo to thermalize which leads to an isothermal (cored) profile [75].

In this work we are less concerned with specific models of particulate dark matter, but rather with the general macroscopic features that CDM or SIDM produce in galaxies. This has been studied in detail for both. We will use the results from these studies in the following sections to build appropriate DM density profiles.

3.2 Navarro-Frenk-White Profiles

We will assume that the outermost region of any dark matter halo can be modelled by the Navarro-Frenk-White (NFW) density profile. The NFW profile was found from N-body simulations of CDM halos [31]. It is given by

$$\rho_{NFW}(r) = \frac{\rho_s}{(\frac{r}{r_s})(1+\frac{r}{r_s})^2} \tag{3.1}$$

where r_s and ρ_s are the scale radius and scale density, respectively. The scale radius marks the transition of the profile from the inner r^{-1} power law to the outer r^{-3} power law.

To make use of the profile in our model, we will need to determine two parameters. The first is the total mass that is enclosed in the halo, while the second is the concentration parameter.

3.2.1 Halo Mass

The mass enclosed in the NFW profile is defined by the virial radius, R_{200}. This is a convention where the total halo mass within R_{200} is equal to the mass of a uniform sphere with an overdensity of 200 times the critical density, $\rho_c(z)$. That is to say,

$$\begin{aligned} M_{200}(z) &\equiv 200\rho_c(z) \times \frac{4}{3}\pi R_{200}^3 \\ &= 4\pi \int_0^{R_{200}} r^2 dr \rho_{NFW}(r) \end{aligned}, \tag{3.2}$$

where M_{200} is referred to as the virial mass. In the previous chapter we calculated the virial mass using a series a scaling relations that depend upon the initial SMBH binary mass and the redshift. We can therefore calculate the virial radius by rearranging the first line in equation (3.2).

The critical density $\rho_c(z)$ can be determined from ΛCDM cosmology and is given by

$$\begin{aligned} \rho_c(z) &= \frac{3H^2(z)}{8\pi G} \\ &= \frac{3H_0^2}{8\pi G}[\Omega_{\Lambda,0} + \Omega_{m,0}(1+z)^3] \end{aligned}, \tag{3.3}$$

where H_0 is the present-day Hubble constant, $\Omega_{\Lambda,0}$ is the present-day dark energy density parameter, and $\Omega_{m,0} \approx 1 - \Omega_{\Lambda,0}$ is the present-day matter density parameter. We take the values of these physical constants from Planck 2018 [73].

Using the second line in equation (3.2), we can determine the scale density associated with a given virial mass. This gives us

$$\begin{aligned} M_{200} &= 4\pi \int_0^{R_{200}} r^2 dr \frac{\rho_s}{(\frac{r}{r_s})(1+\frac{r}{r_s})^2} \\ &= 4\pi\rho_s r_s^3 \int_0^{u_{200}} du \frac{u-1}{u^2} \\ &= 4\pi\rho_s r_s^3 \left[\ln\left(1+\frac{R_{200}}{r_s}\right) - \frac{\frac{R_{200}}{r_s}}{1+\frac{R_{200}}{r_s}}\right], \\ \rho_s &= \frac{M_{200}}{4\pi r_s^3}\left[\ln\left(1+\frac{R_{200}}{r_s}\right) - \frac{\frac{R_{200}}{r_s}}{1+\frac{R_{200}}{r_s}}\right]^{-1} \end{aligned} \tag{3.4}$$

where we utilized the change of variable $u = r/r_s + 1$. In order to fully calculate the scale density, however, we will need to determine the scale radius. This is calculated in the next section.

3.2.2 The Concentration Parameter

The scale radius is related to the virial radius by the concentration parameter, given by $c_{200} = R_{200}/r_s$. This parameter determines where the scale transition occurs in the NFW profile and can vary depending on the size and cosmological history of the halo. To determine c_{200} in our model, we use the relation [76]

$$C(M_{200}, z) = C_0(z)\left(\frac{M_{200}}{10^{12}h^{-1}M_\odot}\right)^{-\gamma(z)}\left[1+\left(\frac{M_{200}}{M_0(z)}\right)^{0.4}\right], \tag{3.5}$$

which was derived from studying halo evolution in ΛCDM cosmological simulations. The parameters C_0, γ, and M_0 are themselves functions of z that need to be determined by interpolating the values in table 2 of Klypin et. al [76]. We compute this interpolation using the `interp` function in `NumPy`.

The concentration parameter as a function of SMBH mass is shown plotted in figure 3.1 for several values of redshift. Having determined the concentration parameter we can then obtain the scale density from equation (3.4).

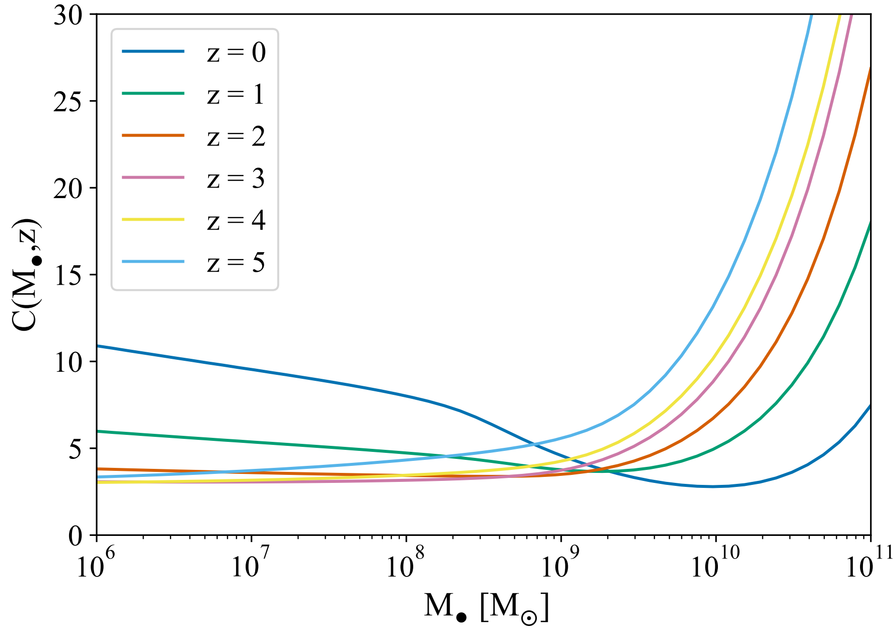

Figure 3.1: The concentration parameter as a function of black hole mass from equation (3.5) for redshifts $z = 0, 1, ..., 5$. The halo mass is calculated using the black hole to halo mass relation shown in figure 2.1.

3.3 Isothermal Profiles

The core-cusp problem poses a significant issue in the construction of our density profiles since realistic halos tend to be flatter in the centre compared to the predicted $\rho \propto r^{-1}$ of the NFW profile. We are therefore interested in building a complementary SIDM profile. We do so following the approach that is shown in reference [75]. They assume that there is a boundary at some radius r_c where the number of self-interactions in the age of the systems is unity. This means that for $r > r_c$, SIDM will behave like CDM and we can model the outer region by the NFW profile. This critical condition is determined by multiplying the scattering interaction rate by the age of the system, and is given by

$$N_{int}(r_c) = \frac{\langle \sigma v \rangle}{m} \rho_c t_{age} \approx 1, \tag{3.6}$$

where $\frac{\langle \sigma v \rangle}{m}$ is the velocity-weighted SIDM cross section per unit mass, ρ_c is the density of the NFW profile at r_c, and t_{age} is the age of the system. Note here that we fix the parameter $t_{age} = 0.1$ Gyr since the time it takes the system to form is an unknown.

In the core of the SIDM halo ($r < r_c$), the $2 \to 2$ scatterings will lead the DM to reach hydrostatic equilibrium. The density profile can then be determined by solving the following Poisson equation [75]:

$$v_c^2 \nabla^2 \ln \rho = -4\pi G \rho, \tag{3.7}$$

where v_c is the velocity dispersion of the core. Since we are assuming that the density distribution is spherically symmetric, this means that the Poisson equation becomes

$$\begin{aligned}
\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 \frac{\partial \ln \rho}{\partial r}\right) &= -\frac{4\pi G \rho}{v_c^2} \\
\frac{\partial}{\partial r}\left(\frac{r^2}{\rho}\frac{\partial \rho}{\partial r}\right) &= -\frac{4\pi G \rho r^2}{v_c^2}. \\
\left(\frac{2r}{\rho} - \frac{r^2}{\rho^2}\right)\frac{\partial \rho}{\partial r} + \frac{r^2}{\rho}\frac{\partial^2 \rho}{\partial r} &= -\frac{4\pi G \rho r^2}{v_c^2}
\end{aligned}$$

In order to simplify the algebra, we can introduce dimensionless variables $x = r/r_c$ and $\Lambda = \ln \rho/\rho_c$. This is done so that $x = 1$ and $\Lambda = 0$ at the boundary between the two profiles. The differential equation then becomes

$$\Lambda'' + \frac{2}{x}\Lambda' + Ce^{\Lambda} = 0, \tag{3.8}$$

where the prime denotes derivatives with respect to x and we defined $C = 4\pi G \rho_c r_c^2 / v_c^2$.

3.3.1 Solving the Poisson Equation

To solve for the isothermal profile in full, we need to define five constraints. The first two of these will come from setting the boundary conditions $\rho_{iso}(0) = \rho_0$ and $\rho'_{iso}(0) = 0$. This will be enough to solve equation (3.8) for a given choice of C. The second two constraints come from matching the isothermal and NFW profiles at the core radius r_c. This includes both the density profiles so that $\rho_c = \rho_{NFW}(r_c)$ and the enclosed mass $M_{iso}(r_c) = M_{NFW}(r_c)$. The remaining condition is given by the scattering-interaction constraint in equation (3.6). This is enough to solve for ρ_c, r_c, and v_c to give a unique value of C.

We solve equation (3.8) using the `ParametricNDSolveValue` function in `Mathematica`. While we have the boundary conditions $\Lambda(0) = \Lambda_0$ and $\Lambda'(0) = 0$, we don't know *a priori* which value of Λ_0 will solve the equation for a given choice of C. To overcome this issue, we vary the value of Λ_0 until we find one that gives us $\Lambda(1) = 0$. The solution is shown by the orange curve in figure 3.2 (left).

It is straightforward to determine the value of the core density, ρ_c, by substituting the core radius r_c into the NFW profile. From here, there are two more conditions remaining to solve for the profile in full.

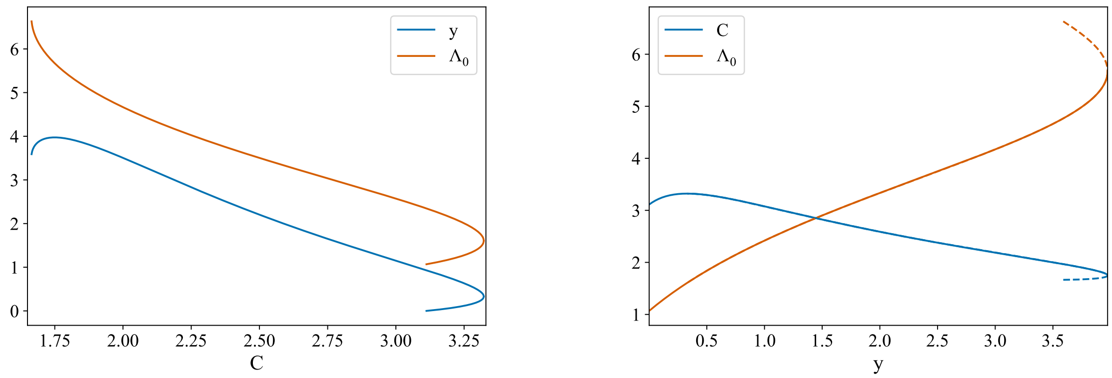

Figure 3.2: *Left*: The values of y and Λ_0 that satisfy both the Poisson equation and the conservation of mass for a given value of C. There exists both an upper and lower branch of solutions. *Right*: C and Λ_0 as a function of y. The dashed lines show where the solutions for $C < 1.75$ are cut in order for y to yield a unique C and Λ_0

The mass enclosed in the halo at r_c can be determined by integrating the density profile over the volume of a sphere. Note that for simplicity we let the profile extend to the origin (past the Schwarzschild radius). This is reasonable given that the mass enclosed in the SIDM core will be several orders of magnitude larger than the black hole binary mass. Equating the enclosed mass in the isothermal and NFW profiles gives us that

$$\begin{aligned} \int_0^{r_c} r^2 \rho_{iso}(r) dr &= \int_0^{r_c} r^2 \frac{\rho_s}{(\frac{r}{r_s})(1+\frac{r}{r_s})^2} dr \\ \int_0^1 x^2 \rho_c e^{\Lambda(x)} dx &= \int_0^1 x^2 \frac{\rho_s}{(xy)(1+xy)^2} dx, \\ \int_0^1 x^2 e^{\Lambda(x)} dx &= \int_0^1 \frac{x(1+y)^2}{(1+xy)^2} dx \end{aligned} \tag{3.9}$$

where we introduced another dimensionless variable $y = \frac{r_c}{r_s}$. With the solution to $\Lambda(x)$ for a given choice of C, we can determine the values of y that solve equation (3.9). The possible values of y are shown by the blue curve in figure 3.2 (left).

If we cut off values of C below 1.75, then for a given choice of y we can get a unique value of C and Λ_0. The functions $C(y)$ and $\Lambda_0(y)$ are shown plotted in right-hand side of figure 3.2. One of the motivations for doing this is so that we can keep the solutions that extend down to the $y \sim 0.001$ from the lower branch of $y(C)$.

Finally, we introduce the condition that the number of SIDM scatterings is $N_{int}(r_c) \sim 1$ at r_c using equation (3.6). For a general SIDM model, we assume that the cross section has some power-law velocity dependence following Shapiro and Paschalidis [77]. This gives us that

$$\sigma(v) = \sigma_0 \left(\frac{v}{v_0}\right)^{-a}, \tag{3.10}$$

where v_0 is an arbitrary reference velocity which we take to be $v_0 = 100$ km/s, σ_0 is the SIDM cross section at v_0, and $a = 0, 1, ..., 4$. The powers of $a = 0$ and $a = 4$ have important physical significance: $a = 0$ corresponds to a contact interaction; while $a = 4$ corresponds to a Coulomb-like interaction. We will also consider powers 1, 2, and 3, here for completeness.

From equation (3.10), we can re-express the velocity weighted cross section in terms of the core velocity dispersion and the reference values to get that

$$\begin{aligned} \langle \sigma v \rangle &= \sigma_c v_c \\ &= \sigma_0 \left(\frac{v_c}{v_0}\right)^{-a} v_c \\ &= \sigma_0 v_0 \left(\frac{v_c}{v_0}\right)^{-a+1} \end{aligned}. \tag{3.11}$$

If we substitute equation (3.11) back into our condition that $N_{int}(r_c) \sim 1$, we get that

$$\begin{aligned} 1 &= \frac{\langle \sigma v \rangle}{m} \rho_c t_{age} \\ &= \frac{\sigma_0}{m} v_0 \left(\frac{v_c}{v_0}\right)^{-a+1} \rho_c t_{age} \\ &= \frac{\sigma_0}{m} v_0 \left(\frac{v_c}{v_0}\right)^{-a+1} \frac{\rho_s}{y(1+y)^2} t_{age} \\ &= S \left(\frac{v_c}{v_0}\right)^{-a+1} \frac{1}{y(1+y)^2} \end{aligned}, \tag{3.12}$$

where we defined $S = (\sigma_0/m) v_0 \rho_s t_{age}$. Note that in the third line we used $\rho_c = \rho_{NFW}(r_c)$ and $y = r_c/r_s$. Given that $C = 4\pi G \rho_c r_c^2 / v_c^2$, we can rearrange this expression to solve for v_c and find that

$$\begin{aligned} \left(\frac{v_c}{v_0}\right)^2 &= \frac{4\pi G \rho_c r_c^2}{C v_0^2} \\ \left(\frac{v_c}{v_0}\right)^2 &= \frac{4\pi G \rho_s r_s^2 y}{C v_0^2 (1+y)^2} \\ \left(\frac{v_c}{v_0}\right)^{2-2a} &= \left(\frac{4\pi G \rho_s r_s^2 y}{C v_0^2 (1+y)^2}\right)^{(1-a)} \end{aligned}. \tag{3.13}$$

When we square both sides of equation (3.12) and substitute in equation (3.13), we then

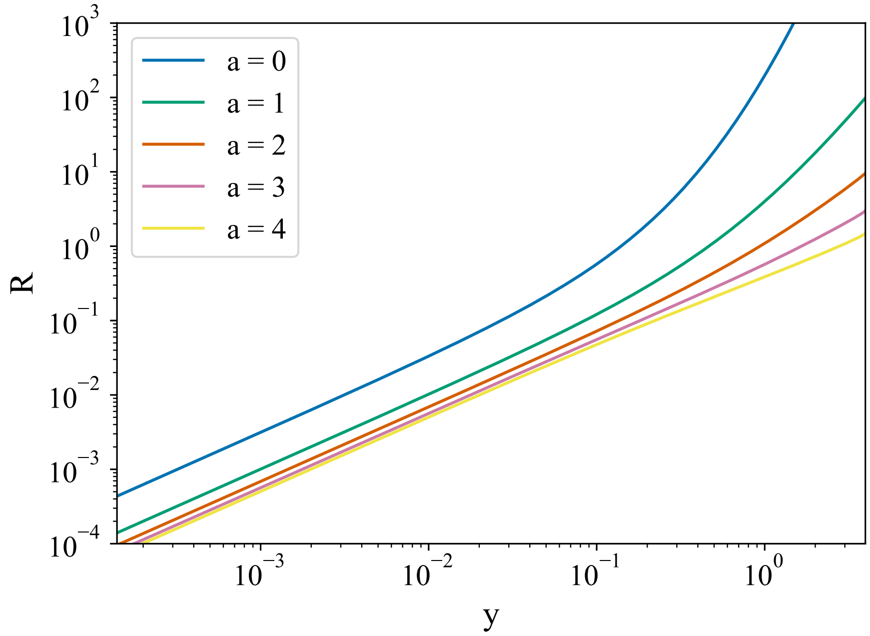

Figure 3.3: R vs. y for several values of a. We can determine the value of y by calculating the R associated to a given set of parameters (M_1, q, z, σ_0/m, a).

have that

$$\begin{aligned} 1 &= S^2 \left(\frac{v_c}{v_0}\right)^{-2a+2} \frac{1}{y^2(1+y^2)^4} \\ &= S^2 \left(\frac{4\pi G \rho_s r_s^2 y}{C v_0^2 (1+y)^2}\right)^{(1-a)} \frac{1}{y^2(1+y^2)^4} \\ &= S^2 C^{a-1} y^{-1-a} (1+y^2)^{2a-6} \left(\frac{4\pi G \rho_s r_s^2}{v_0^2}\right)^{(1-a)}, \\ S^2 \left(\frac{4\pi G \rho_s r_s^2}{v_0^2}\right)^{(1-a)} &= C^{-a+1} y^{1+a} (1+y^2)^{6-2a} \\ R &= C^{(1-a)/(1+a)} y (1+y^2)^{(6-2a)/(1+a)} \end{aligned} \tag{3.14}$$

where we defined $R^{1+a} = S^2 \left(4\pi G \rho_s r_s^2 / v_0^2\right)^{(1-a)}$. We plot equation (3.14) as a function of y for several choices of a in figure 3.3.

The end result is that a given choice of y will determine a unique value of R and vice versa. Since R depends only on our selected parameters (M_1, q, z, σ_0/m, a), this is enough to solve for the isothermal profile in full. We show a comparison of an NFW profile and an isothermal profile for $y = 0.5$ in figure 3.4.

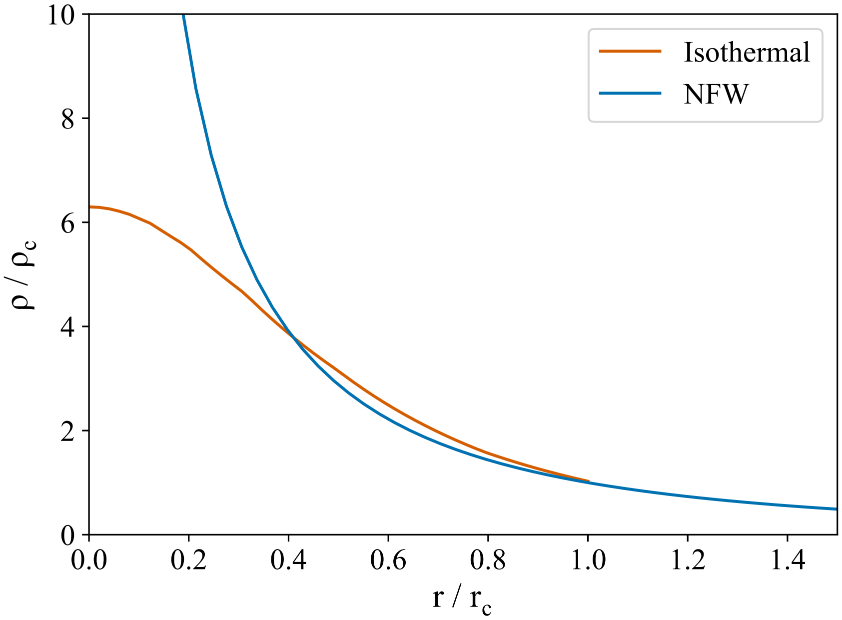

Figure 3.4: Comparison of a "cuspy" NFW profile (blue) to a "cored" isothermal profile (orange) for y=0.5. The figure is normalized with respect the core density ρ_c and core radius r_c.

3.3.2 The Massive Mediator Model

In realistic models of SIDM, the velocity dependence of the cross section may not be given by a simple power law. This is the case when SIDM interactions are mediated by the exchange of a massive mediator (MM) of mass m_ϕ. As an example, consider the differential cross section listed in equation (25) of reference [74]. In the limit where $m_\phi/m_\chi \gg v/c$, the interactions will behave like the $a = 0$ case. When $m_\phi/m_\chi \ll v/c$, the interactions will behave like the $a = 4$ case.

We would like to consider the MM model as a potential solution to the final parsec problem. To approximate the behaviour of the cross section in the limiting cases shown above, we will assume that there is a transition velocity $v_t \sim cm_\phi/m_\chi$ where the power law switches from $a = 0$ to $a = 4$. We will therefore use the following piecewise cross section for the MM model:

$$\sigma(v) = \frac{\sigma_0}{1 + (v/v_t)^4} \approx \sigma_0 \begin{cases} 1, & v < v_t \\ (v_t/v)^4, & v > v_t \end{cases}. \tag{3.15}$$

The location of the transition is straightforward to determine since the velocity is known in the isothermal and SIDM spike profiles (see the next section). Therefore, the MM model will be fully determined by the set of parameters (M_1, q, z, σ_0/m, v_t).

It is important to note that for some choices of parameters we find $v_c > v_t$, indicating that the transition should occur somewhere in the SIDM core. When this is the case, we will redefine the isothermal profile using $a = 4$ and replace σ_0/m with the effective cross section

$$\sigma_{eff} = \sigma_0 \left(\frac{v_t}{v_0} \right)^4 . \tag{3.16}$$

Note that this also changes the value of the velocity dispersion in the core.

3.4 Dark Matter Spikes

In the vicinity of the BH, the DM from the halo will be drawn gravitationally into an overdensity referred to as a spike. DM spikes were first proposed by Gondolo and Silk to constrain CDM annihilations from excess gamma rays in the galactic core [28]. Since CDM is collisionless, the spike would grow via adiabatic accretion. This means that the profile can be determined by relating the conserved thermodynamic quantities (angular momentum and radial action) before and after the formation of the spike. Gondolo and Silk showed that if the initial profile of the halo cusp is given by a power law with slope $0 \leq \gamma_0 \leq 2$, then the spike would also be a power law with slope $\gamma = (9 - 2\gamma_0)/(4 - \gamma_0)$. The spike profile would therefore be given by

$$\rho_{sp}(r) = \rho_{sp} \left(\frac{r}{r_{sp}} \right)^{-\gamma} , \tag{3.17}$$

where ρ_{sp} is the spike density at the spike radius r_{sp}.

The radius of the spike is determined by the radius of influence of the black hole. In the case where the region outside of the spike is an isothermal core with a constant velocity dispersion v_c, the radius of influence is defined as [78]

$$r_h = GM_\bullet / v_c^2 . \tag{3.18}$$

The mass enclosed in this region is of the same order of magnitude as the central BH. We will use r_h as the radius of the SIDM spike since we obtained v_c earlier from solving the Poisson equation. This is also the definition that was used in reference [77] to derive the SIDM spike profile.

Since we assume that the region outside of the CDM spike is given by the NFW profile, then we no longer have a constant velocity dispersion outside of the radius of influence.

Instead, we will define r_h for CDM as

$$2M_\bullet = 4\pi \int_0^{r_h} dr \rho(r) r^2, \tag{3.19}$$

where $\rho(r)$ is the foundational distribution (such as the NFW profile), and $r_{sp} \approx 0.2 r_h$ [78]. The definitions in equations (3.18) and (3.19) for the radius of influence are equivalent when the initial density profile is given by a singular isothermal sphere $\rho(r) = v_c/2\pi G r^2$.

3.4.1 Cold Dark Matter Spikes

Assuming that we are in a region where $r_{sp} \ll r_s$, we can approximate the NFW profile as a $\rho \propto r^{-1}$ power law. This means that the CDM spike radius is given by

$$\begin{aligned} 2M_\bullet &\approx 4\pi \int_0^{r_{sp}/0.2} dr \rho_s \frac{r_s}{r} r^2 \\ &= 4\pi \rho_s r_s \int_0^{r_{sp}/0.2} dr r \\ M_\bullet &= \pi \rho_s r_s (r_{sp}/0.2)^2 \\ r_{sp} &= 0.2 \left(\frac{M_\bullet}{\pi \rho_s r_s} \right)^{1/2} \end{aligned} . \tag{3.20}$$

From the spike radius, we can then calculate the spike density ρ_{sp} by inserting r_{sp} into the NFW profile.

The slope of the CDM spike is highly dependent on the environment of the black hole. One commonly cited value is $\gamma = 7/3$ which forms as a result of adiabatic accretion starting from an NFW ($\rho \propto r^{-1}$) profile around an isolated black hole [28]. Other common choices γ for CDM include: $\gamma = 1/2$ which forms after a galactic merger [79, 80]; $\gamma = 3/2$, which forms as a result of adiabatic accretion from a constant density isothermal core [78]; and $\gamma = 9/4$, which forms from the accretion around a primordial black hole [81].

3.4.2 Self-Interacting Dark Matter Spikes

The spike radius for SIDM will be given by the radius of influence that is defined in equation (3.18), since we already calculated v_c. We can then determine the spike density by inserting r_h into the isothermal profile. As for γ, it was shown by Shapiro and Paschalidis that the slope of the SIDM spike is related to the velocity dependence of the cross section in the collisional fluid approximation [77]. If the cross section obeys the power law dependence $\sigma \propto v^{-a}$, then the slope of the spike is given by $\gamma = (3 + a)/4$. This means that we will be looking at values of $\gamma = 3/4$, $4/4$, ..., $7/4$ for the general SIDM model.

For the MM model, there will be two important cases that we need to take into consideration: $v_c > v_t$ and $v_c < v_t$. In the case where our choice of v_t puts the transition in the SIDM core, we must recalculate the parameters of the isothermal profile by letting $a = 0 \rightarrow 4$ and $\sigma_0/m \rightarrow \sigma_{eff}/m$. This yields a new value of the v_c that will be used to determine the spike radius. The power of the spike will be given by $\gamma = 7/4$. If instead the threshold occurs in the spike, then we will need to determine where this transition occurs. Up until then we can calculate the profile identically to the $a = 0$ case which produces a spike with a slope of $\gamma = 3/4$.

It was also shown by Shapiro and Paschalidis that the velocity dispersion of the DM in the spike is given by [77]

$$v(r)/v_c \approx \frac{7}{11} + \frac{4}{11}(\frac{r_h}{r})^{1/2}. \tag{3.21}$$

Note that our choice of constant 7/11 deviates slightly from their definition ($7/11 \rightarrow 1$). This was selected to ensure that velocity dispersion at r_h matched the constant velocity dispersion of the isothermal core. We found that otherwise there would be a noticeable discontinuity in our results, like figure 4.1.

Equation (3.21) can then be rearranged to give the radius

$$r_t = \frac{121}{16} r_h (v_t/v_c - 7/11)^2 \tag{3.22}$$

at which the transition velocity v_t is reached in the spike. The density at the threshold will then be given by $\rho_t = \rho_{sp}(r_h/r_t)^{3/4}$. In the end we get a piecewise spike profile given by

$$\rho_{sp,MM}(r) = \begin{cases} \rho_t (r/r_t)^{-7/4} & 2R_{sch} < r \leq r_t \\ \rho_{sp}(r/r_h)^{-3/4} & r_t \leq r \leq r_h \end{cases} \tag{3.23}$$

when $v_t > v_c$.

3.5 Summary of Dark Matter Profiles

The complete CDM and SIDM halo density profiles are given by the piecewise functions listed below. They are cut off at twice the Schwarzschild radius R_{sch}, which was found to be the distance at which dark matter vanishes around a Schwarzschild black hole [82]. Some examples of these profiles are shown in figures 3.5 and 3.6 for an equal mass SMBH binary with $M_1 = 3 \times 10^9 M_\odot$ at $z = 0$.

The density profile for the case of CDM is given by

$$\rho_{CDM}(r) = \begin{cases} 0 & r \leq 2R_{sch} \\ \rho_{sp}(r/r_{sp})^{-\gamma} & 2R_{sch} < r \leq r_{sp} \\ \rho_{NFW}(r) & r \geq r_{sp} \end{cases}, \tag{3.24}$$

while the density profile for the general case of SIDM is given by

$$\rho_{SIDM}(r) = \begin{cases} 0 & r \leq 2R_{sch} \\ \rho_{sp}(r/r_{sp})^{-(3+a)/4} & 2R_{sch} < r \leq r_{sp} \\ \rho_{iso}(r) & r_{sp} \leq r \leq r_c \\ \rho_{NFW}(r) & r \geq r_c \end{cases}. \tag{3.25}$$

A little more care must be taken for the massive mediator model, since the transition from $a = 0$ to $a = 4$ could occur in either the core or the spike. We get that

$$\rho_{MM}(r) = \begin{cases} \begin{cases} 0 & r \leq 2R_{sch} \\ \rho_{sp}(r/r_{sp})^{-7/4} & 2R_{sch} < r \leq r_{sp} \\ \rho_{iso}(r) & r_{sp} \leq r \leq r_c \\ \rho_{NFW}(r) & r \geq r_c \end{cases} & v_t \leq v_c \\ \\ \begin{cases} 0 & r \leq 2R_{sch} \\ \rho_t(r/r_t)^{-7/4} & 2R_{sch} < r \leq r_t \\ \rho_{sp}(r/r_{sp})^{-3/4} & r_t \leq r \leq r_{sp} \\ \rho_{iso}(r) & r_{sp} \leq r \leq r_c \\ \rho_{NFW}(r) & r \geq r_c \end{cases} & v_t > v_c \end{cases}, \tag{3.26}$$

where the core parameters must be recalculated using the effective cross section when $v_t \leq v_c$.

All in all, the profiles require 4-5 parameters in order to be described in full. For CDM, we require (M_1, q, z, γ). For SIDM, we require (M_1, q, z, σ_0/m, a) in the general case, and (M_1, q, z, σ_0/m, v_t) for the MM case. Since the characteristic strain of the GW background will depend on the BH merger rate, we will need to integrate over parameters M_1, q, and z to model the spectrum. This leaves parameters like γ, σ_0/m, a, and v_t free to be selected, which we will do based on some phenomenological choices established in other works.

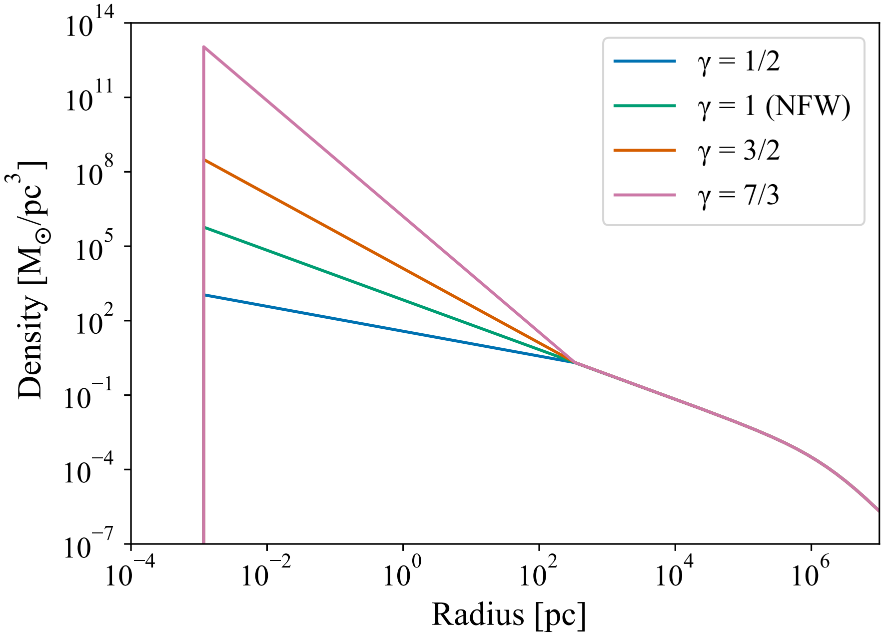

Figure 3.5: Example of several CDM profiles for an equal mass binary of $M_1 = 3 \times 10^9 M_\odot$ at redshift $z = 0$.

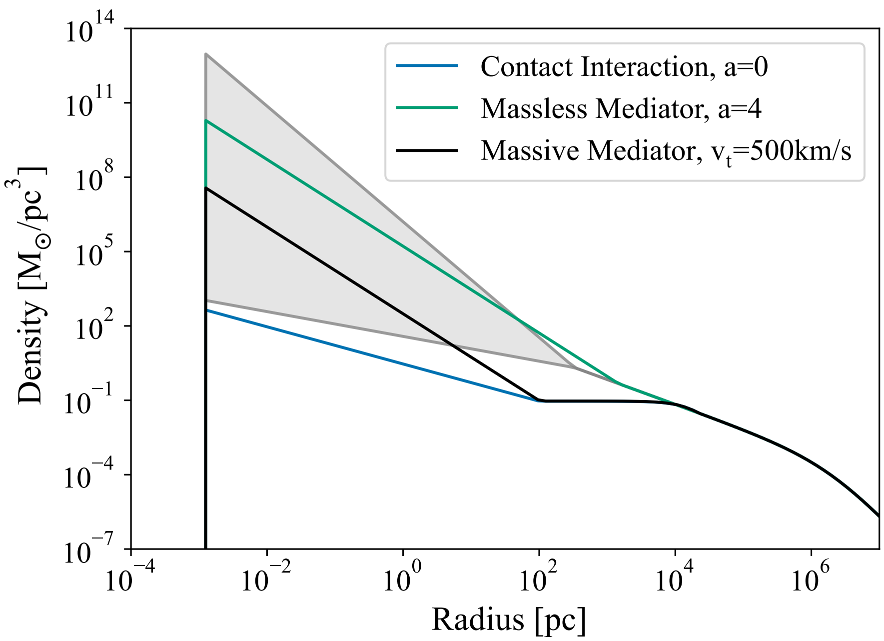

Figure 3.6: Example of several SIDM profiles with cross section $\sigma_0/m = 3$ cm^2/g for an equal mass binary of $M_1 = 3 \times 10^9 M_\odot$ at redshift $z = 0$. The CDM profiles from figure 3.5 are shown outlined in grey.

Chapter 4

Orbital Dynamics

4.1 Binary Orbits

We will model the orbital dynamics of the SMBH binary by assuming that we have two BHs of masses M_1 and M_2 that evolve in circular orbits. It is common to express the mass of the lighter BH, M_2, in terms of the heavier M_1 using the mass fraction q such that $M_2 \equiv qM_1$. This parametrization will become useful for when we calculate the SMBH merger rate. We therefore have that the total mass and the reduced mass of the binary in terms of M_1 and q are given by

$$M_\bullet = M_1 + M_2 = M_1(1+q) \tag{4.1}$$

and

$$\mu = \frac{M_1 M_2}{M_1 + M_2} = M_1 \frac{q}{(1+q)}. \tag{4.2}$$

The positions of the BHs relative to the centre of mass are given by the vectors $\vec{r}_1$ and $\vec{r}_2$ for M_1 and M_2, respectively. If we define the centre of mass as the origin of our coordinate system so that $\vec{r}_{cm} = \vec{0}$, we then find that the positions are proportional and related as

$$\begin{aligned} \vec{r}_{cm} &= \frac{M_1\vec{r}_1 + M_2\vec{r}_2}{M_1 + M_2} \\ \vec{0} &= \frac{\vec{r}_1 + q\vec{r}_2}{1+q} \\ \vec{r}_1 &= -q\vec{r}_2 \end{aligned} \quad . \tag{4.3}$$

In addition to the individual positions, we would also like to calculate the total separation. This is given by $R = |\vec{r}_1 - \vec{r}_2|$. When we plug in our result from equation (4.3) and rearrange, we find that the distance of each black hole from the origin can be expressed in terms of R as

$$r_1 = \frac{q}{(1+q)} R \tag{4.4}$$

and

$$r_2 = \frac{1}{(1+q)} R, \tag{4.5}$$

where it is clear that $r_1 + r_2 = R$.

We would like to determine the angular frequency of the binary since this determines the frequency in GWs. We know from the study of central forces in classical mechanics that the orbital motion of a two-body system can be described simply by the total mass, the reduced mass, and the binary separation (see chapter 7.4 in [83] as an example). Since we are assuming that the orbits are circular, this means that the force of gravity on the reduced mass acts as a centripetal force. We get that

$$\begin{aligned} \frac{GM_\bullet \mu}{R^2} &= \mu R \omega_s^2 \\ \omega_s^2 &= \frac{GM_1(1+q)}{R^3} \end{aligned}, \tag{4.6}$$

where ω_s is the angular frequency of the binary at the source. The circular nature of the orbits also means that all of the power in GWs is radiated in the second harmonic [84]. Therefore, the frequency in GWs is given by $f_{GW} = 2f_s = \omega_s/\pi$.

We now have an expression that relates the frequency in GWs to the separation of the binary, and we can take the derivative of this expression to find the inspiral rate dR/dt. We get that

$$\begin{aligned} f_{gw} &= \frac{\omega_s}{\pi} \\ &= \frac{1}{\pi}\sqrt{\frac{GM_1(1+q)}{R^3}} \\ \frac{df_{gw}}{dt} &= \frac{-3}{2} f_{gw} R^{-1} \frac{dR}{dt} \\ \frac{dR}{dt} &= \frac{-2}{3} f_{gw}^{-1} R \frac{df_{gw}}{dt} \end{aligned}. \tag{4.7}$$

Finally, we would like to determine the rate of orbital energy loss of the system. The orbital energy of the binary is given by the sum of the kinetic and gravitational potential energy

$$E_{orb} = \frac{1}{2}M_1\omega_s^2 r_1^2 + \frac{1}{2}M_2\omega_s^2 r_2^2 - \frac{GM_1M_2}{R} \\ = -\frac{qGM_1^2}{2R} . \tag{4.8}$$

When we take the time derivative of equation (4.8), we get that the rate at which the orbital energy changes is given by

$$P_{orb} = \frac{dE_{orb}}{dt} = \frac{qGM_1^2}{2R^2}\frac{dR}{dt}. \tag{4.9}$$

This gives us an expression that relates the power loss in orbital energy to the decrease in binary separation.

4.2 Energy Balance Equation

In our model there are two mechanisms that drive the binary to inspiral: gravitational wave emission and dynamical friction (DF) from a DM spike. From this, we have the energy balance equation

$$-\frac{dE_{orb}}{dt} = \frac{dE_{GW}}{dt} + \frac{dE_{DF}}{dt} \\ -P_{orb} = P_{GW} + P_{DF} , \tag{4.10}$$

which can be used to study the orbital evolution of the binary. There have been several papers that have looked at the role that dynamical friction from a spike may play in the inspiralling of intermediate mass BH binaries [21, 22], primordial BH binaries [85, 86], and more recently SMBH binaries [18]. Note that we do not include any back-reaction on the DM caused by the inspiralling in our model. We will however make an argument for the validity of this approach in the last section of this chapter.

The first of these effects is the energy loss due to GW emission, P_{GW}. Since we are assuming that we have a circular Newtonian binary, then the rate of energy loss in GWs is given by the lowest (quadrupole) order in the multipole expansion [87]

$$P_{GW} = \frac{dE_{GW}}{dt} = \frac{32}{5}\frac{G\mu^2}{c^5}R^4\omega_s^6. \tag{4.11}$$

The second effect that is driving the inspiralling is the dynamical friction exerted by the DM spike. Dynamical friction is the drag force exerted on an object by a continuous distribution of particles (often stars) with speeds smaller than that of the subject [29]. As the SMBHs pass through the DM spike, the drag force will cause the binary to decelerate and thereby lose some of its orbital energy. In the limit where the subject mass is much larger than the particle mass, this deceleration can be determined by the first order diffusion coefficient of the Fokker-Planck equation [88].

4.2.1 Dynamical Friction from CDM

We will assume in the case of CDM that the velocity of each BH is much larger than the velocity of the dark matter. This means that the acceleration due to dynamical friction is given by [88]

$$\frac{dv}{dt}(r) = 4\pi \ln \Lambda G^2 M \frac{\rho(r)}{v^2}. \tag{4.12}$$

The factor $\ln \Lambda$ is the Coulomb logarithm which is the ratio of the maximum to minimum impact parameters. There is some ambiguity in how to define $\ln \Lambda$ for an SMBH binary, with typical values ranging from $2 \leq \ln \Lambda \leq 5$ [89, 90]. We follow the approach of Shen et. al and take $\ln \Lambda \approx 3$ [18].

The net energy loss due to dynamical friction acting on both BHs is given by

$$\begin{aligned} \frac{dE_{DF}}{dt} &= M_1 v_1 \frac{dv_1}{dt} + M_2 v_2 \frac{dv_2}{dt} \\ &= 12\pi G^2 \frac{\mu^2 R^2}{\omega_s} \left[\frac{\rho_{DM}(r_1)}{r_1^3} + \frac{\rho_{DM}(r_2)}{r_2^3} \right]. \end{aligned} \tag{4.13}$$

If we use equations (4.1-4.5) to rewrite this expression in terms of M_1, q, and R, we get that

$$\begin{aligned} \frac{dE_{DF}}{dt} &= 12\pi G^2 \frac{\mu^2 R^2}{\omega_s} \left[\frac{\rho_{DM}(r_1)}{r_1^3} + \frac{\rho_{DM}(r_2)}{r_2^3} \right] \\ &= 12\pi G^2 \frac{M_1^2 q^2 R^2 R^{3/2}}{(1+q)^2 (GM_1(1+q))^{1/2}} \left[\frac{\rho_{DM}(r_1)}{r_1^3} + \frac{\rho_{DM}(r_2)}{r_2^3} \right] \\ &= 12\pi G^{3/2} \frac{M_1^{3/2} q^2 R^{7/2}}{(1+q)^{5/2}} \left[\frac{(1+q)^3 \rho_{DM}(\frac{qR}{1+q})}{(qR)^3} + \frac{(1+q)^3 \rho_{DM}(\frac{R}{1+q})}{R^3} \right] \\ &= 12\pi G^{3/2} M_1^{3/2} q^2 (1+q)^{1/2} R^{1/2} \left[q^{-3} \rho_{DM}(\frac{qR}{1+q}) + \rho_{DM}(\frac{R}{1+q}) \right] \\ &= 12\pi G^{3/2} M_1^{3/2} q^2 (1+q)^{1/2} R^{1/2} \left[q^{-3} \rho_{sp}(\frac{qR}{(1+q) r_{sp}})^{-\gamma} + \rho_{sp}(\frac{R}{(1+q) r_{sp}})^{-\gamma} \right] \\ &= 12\pi G^{3/2} M_1^{3/2} q^2 (1+q)^{1/2+\gamma} R^{1/2-\gamma} \rho_{sp} r_{sp}^{\gamma} \left[q^{-3-\gamma} + 1 \right] \end{aligned} \tag{4.14}$$

is the total dynamical friction from a CDM spike that is exerted on the binary.

4.2.2 Dynamical Friction from SIDM

In the case where we have SIDM, we can no longer assume that the velocities of the BHs surpass that of the DM entirely due to the scattering interactions. Instead, we must use the generalized version of equation (4.12) given by

$$\frac{dv}{dt}(r) = 12\pi G^2 M \frac{\rho(r)}{v^2}\left[\mathrm{erf}(X) - \frac{2X}{\sqrt{\pi}}e^{-X^2}\right], \tag{4.15}$$

where $X = v/\sqrt{2}\sigma$ and σ is the velocity of the dark matter [88]. The term in square brackets, which we define as

$$N(X) \equiv \mathrm{erf}(X) - \frac{2X}{\sqrt{\pi}}e^{-X^2}, \tag{4.16}$$

represents the fraction of DM with speeds less than v for a Maxwellian distribution of particles. When we have $v \gg \sigma$, this gives us that $N(X \gg 1) \to 1$ and we recover equation (4.12).

We can show that the value of X_i ($i = 1, 2$) for SIDM is approximately independent of the BH and DM velocities and only dependent on q. The velocity of the i^{th} SMBH is given by

$$\begin{aligned} v_i &= \omega_s r_i \\ &= \sqrt{\frac{GM_1(1+q)}{R^3}}r_i \end{aligned}. \tag{4.17}$$

Meanwhile, the velocity of the DM in the spike is given by

$$\begin{aligned} \sigma_i &\approx \frac{4}{11}\sigma_c r_h^{1/2} r_i^{-1/2} \\ &= \frac{4}{11}\sigma_c\sqrt{\frac{GM_1(1+q)}{\sigma_c^2}}r_i^{-1/2} \\ &= \frac{4}{11}\sqrt{GM_1(1+q)}r_i^{-1/2} \end{aligned} \tag{4.18}$$

as an approximation of equation (3.21). When we substitute equations (4.17) and (4.18) back into our expression for X_i, we get that

$$
\begin{aligned}
X_i(q) &= \frac{v_i}{\sqrt{2}\sigma_i} \\
&\approx \frac{11}{4\sqrt{2}}(\frac{r_i}{R})^{3/2} \\
&= \frac{11}{4\sqrt{2}} \begin{cases} q^{3/2}(1+q)^{-3/2} & i=1 \\ (1+q)^{-3/2} & i=2 \end{cases}
\end{aligned} . \tag{4.19}
$$

The rate of energy loss due to dynamical friction from an SIDM spike (in the general case of $a = 0, 1, ..., 4$) is therefore given by

$$\frac{dE_{DF}}{dt} = 12\pi G^{3/2} M_1^{3/2} q^2 (1+q)^{1/2+\gamma} R^{1/2-\gamma} \rho_{sp} r_{sp}^{\gamma} \left[N(X_1(q)) q^{-3-\gamma} + N(X_2(q)) \right] . \tag{4.20}$$

As for the massive mediator case, the power loss due to dynamical friction will be nearly identical to equation (4.20). The main differences are that: for $v_c \geq v_t$ we will use the effective cross section in equation (3.16) with $\gamma = 7/4$; and when $v_c < v_t$ we will let $r_{sp} \to r_t$, $\rho_{sp} \to \rho_t$ and $\gamma = 3/4 \to 7/4$ when $r < r_t$.

4.3 Timescales

From the energy balance equation, we see that there are two important regimes that dictate the orbital motion: the dynamical friction driven regime at large separations; and the GW driven regime at smaller separations. If we are in a region where one of these two effects dominates, then we can drop the subleading term on the right-hand side of equation (4.10) to find analytical solutions.

We would like to determine whether the dynamical friction is enough to bring the binary to the GW driven regime in less than a Hubble time. As a case study, we will pick a benchmark binary model with parameters $M_1 = 3 \times 10^9 M_\odot$, $q = 1$, and $z = 0$. This is a typical representative of the SMBH binary population that can be found from the peak in the BH merger rate density spectrum (see equation (5.2)). We assume conservatively that interactions with stars and gas only bring the binary to an initial separation of $R_\star = 10$ pc. This is roughly the separation at which a SMBH binary hardens [56]. After this point, the inspiralling will be driven by dynamical friction from the spike until the binary reaches a separation of $R_{GW} = 0.1$ pc where GW emission will dominate.

For an equal mass binary in an SIDM spike, the energy balance equation in the dynamical friction driven regime is given by

$$\begin{aligned}
P_{orb} &= -P_{DF} \\
\frac{GM_1^2}{2R^2}\frac{dR}{dt} &= -24(2)^{1/2+\gamma}\pi N(X(1))G^{3/2}M_1^{3/2}R^{1/2-\gamma}\rho_{sp}r_{sp}^{\gamma} \\
\frac{dR}{dt} &= -48(2)^{1/2+\gamma}\pi N(X(1))G^{1/2}M_1^{-1/2}R^{5/2-\gamma}\rho_{sp}r_{sp}^{\gamma} \\
\frac{dR}{dt} &= -48(2)^{1/2+\gamma}\pi N(X(1))G^{1/2}M_1^{-1/2}R^{5/2-\gamma}\rho_{sp}r_{sp}^{\gamma} \\
\frac{2r_{sp}}{t_{sp}}\frac{dx}{d\tau} &= -48(2)^{1/2+\gamma}\pi N(X(1))G^{1/2}M_1^{-1/2}(2xr_{sp})^{5/2-\gamma}\rho_{sp}r_{sp}^{\gamma}, \\
\frac{dx}{d\tau} &= -192\pi N(X(1))G^{1/2}M_1^{-1/2}x^{5/2-\gamma}\rho_{sp}r_{sp}^{3/2}t_{sp} \\
&= -192\pi N(X(1))G^{1/2}M_1^{-1/2}x^{5/2-\gamma}\rho_{sp}r_{sp}^{3/2}(\frac{r_{sp}^3}{GM_1})^{1/2} \\
&= -192\pi N(X(1))M_1^{-1}\rho_{sp}r_{sp}^{3}x^{5/2-\gamma} \\
&= -Bx^{5/2-\gamma}
\end{aligned} \tag{4.21}$$

where we've introduced variables $B = 192\pi N(X(1))M_1^{-1}\rho_{sp}r_{sp}^3$, $t_{sp} = \sqrt{r_{sp}^3/GM_1}$, $x = R/(2r_{sp})$, and $\tau = t/t_{sp}$ to simplify the calculations. Note that from equation (4.19) we have $X_1(1) = X_2(1) \approx 0.6875$ for an equal mass binary, which means that $N(X_1(1)) = N(X_2(1)) \approx 0.1855$. We can also obtain the inspiral rate in a CDM spike from equation (4.21) since we've assumed that $N(X_i(q)) \approx 1$ for CDM.

Using the method of separation of variables, we then have that the timescale for dynamical friction is given by

$$\begin{aligned}
\tau_{DF} &= -\frac{1}{B}\int_{x_\star}^{x_{GW}} dx x^{\gamma-5/2} \\
t_{DF} &= \frac{t_{sp}}{B}\begin{cases} \frac{x_\star^{\gamma-3/2}-x_{GW}^{\gamma-3/2}}{B(\gamma-3/2)} & \gamma \neq 3/2 \\ \ln(\frac{x_\star}{x_{GW}}) & \gamma = 3/2 \end{cases}.
\end{aligned} \tag{4.22}$$

This is sufficient to calculate any of the CDM or general SIDM profiles, but for the massive mediator case we must be mindful of where the $a = 0$ to $a = 4$ transition occurs. If the transition occurs in the isothermal core, then we simply need to use the spike parameters that are determined by the effective cross section with $\gamma = 7/4$. However, if the transition occurs somewhere in the spike we could have that either $R_\star < r_t$ or $R_{GW} < r_t < R_\star$. This

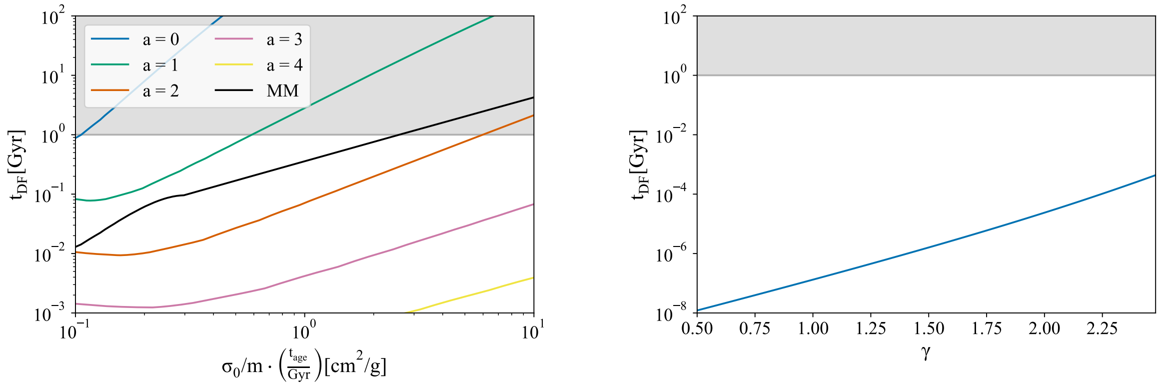

Figure 4.1: The dynamical friction timescale for SIDM (left) and CDM (right). The shaded region with $t_{DF} > 1$ Gyr is excluded as a potential solution to the final parsec problem. The MM case is shown by the black curve (left) with $v_t = 500$ km/s.

means that t_{DF} for $v_t > v_c$ becomes

$$t_{DF} = \frac{2t_{sp}}{Bx_t}\begin{cases}(x_\star^{1/4} - x_{GW}^{1/4}) & x_\star < x_t \\ (\frac{5}{3}x_t^{1/4} - x_{GW}^{1/4} - \frac{2}{3}x_\star^{-3/4}x_t) & x_{GW} < x_t < x_\star\end{cases}. \tag{4.23}$$

The timescale for dynamical friction using our benchmark binary is shown for SIDM (left) and CDM (right) in figure 4.1. From the parts of the curves that lie outside of the shaded grey area, we see that there are several regions in the parameter space where $t_{DF} \leq$ 1 Gyr. While this is reassuring, we have yet to investigate how the binary may disrupt the DM spike during inspiralling. This is done in the next section.

4.4 The Disruption of the Spike

We assumed in the previous calculations that the spike profile remained constant during inspiralling. However, this assumption is valid only if the energy dissipated to the DM via dynamical friction can be absorbed without disrupting its distribution. To determine if such a back-reaction could occur, we will make an order of magnitude estimate by comparing the change in orbital energy ΔE_{orb} to the binding energy of the spike U_{sp}. If $\Delta E_{orb} > U_{sp}$, then the injected energy would disrupt the spike and complicate the previous results.

From equation (4.8), we have that the change in orbital energy of the SMBHs from $R_\star$ to R_{GW} will be given by

$$\Delta E_{orb} = \frac{q}{2}GM_1^2(R_{GW}^{-1} - R_\star^{-1}). \tag{4.24}$$

This is equivalent to the energy gained by the DM since we are assuming that inspiralling is driven by dynamical friction from $R_\star$ to R_{GW}.

4.4.1 Binding Energy of the Spike

The gravitational binding energy of the spike can be calculated analytically from the shell theorem of integration. Consider a spherical shell of radius r whose mass is given by $dM_{shell}(r) = 4\pi r^2 \rho_{sp} r_{sp}^{\gamma} r^{-\gamma} dr$. The binding energy of the shell will depend on the enclosed mass, which in this case has two components. The first is the SMBH binary of mass $M_\bullet = M_1(1+q)$ which we will assume is located at the origin for simplicity. The second is the enclosed dark matter sphere, $M_{sp}(r)$, given by

$$\begin{aligned} M_{sp}(r) &= 4\pi \int_{R_{sch}}^{r} dr' r'^2 \rho_{sp} r_{sp}^{\gamma} r'^{-\gamma} \\ &= 4\pi \rho_{sp} r_{sp}^{\gamma} \int_{R_{Sch}}^{r} dr' r'^{2-\gamma} \\ &= \frac{4}{3-\gamma} \pi \rho_{sp} r_{sp}^{\gamma} (r^{3-\gamma} - R_{Sch}^{3-\gamma}) \\ &= \frac{4}{3-\gamma} \pi \rho_{sp} r_{sp}^{3} \left(\left(\frac{r}{r_{sp}} \right)^{3-\gamma} - \epsilon^{3-\gamma} \right) \end{aligned} , \tag{4.25}$$

where in the last line we've factored out $r_{sp}^{3-\gamma}$ from the term in brackets and defined the ratio $\epsilon = R_{sch}/r_{sp}$. The binding energy of the shell with the enclosed mass is then given by

$$\begin{aligned} dU_{SP} &= \frac{G(M_\bullet + M_{sp}(r)) dM_{shell}(r)}{r} \\ &= 4\pi G(M_\bullet + M_{sp}(r)) \rho_{sp} r_{sp}^{\gamma} r^{1-\gamma} dr \end{aligned} . \tag{4.26}$$

From equation (4.26), we can determine the binding energy of the spike by integrating over all possible shells. We therefore have that

$$\begin{aligned} U_{SP} &= 4\pi G \rho_{sp} r_{sp}^{\gamma} \int_{R_{sch}}^{r_{sp}} (M_\bullet + M_{sp}(r)) r^{1-\gamma} dr \\ &= 4\pi G \rho_{sp} r_{sp}^{\gamma} M_\bullet \int_{R_{sch}}^{r_{sp}} r^{1-\gamma} dr + 4\pi G \rho_{sp} r_{sp}^{\gamma} \int_{R_{sch}}^{r_{sp}} M_{sp}(r) r^{1-\gamma} dr \end{aligned} . \tag{4.27}$$

Let's break the calculation down into the two separate integrals. The first is the integral to find the binding energy of the spike with the black holes. We get that

$$
\begin{aligned}
U_{BH-SP} &= 4\pi G M_1(1+q)\rho_{sp} r_{sp}^{\gamma} \int_{R_{sch}}^{r_{sp}} dr r^{1-\gamma} \\
&= 4\pi G M_1(1+q)\rho_{sp} r_{sp}^{\gamma} \frac{(r_{sp}^{2-\gamma} - R_{Sch}^{2-\gamma})}{2-\gamma}. \\
&= 4\pi G M_1(1+q)\rho_{sp} r_{sp}^{2} \frac{1-\epsilon^{2-\gamma}}{2-\gamma}
\end{aligned}
\tag{4.28}
$$

The second integral is the one involving the self binding energy, and is given by

$$
\begin{aligned}
U_{SP-SP} &= \frac{16}{3-\gamma}\pi^2 G \rho_{sp}^2 r_{sp}^{\gamma+3} \int_{R_{sch}}^{r_{sp}} dr (r^{4-2\gamma} r_{sp}^{\gamma-3} - \epsilon^{3-\gamma} r^{1-\gamma}) \\
&= \frac{16}{3-\gamma}\pi^2 G \rho_{sp}^2 r_{sp}^{\gamma+3} \left[\frac{r_{sp}^{2-\gamma}}{5-2\gamma} - \frac{R_{sch}^{5-2\gamma} r_{sp}^{\gamma-3}}{5-2\gamma} - \frac{\epsilon^{3-\gamma} r_{sp}^{2-\gamma}}{2-\gamma} + \frac{\epsilon^{3-\gamma} R_{Sch}^{2-\gamma}}{2-\gamma} \right]. \\
&= \frac{16}{3-\gamma}\pi^2 G \rho_{sp}^2 r_{sp}^{5} \left[\frac{1-\epsilon^{5-2\gamma}}{5-2\gamma} + \frac{\epsilon^{5-2\gamma} - \epsilon^{3-\gamma}}{2-\gamma} \right]
\end{aligned}
\tag{4.29}
$$

The total binding energy will therefore be $U_{SP} = U_{BH-SP} + U_{SP-SP}$.

In figure 4.2 we show the ratio of the orbital energy loss (equation (4.24)) to the binding energy of the spike (equation (4.27)) for our benchmark case. For both CDM and SIDM there are no viable choice of slope γ (and σ_0/m for SIDM) that would prevent the disruption of the spike. Note that the massive mediator case will lie somewhere between the $a = 0$ and $a = 4$ cases, so it also will be disrupted.

However, not all hope is lost. If the relaxation timescale for self-interactions is sufficiently small compared to the dynamical friction timescale, then it may still be possible for the SIDM spike to avoid the back-reaction. This will depend on whether or not the isothermal core is disrupted by the inspiralling. If not, then the core could act as a reservoir that replenishes the spike through $2 \to 2$ scatterings.

4.4.2 Binding Energy of the SIDM Core

Let's now look at the binding energy of the SIDM core. In this case we are only concerned with determining the self-binding energy and not the binding energy with the binary. This is because the mass enclosed in the spike is typically of the same order of magnitude as the BH mass. The mass of the core, however, is several orders of magnitude larger and so the self-binding energy will dominate.

In this case we will not be able to calculate the binding energy analytically from the shell theorem. This because the density profile of the core is numerically determined from the

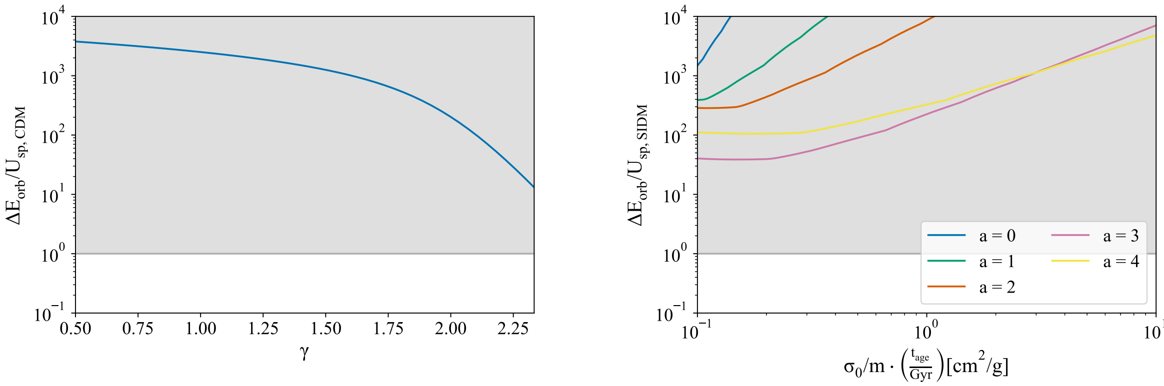

Figure 4.2: *Left*: Ratio of orbital energy loss to the binding energy of the CDM spike versus the spike slope γ. For the benchmark model there is no viable choice of γ where the spike avoids disruption. *Right*: Ratio of orbital energy loss to the binding energy of the SIDM spike versus the SIDM cross section for several choices of velocity dependence a. Similar to the case of CDM, there is no viable choice of parameters for the benchmark model that would prevent the disruption of the spike. Note that the massive mediator case would interpolate between the cases of $a = 0$ (blue) and $a = 4$ (yellow).

Poisson equation. Instead, we must find the binding energy between two points in the core and integrate over all such points. This gives us the expression

$$\begin{aligned} U_{core} &= G \int d^3r_1 d^3r_2 \frac{\rho(\vec{r_1})\rho(\vec{r_2})}{2|\vec{r_1} - \vec{r_2}|} \\ &= 4\pi^2 G\rho_c^2 r_c^5 \int dx_1 dx_2 \sin\alpha d\alpha \frac{e^{\Lambda(x_1)+\Lambda(x_2)}}{\sqrt{x_1^2 + x_2^2 - 2x_1x_2\cos\alpha}}, \\ &= 4\pi^2 G\rho_c^2 r_c^5 \int dx_1 dx_2 du \frac{e^{\Lambda(x_1)+\Lambda(x_2)}}{\sqrt{x_1^2 + x_2^2 - 2x_1x_2 u}} \end{aligned} \tag{4.30}$$

where the factor of 2 in the denominator is included to prevent overcounting and we've defined $x_i = r_i/r_c$ and $u = \cos\alpha$. Given the symmetries of the sphere, we can extract a factor of $8\pi^2$ so that we only need to integrate over dimensionless variables x_1, x_2, and u. Note that we get a factor of 4π from integrating over the sphere once, and 2π from the azimuthal symmetry of the denominator. The integral in equation (4.30) is computed using the `NIntegrate` function in `Mathematica` and the results are plotted in figure 4.3. We see that there is now a region in the parameter space where the core is not disrupted by the inspiralling.

In summary, we have found that while the timescale for dynamical friction from CDM is less than 1 Gyr, the spike will be completely disrupted by the inspiralling and therefore cannot be used to solve the final parsec problem. While SIDM suffers from a similar issue,

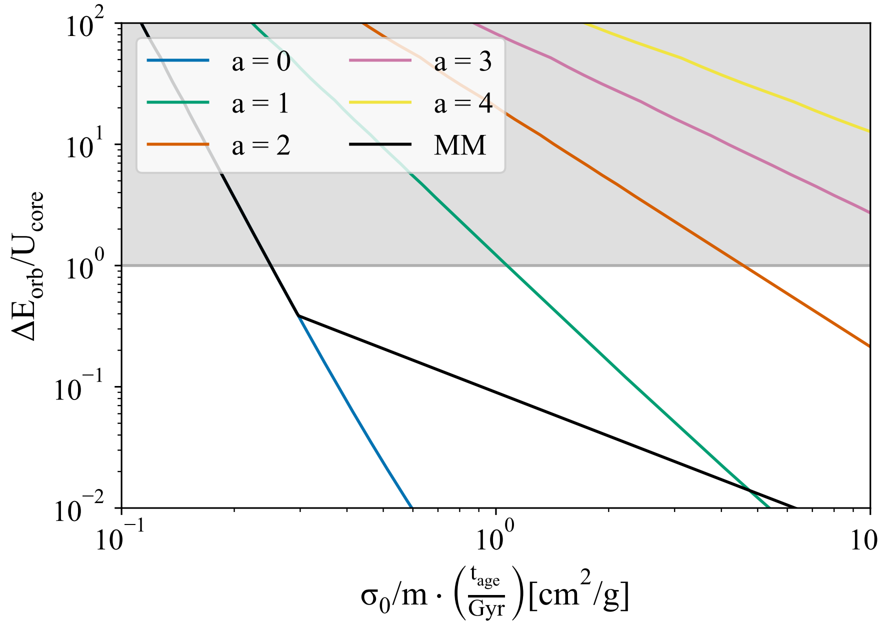

Figure 4.3: Ratio of the orbital energy lost from inspiralling to the binding energy of the SIDM core. In contrast with the spike (CDM or SIDM), there is a viable range of cross sections for which the core can withstand the disruption.

it may be possible for the spike to be replenished via $2 \to 2$ scattering interactions from the isothermal core. Since there is a region in the parameter space where the core can withstand the energy injected by the binary, then it may still be possible for SIDM (including the MM model) to solve the final parsec problem. We will investigate the implications of these results for the PTA data in the next chapter.

Chapter 5

Pulsar Timing Array Data

5.1 Pulsars Timing Arrays

Pulsars are neutron stars that emit beams of electromagnetic radiation along their magnetic poles. These poles are generally misaligned with the axis of rotation of the star, leading to the observation of pulses by a distant observer such as the Earth. A common analogy is that these pulses are like the flashes of light from a lighthouse [91]. The timing of these pulses is so consistent that it is possible to create "time-of-arrival" (TOA) functions that predict when they are set to arrive.

The principle behind using pulsar signals to uncover GW physics is simple. Since gravity alters the geometry of spacetime, a GW that passes between a pulsar and the Earth will alter the travel time of the pulses. This delay would be proportional to the amplitude of the GW. Using the TOA function, one could therefore study the features of the wave by calculating the timing residual.

The challenge with using the pulses from a single source is that it is unlikely for a realistic pulsar to produce a high signal-to-noise ratio [89]. However, it was shown by Hellings and Downs (HD) that cross-correlating the timing residuals from a set of pulsars could reveal a quadrupole signature that is unique to a GW background [4]. It was announced in 2023 that the PTA collaborations NANOGrav, the PPTA, the EPTA, and the CPTA had observed this HD correlation to $\sim 3\sigma$ after over 15 years of data collection [3, 5, 6, 7].

From the timing residual spectrum, it is then possible to Fourier transform the data into a characteristic strain spectrum. The characteristic strain represents the average amplitude of the stochastic GW background. It was shown by NANOGrav and the EPTA that the characteristic strain from the recent PTA dataset closely followed an $f^{-2/3}$ power law [2, 10], which is the predicted spectrum for a circularly inspiralling SMBH binary population [8, 9].

A notable (albeit preliminary) feature of the characteristic strain spectrum suggests that there may be a low frequency turnover away from the predicted power law. Assuming that the GWs are produced by SMBH inspiralling, it is likely this effect would be produced by interactions with the environment before reaching the GW dominated phase [15, 16, 17]. This can be understood intuitively from the reciprocal proportionality between frequency and separation in equation (4.6).

With this in mind, we would like to determine how well our DM models can reproduce this turnover in the characteristic strain spectrum compared to a case without DM. In particular, we want to determine if the viable solutions to the final parsec problem also show promise here. In the following sections we calculate the characteristic strain for several DM models and compare them to the PTA data.

5.2 Modelling Characteristic Strain

The characteristic strain $h_c(f)$ represents the average amplitude of the GW background. To find the $h_c(f)$ that is produced by a SMBH binary population, one must integrate over the strain contributed by each individual source. For a circular binary, this individual strain is proportional to the luminosity in GWs as $h_s^2 \propto P_{GW}$ [92]. One can therefore show that the characteristic strain produced by a population of circular SMBH binaries is given by [8]

$$h_c^2(f) = \frac{4G}{\pi c^2 f} \int dM_1 dq dz \frac{d^3 n}{dM_1 dq dz} \frac{dE_{GW}}{df_s}, \tag{5.1}$$

where f is the detected GW frequency and $f_s = f(1+z)$ is the GW frequency at the source. The integrand is weighted by the SMBH merger rate density (per unit comoving volume) [2]

$$\frac{d^3 n}{dM_1 dq dz} = \frac{d^3 n_{gal}}{dM_{\star 1} dq_\star dz} \frac{dM_{\star 1}}{dM_1} \frac{dq_\star}{dq}, \tag{5.2}$$

where $d^3 n_{gal}/dM_{\star 1} dq_\star dz$ is the galactic merger rate density, $M_{\star 1}$ is the stellar mass surrounding M_1, and $q_\star \equiv M_{\star 2}/M_{\star 1}$ is the stellar mass fraction.

The derivative of the stellar mass with respect to black hole mass can be found using the mass scaling relations in chapter 2. By combining equations (2.1) and (2.2), we have that the black hole to stellar mass relation is given by

$$\begin{aligned}\log_{10}\left(\frac{M_1}{M_\odot}\right) &= 8.7 + 1.1\log_{10}\left(\frac{f_\star M_{\star 1}}{10^{11}M_\odot}\right)\\ M_1 &= 10^{8.7}\left(\frac{f_\star M_{\star 1}}{10^{11}M_\odot}\right)^{1.1} M_\odot \\ M_{\star 1} &= \frac{1}{f_\star}10^{11}\left(\frac{M_1}{10^{8.7}M_\odot}\right)^{1/1.1} M_\odot\end{aligned} . \tag{5.3}$$

Here we've assumed that $f_\star \approx 0.615$ for all values of stellar mass as an order of magnitude estimate of equation (2.3). This is reasonable given that the peak of the distribution $f_\star \approx 0.783$ at $M_\star \sim 10^{12}M_\odot$ offers only a small correction to the final result. For constant $f_\star$, we then have that $dM_{\star 1}/dM_1$ is given simply by differentiating equation (5.3). We find that

$$\frac{dM_\star}{dM_1} = \frac{1}{1.11}\frac{M_\star}{M_1}. \tag{5.4}$$

Next, we would like to calculate the differential mass fraction $dq_\star/dq$. We know that the stellar mass fraction is given by $q_\star = M_{\star 2}/M_{\star 1}$. When we use our black hole to stellar mass relation in $q_\star$ and differentiate, we find that

$$\begin{aligned} q_\star &= \frac{M_{\star 2}}{M_{\star 1}} \\ &= \left(\frac{M_2}{M_1}\right)^{1/1.11} \\ &= q^{1/1.11} \\ \frac{dq_\star}{dq} &= \frac{1}{1.11}q^{-0.11/1.11}\end{aligned} . \tag{5.5}$$

For the galactic merger rate we take the function that was used in the recent analysis from NANOGrav [2, 69], given by

$$\frac{d^3 n_{gal}}{dM_{\star 1}dq_\star dz} = \frac{\Psi(M_{\star 1}, z')}{M_{\star 1}\ln(10)}\frac{P(M_{\star 1}, q_\star, z')}{T_{gal-gal}(M_{\star 1}, q_\star, z')}\frac{dt}{dz'}, \tag{5.6}$$

where $\Psi(M_{\star 1}, z)$ is the galactic stellar mass function, $P(M_{\star 1}, q_\star, z)$ is the galaxy pair fraction, $T_{gal-gal}(M_{\star 1}, q_\star, z)$ is the galaxy merger time, and dt/dz accounts for the span in the merger timescale.

For dt/dz we use the function

$$\begin{aligned}\frac{dt}{dz} &= \frac{1}{H(z)(1+z)} \\ &= \frac{1}{H_0(1+z)\sqrt{\Omega_{\Lambda,0}+(1+z)^3\Omega_{m,0}}},\end{aligned} \tag{5.7}$$

where H(z) is the Hubble parameter at redshift z and the parameters H_0, $\Omega_{\Lambda,0}$, $\Omega_{m,0}$ are taken from Planck [73]. For the parametrization of the galaxy pair fraction and merger time we have that

$$P(z) = P_0(1+z)^{\beta_{p0}} \tag{5.8}$$

and

$$T_{gal-gal}(q_\star, z) = T_0(1+z)^{\beta_{t0}} q_\star^{\gamma_{t0}}, \tag{5.9}$$

where the parameters P_0, β_{p0}, T_0, β_{t0}, and γ_{t0} are given in table 5.1.

As for the galaxy stellar mass function, we have that

$$\Psi(M_{\star 1}, z') = \ln(10)\Psi_0 \left(\frac{M_\star}{M_\psi}\right)^{\alpha_\psi} \exp\left(-\frac{M_\star}{M_\psi}\right) \tag{5.10}$$

where the redshift dependent parameters are given by

$$\log_{10}(\Psi_0(z)/Mpc^3) = \psi_0 + \psi_z z \tag{5.11}$$

$$\log_{10}(M_\psi(z)/M_\odot) = m_{\psi 0} + m_{\psi z} z \tag{5.12}$$

$$\alpha_\psi = 1 + \alpha_{\psi 0} + \alpha_{\psi z} z. \tag{5.13}$$

The fixed parameters ψ_z, $m_{\psi 0}$, $m_{\psi z}$, $\alpha_{\psi 0}$, and $\alpha_{\psi z}$ are again given in table 5.1. However, we leave ψ_0 as a free parameter. Since ψ_0 can be factored out of the integral without affecting M_1, q, or z, we can adjust its value to change the overall normalization. One reason that we do this is so that we can determine if our fit results yield a value of ψ_0 that is closer to the astrophysical prior $\psi_0 = -2.56 \pm 0.40$, since the posterior for ψ_0 in the NANOGrav analysis favoured a larger value due to the low frequency turnover [2].

Given that typical SMBH masses range from $10^6 - 10^{11} M_\odot$, it will be useful to perform the change of variable $\mu \equiv \log_{10}(M_1/M_\odot)$ to simplify the numerical integration. This gives us that the integral in equation (5.1) becomes

$$h_c^2(f) = \frac{4GM_\odot \ln(10)}{\pi c^2 f} \int d\mu dq dz \frac{d^3 n}{d\mu dq dz} \frac{dE}{df_s} 10^\mu. \tag{5.14}$$

Parameter	Value
ψ_0	Free
ψ_z	-0.6
$\mathrm{m}_{\psi 0}$	11.5
$\mathrm{m}_{\psi z}$	0.11
$\alpha_{\psi 0}$	-1.21
$\alpha_{\psi z}$	-0.03
P_0	0.033
β_{p0}	1
T_0	0.5 Gyr
β_{t0}	-0.5
γ_{t0}	-1
δ_M	1.1

Table 5.1: List of parameters needed to determine the SMBH merger rate density given by equation (5.2). These are the fiducial values from table B1 in reference [2]

.

We will set the integration limits to be $\mu \in [6, 11]$, $q \in [0, 1]$, and $z \in [0, 5]$ which follows the example of Shen et. al [18]. As seen in figure 12 of reference [2], this sample of the SMBH binary population contributes almost entirely to the GW background.

5.2.1 The Differential Energy Spectrum

The model dependence of the characteristic strain will come from the GW differential energy spectrum dE/df_s. Since we know that the rate of energy loss due to GW emission is given by equation (4.11), we can apply the chain rule to get that

$$\begin{aligned} P_{GW} &= \frac{dE_{GW}}{df_s}\frac{df_s}{dt} \\ \frac{dE_{GW}}{df_s} &= P_{GW}\dot{f_s}^{-1} \end{aligned} . \tag{5.15}$$

Now, we found from equation (4.7) that the time rate of change in GW frequency is related to the inspiral rate dR/dt. In addition, we know from equation (4.9) that the inspiral rate is related to the rate of orbital energy loss. With this in mind, we have that the differential energy spectrum is given by

$$\begin{aligned}
\frac{dE_{GW}}{df_s} &= P_{GW}\left(-\frac{2}{3}\frac{R}{\dot{R}}f_s^{-1}\right)\\
&= -\frac{2}{3}P_{GW}R\dot{R}^{-1}f_s^{-1}\\
&= -\frac{2}{3}P_{GW}R\left(\frac{qGM_1^2}{2R^2P_{orb}}\right)f_s^{-1}\\
&= -\frac{1}{3}qR^{-1}(GM_1^2)f_s^{-1}\frac{P_{GW}}{P_{orb}} \qquad\qquad , \\
&= -\frac{1}{3}q\left(\frac{(\pi f_s)^{2/3}}{G^{1/3}M_\bullet^{1/3}}\right)(GM_1^2)f_s^{-1}\frac{P_{GW}}{P_{orb}}\\
&= \frac{1}{3}q(1+q)^{-1/3}\pi^{2/3}G^{2/3}M_1^{5/3}f^{-1/3}(1+z)^{-1/3}\frac{P_{GW}}{P_{GW}+P_{DF}}
\end{aligned} \tag{5.16}$$

where the last step uses the relationship between the source frequency at redshift z and the detected frequency $f = f_s/(1+z)$. We see that the only part of the expression which is affected by the choice of DM model is the ratio of P_{GW}/P_{orb} since the total orbital energy loss is the sum in energy loss from dynamical friction and gravitational wave emission. With this, we are ready to calculate the characteristic strain. We perform the integration of equation (5.14) using the `Python` release of the numerical integration package `Vegas`.

5.3 Comparison of our model with PTA data

To obtain the PTA dataset, we digitized the timing residual plots from figure 1.a) of reference [2], figure 6 of reference [5], and figure 1 of reference [6]. The mean and upper/lower uncertainties are calculated by fitting a modified Gaussian distribution to the violin plots. The binned frequencies are determined by the duration T of the observations such that the i^{th} frequency is given by $f_i = i/T$. The extracted timing residuals are then converted to the corresponding characteristic strain using [2]

$$\Phi(f) = \frac{h_c^2}{12\pi^2 f^3}. \tag{5.17}$$

The results from the data digitization are shown in table 5.2.

We will now compare the characteristic strain that we get from integrating equation (5.14) numerically to the PTA data. For the model with no dark matter (NDM), we find a minimum $\chi^2_{NDM} = 18.2$ with $\psi_0 = -2.8$. We will use this as a reference when comparing our CDM and SIDM models to the PTA data. Note that $\psi_0 = -2.56 \pm 0.40$ is the astrophysical

Group	f [yr^{-1}]	$\mathrm{h}_c/10^{-15}$
NANOGrav [3]	0.062	$6.5^{+4.5}_{-2.4}$
	0.12	$7.9^{+3.0}_{-1.8}$
	0.19	$7.4^{+3.1}_{-2.0}$
	0.25	$6.4^{+3.3}_{-1.9}$
	0.31	$9.3^{+4.8}_{-4.2}$
PPTA [5]	0.055	$8.3^{+6.8}_{-3.4}$
	0.11	$9.6^{+4.2}_{-3.6}$
	0.17	$7.4^{+3.1}_{-1.8}$
	0.22	$6.3^{+4.5}_{-2.5}$
	0.28	$1.0^{+3.2}_{-0.8}$
	0.33	$6.3^{+4.4}_{-4.0}$
	0.39	$3.7^{+3.7}_{-1.5}$
	0.44	$7.1^{+3.9}_{-2.6}$
	0.50	$1.6^{+3.2}_{-1.0}$
	0.55	$3.6^{+3.6}_{-2.1}$
EPTA [6]	0.097	$8.0^{+4.0}_{-2.7}$
	0.19	$9.6^{+2.9}_{-1.9}$
	0.29	$8.2^{+3.8}_{-8.2}$
	0.39	$11.0^{+4.5}_{-2.9}$
	0.48	$5.2^{+10.0}_{-5.5}$
	0.58	$0.60^{+4.80}_{-0.40}$

Table 5.2: Tabulated values of the characteristic strain and frequency obtained by digitizing the timing residual plots from the 2023 PTA data release. This work was done in reference [1].

γ	ψ_0	χ^2_{min}	$\chi^2_{min} - \chi^2_{NDM}$
1/2	-2.8	18.0	-0.2
1	-2.7	14.7	-3.5
3/2	-2.2	14.5	-3.7
7/3	0.6	30.4	+12.2

Table 5.3: Comparison of the minimum χ^2 for CDM with several choices of γ. The value of ψ_0 is adjusted after numerical integration to determine the optimal fit. Note that all of these models are ruled out by the disruption of the spike shown in figure 4.2.

prior listed in reference [2], so $\psi_0 = -2.8$ in the NDM case falls within 1σ.

In the case where we have a CDM spike, there is an additional free parameter γ corresponding to the density slope. We select the values of γ by hand based off of previous CDM spike studies [28, 79, 80, 78]. We see from table 5.3 that characteristic strain models with CDM tend to perform better on average compared to the case of NDM, with the exception of $\gamma = 7/3$. It should be noted however that while these fits are an improvement, they are ruled out as solutions to the final parsec problem since the spikes are completely disrupted during inspiralling. This turns our attention back to SIDM, which was shown to have a viable range of possible solutions.

We are most interested in studying the SIDM cross sections which are consistent with solving the final parsec problem. The lower bound of this range is determined by calculating the cross section at which $E_{orb}/U_{core} \approx 1$ in figure 4.3 so that the core cannot be disrupted by the inspiralling. Meanwhile, the upper bound is determined by the cross section where $t_{df} \approx 1$ Gyr in figure 4.1.

First lets consider the general SIDM models. The cases of $a = 0$ and $a = 1$ for SIDM are immediately ruled out by these constraints. This is because the cross sections needed for $t_{df} \approx 1$ Gyr fall below the minimum cross sections needed for the SIDM core to be sustained. For $a = 2, 3, 4$, we have that there are not only a viable range of cross sections that can satisfy the final parsec problem, but that the best fits in each range are also an improvement over the NDM case. These results are summarized in table 5.4.

In the case of the massive mediator we are interested in determining which values of cross section and threshold velocity give the best fit to the PTA data. To determine the excluded region we calculate the core energy ratio using the benchmark binary parameters over a range of threshold velocities from 100 km/s to 1000 km/s. The results that do not satisfy the final parsec problem are shaded in grey in figure 5.1. We also plot several dynamical friction timescales from 50 Myr to 1 Gyr for our benchmark binary. Therefore, the viable range of parameters exist between the grey shaded region and the 1 Gyr curve (pink).

From calculating the strain for each pair of parameters (σ_0/m, v_t), we see that solutions

a	Viable σ_0/m [cm^2/g]	Best σ_0/m [cm^2/g]	ψ_0	χ^2_{min}	$\chi^2_{min} - \chi^2_{NDM}$
0	-	-	-	-	-
1	-	-	-	-	-
2	$4.5 \leq \sigma_0/m \leq 6$	4.75	-2.8	17.7	-0.5
3	$20 \leq \sigma_0/m \leq 90$	20	-2.8	17.2	-1.0
4	$50 \leq \sigma_0/m \leq 1600$	50	-2.7	14.7	-3.5

Table 5.4: Comparison of the minimum χ^2 for SIDM with several choices of $\gamma = (3 + a)/4$. The range of viable cross sections are determined from the boundary of the shaded regions in figures 4.1 and 4.3 in accordance with the final parsec problem. The value of ψ_0 is adjusted after numerical integration to determine the optimal fit.

which are consistent with the final parsec problem have a χ^2 value from 16 to 18, which is less than the NDM χ^2 value. Here we've selected a benchmark MM model with $\sigma_0/m = 3$ cm^2/g and 500km/s as a reference, labelled by $\star$ in figure 5.1. We use this benchmark for both its moderate cross section and transition velocity as well as its small $t_{DF} = 100$ Myr. This MM model gives us $\chi^2_{min} = 16.8$ and $\psi_0 = -2.8$, which is again an improvement over the NDM case.

In summary, we see that the vast majority of models with DM perform better than the NDM case. While CDM models provide a reasonable fit to the characteristic strain data (except for $\gamma = 7/3$), they are ruled out because the disruption of the spike spoils them as a solution to the final parsec problem. Meanwhile, SIDM not only provides a good fit to the PTA data but also has a region in its parameter space that is consistent with solving the final parsec problem. This is true for the general case as well as the more realistic massive mediator model. We plot the fit results for the NDM case, the CDM case with $\gamma = 3/2$, and the benchmark MM case in figure 5.2.

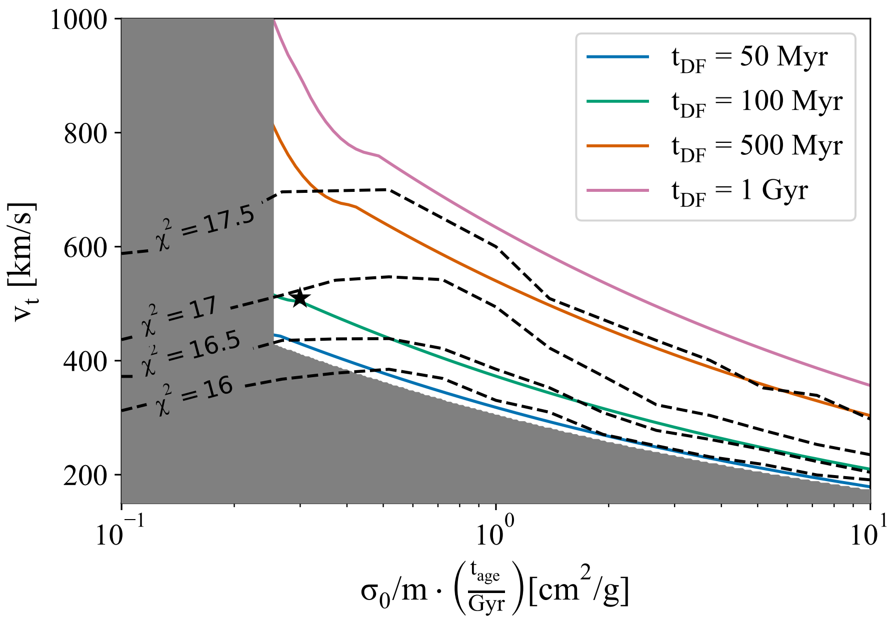

Figure 5.1: The parameter space for the massive mediator model. The shaded regions are excluded due to the disruption of the isothermal core (see figure 4.3). The coloured lines show several dynamical friction timescales for the benchmark binary. The dashed contour lines show the χ^2 fits to the data in table 5.2. The benchmark MM model with parameters $\sigma_0/m = 3$ cm^2/g and 500 km/s is shown by the symbol $\star$, which has $\chi^2_{min} = 16.8$ and $\psi_0 = -2.8$.

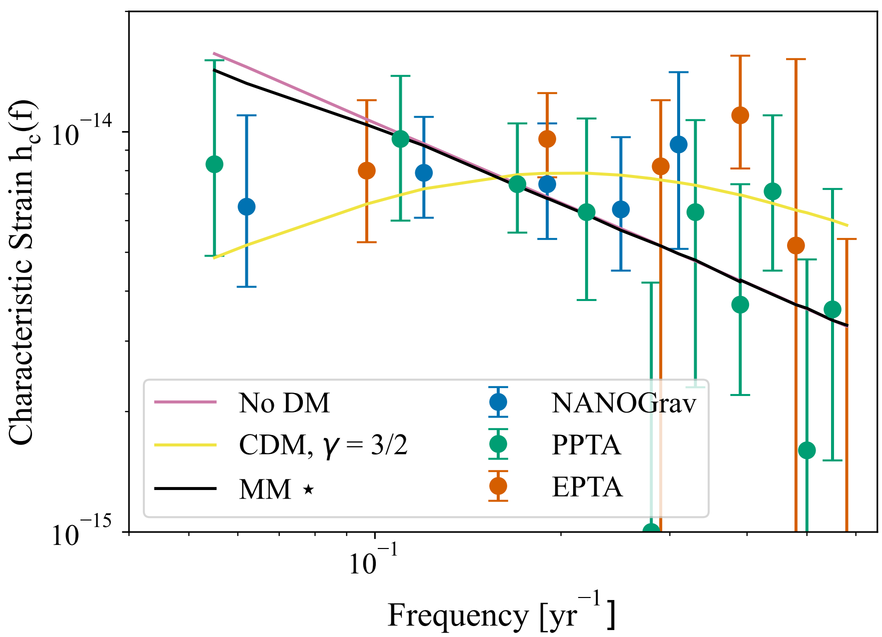

Figure 5.2: Characteristic Strain vs. Frequency. The error bars are from the extracted PTA data in table 5.2. The solid lines are the numerically determined characteristic strain for no DM (pink), CDM with $\gamma = 3/2$ (yellow), and the massive mediator benchmark with $\sigma_0/m = 3\ \mathrm{cm}^2/\mathrm{g}$ and $v_t = 500$ km/s (black). Note that the disruption of the spike rules out the CDM model as a viable solution, but we include it here for completeness.

Chapter 6

Discussion

In this work we have analyzed the effect that dynamical friction from a CDM or SIDM spike has on the inspiralling of SMBH binaries. In particular, we have studied its implication for the final parsec problem. We did this by considering a benchmark binary with parameters ($M_1 = 3 \times 10^9 M_\odot$, $q = 1$, $z = 0$). This was found to be the peak in the SMBH merger rate density from equation (5.2), making it a typical representative of the SMBH binary population.

For our benchmark binary, we determined the time it takes CDM and SIDM to shrink the orbital separation from $R_\star = 10$ pc (where interactions with stars and gas cease), to $R_{GW} = 0.1$ pc (where GW emission begins to dominate). We determined that there are indeed regions in the parameter space for both models where the dynamical friction timescale is less than 1 Gyr.

However, we also found that the energy injected into the spike during inspiralling is enough to disrupt it and complicate our findings. In spite of this we found that the inspiralling is not enough to disrupt the SIDM core. If the SIDM relaxation time is sufficiently small, this may allow the core to act as a reservoir from which $2 \to 2$ scattering interactions can rebuild the SIDM spike. We therefore find that only SIDM, both in general and for the massive mediator case, is compatible with solving the final parsec problem.

After constructing our DM models, we then calculated the resulting characteristic strain spectrum. We found that SMBH binary models which include the dynamical friction from a DM spike are generally preferred over a model without DM. This is likely because they contain the low frequency turnover that is not present in a GW-only model.

A pressing issue to address is the disruption of the spike caused by inspiralling. We showed that for our benchmark model, the energy injected into the spike by the binary is enough to disrupt it. While we made the argument that the SIDM core could help rebuild the spike, the full scope of this issue cannot be understood without further study into the disruption. One possible option may be to include feedback terms in our equations of motion that account for the injected energy, such as in references [21, 22]. Another possible route would be to conduct N-body simulations of SMBH mergers in DM halos.

The upcoming GW experiment LISA is expected to be sensitive to the lower mass range of the SMBH population ($\sim 10^7 M_\odot$) [93]. Unlike PTAs, LISA will be able to detect GW signatures from individual sources. Extensive work has be done to determine how CDM and ultralight dark matter around intermediate mass black holes may lead to the dephasing of the expected GW signature [22, 23, 26, 24]. Therefore, it could be useful to study the implications of our generalized SIDM and MM models in these systems.

With the growing number of GW observatories, there is the possibility that studying the phenomenological features of GWs could further our understanding of both BH populations and DM microphysics. In summary, this work serves as a promising first step for using dynamical friction from a dark matter spike to model the final stages of supermassive black hole binary inspiralling.

Bibliography

[1] G. Alonso-Álvarez, J. M. Cline, and C. Dewar, "Self-Interacting Dark Matter Solves the Final Parsec Problem of Supermassive Black Hole Mergers," *Phys. Rev. Lett.* **133** no. 2, (2024) 021401, `arXiv:2401.14450 [astro-ph.CO]`.

[2] **NANOGrav** Collaboration, G. Agazie *et al.*, "The NANOGrav 15 yr Data Set: Constraints on Supermassive Black Hole Binaries from the Gravitational-wave Background," *Astrophys. J. Lett.* **952** no. 2, (2023) L37, `arXiv:2306.16220 [astro-ph.HE]`.

[3] **NANOGrav** Collaboration, G. Agazie *et al.*, "The NANOGrav 15 yr Data Set: Evidence for a Gravitational-wave Background," *Astrophys. J. Lett.* **951** no. 1, (2023) L8, `arXiv:2306.16213 [astro-ph.HE]`.

[4] R. Hellings and G. Downs, "Upper limits on the isotropic gravitational radiation background from pulsar timing analysis," *Astrophys. J. Lett.* **265** (1983) L39–L42.

[5] D. J. Reardon *et al.*, "Search for an Isotropic Gravitational-wave Background with the Parkes Pulsar Timing Array," *Astrophys. J. Lett.* **951** no. 1, (2023) L6, `arXiv:2306.16215 [astro-ph.HE]`.

[6] **EPTA, InPTA:** Collaboration, J. Antoniadis *et al.*, "The second data release from the European Pulsar Timing Array - III. Search for gravitational wave signals," *Astron. Astrophys.* **678** (2023) A50, `arXiv:2306.16214 [astro-ph.HE]`.

[7] L. Hu, R.-G. Cai, and S.-J. Wang, "Distinctive GWBs from eccentric inspiraling SMBH binaries with a DM spike," `arXiv:2312.14041 [gr-qc]`.

[8] E. S. Phinney, "A Practical theorem on gravitational wave backgrounds," `arXiv:astro-ph/0108028`.

[9] A. H. Jaffe and D. C. Backer, "Gravitational waves probe the coalescence rate of massive black hole binaries," *Astrophys. J.* **583** (2003) 616–631, arXiv:astro-ph/0210148.

[10] **EPTA, InPTA** Collaboration, J. Antoniadis *et al.*, "The second data release from the European Pulsar Timing Array - IV. Implications for massive black holes, dark matter, and the early Universe," *Astron. Astrophys.* **685** (2024) A94, arXiv:2306.16227 [astro-ph.CO].

[11] M. C. Begelman, R. D. Blandford, and M. J. Rees, "Massive black hole binaries in active galactic nuclei," *Nature* **287** (1980) 307–309.

[12] W. C. Saslaw, M. J. Valtonen, and S. J. Aarseth, "The gravitational slingshot and the structure of extragalactic radio sources," *Astrophys. J.* **190** (1974) 253–270.

[13] M. Milosavljevic and D. Merritt, "Long term evolution of massive black hole binaries," *Astrophys. J.* **596** (2003) 860, arXiv:astro-ph/0212459.

[14] M. Milosavljevic and D. Merritt, "The Final parsec problem," *AIP Conf. Proc.* **686** no. 1, (2003) 201–210, arXiv:astro-ph/0212270.

[15] A. Sesana, F. Haardt, P. Madau, and M. Volonteri, "Low - frequency gravitational radiation from coalescing massive black hole binaries in hierarchical cosmologies," *Astrophys. J.* **611** (2004) 623–632, arXiv:astro-ph/0401543.

[16] D. Merritt and M. Milosavljevic, "Massive black hole binary evolution," *Living Rev. Rel.* **8** (2005) 8, arXiv:astro-ph/0410364.

[17] L. Sampson, N. J. Cornish, and S. T. McWilliams, "Constraining the Solution to the Last Parsec Problem with Pulsar Timing," *Phys. Rev. D* **91** no. 8, (2015) 084055, arXiv:1503.02662 [gr-qc].

[18] Z.-Q. Shen, G.-W. Yuan, Y.-Y. Wang, and Y.-Z. Wang, "Dark Matter Spike surrounding Supermassive Black Holes Binary and the nanohertz Stochastic Gravitational Wave Background," arXiv:2306.17143 [astro-ph.HE].

[19] **LIGO Scientific, Virgo** Collaboration, B. P. Abbott *et al.*, "Observation of Gravitational Waves from a Binary Black Hole Merger," *Phys. Rev. Lett.* **116** no. 6, (2016) 061102, arXiv:1602.03837 [gr-qc].

[20] **KAGRA, VIRGO, LIGO Scientific** Collaboration, R. Abbott *et al.*, "GWTC-3: Compact Binary Coalescences Observed by LIGO and Virgo during the Second Part of the Third Observing Run," *Phys. Rev. X* **13** no. 4, (2023) 041039, `arXiv:2111.03606 [gr-qc]`.

[21] K. Eda, Y. Itoh, S. Kuroyanagi, and J. Silk, "Gravitational waves as a probe of dark matter minispikes," *Phys. Rev. D* **91** no. 4, (2015) 044045, `arXiv:1408.3534 [gr-qc]`.

[22] B. J. Kavanagh, D. A. Nichols, G. Bertone, and D. Gaggero, "Detecting dark matter around black holes with gravitational waves: Effects of dark-matter dynamics on the gravitational waveform," *Phys. Rev. D* **102** no. 8, (2020) 083006, `arXiv:2002.12811 [gr-qc]`.

[23] A. Coogan, G. Bertone, D. Gaggero, B. J. Kavanagh, and D. A. Nichols, "Measuring the dark matter environments of black hole binaries with gravitational waves," *Phys. Rev. D* **105** no. 4, (2022) 043009, `arXiv:2108.04154 [gr-qc]`.

[24] K. Kadota, J. H. Kim, P. Ko, and X.-Y. Yang, "Gravitational Wave Probes on Self-Interacting Dark Matter Surrounding an Intermediate Mass Black Hole," `arXiv:2306.10828 [hep-ph]`.

[25] B. C. Bromley, P. Sandick, and B. Shams Es Haghi, "Supermassive Black Hole Binaries in Ultralight Dark Matter," `arXiv:2311.18013 [astro-ph.GA]`.

[26] L. Berezhiani, G. Cintia, V. De Luca, and J. Khoury, "Dynamical friction in dark matter superfluids: The evolution of black hole binaries," `arXiv:2311.07672 [astro-ph.CO]`.

[27] A. Boudon, P. Brax, and P. Valageas, "Supersonic friction of a black hole traversing a self-interacting scalar dark matter cloud," *Phys. Rev. D* **108** no. 10, (2023) 103517, `arXiv:2307.15391 [astro-ph.CO]`.

[28] P. Gondolo and J. Silk, "Dark matter annihilation at the galactic center," *Phys. Rev. Lett.* **83** (1999) 1719–1722, `arXiv:astro-ph/9906391`.

[29] S. Chandrasekhar, "Dynamical Friction. I. General Considerations: the Coefficient of Dynamical Friction," *Astrophys. J.* **97** (1943) 255.

[30] V. Springel *et al.*, "Simulating the joint evolution of quasars, galaxies and their large-scale distribution," *Nature* **435** (2005) 629–636, `arXiv:astro-ph/0504097`.

[31] J. F. Navarro, C. S. Frenk, and S. D. M. White, "The Structure of cold dark matter halos," *Astrophys. J.* **462** (1996) 563–575, `arXiv:astro-ph/9508025`.

[32] R. A. Flores and J. R. Primack, "Observational and theoretical constraints on singular dark matter halos," *Astrophys. J. Lett.* **427** (1994) L1–4, `arXiv:astro-ph/9402004`.

[33] B. Moore, "Evidence against dissipationless dark matter from observations of galaxy haloes," *Nature* **370** (1994) 629.

[34] B. Moore, T. R. Quinn, F. Governato, J. Stadel, and G. Lake, "Cold collapse and the core catastrophe," *Mon. Not. Roy. Astron. Soc.* **310** (1999) 1147–1152, `arXiv:astro-ph/9903164`.

[35] D. N. Spergel and P. J. Steinhardt, "Observational evidence for selfinteracting cold dark matter," *Phys. Rev. Lett.* **84** (2000) 3760–3763, `arXiv:astro-ph/9909386`.

[36] D. Richstone *et al.*, "Supermassive black holes and the evolution of galaxies," *Nature* **395** (1998) A14–A19, `arXiv:astro-ph/9810378`.

[37] L. Ferrarese and H. Ford, "Supermassive black holes in galactic nuclei: Past, present and future research," *Space Sci. Rev.* **116** (2005) 523–624, `arXiv:astro-ph/0411247`.

[38] L. Ferrarese and D. Merritt, "A Fundamental relation between supermassive black holes and their host galaxies," *Astrophys. J. Lett.* **539** (2000) L9, `arXiv:astro-ph/0006053`.

[39] N. J. McConnell and C.-P. Ma, "Revisiting the Scaling Relations of Black Hole Masses and Host Galaxy Properties," *Astrophys. J.* **764** (2013) 184, `arXiv:1211.2816 [astro-ph.CO]`.

[40] J. Kormendy and L. C. Ho, "Coevolution (Or Not) of Supermassive Black Holes and Host Galaxies," *Ann. Rev. Astron. Astrophys.* **51** (2013) 511–653, `arXiv:1304.7762 [astro-ph.CO]`.

[41] G. R. Blumenthal, S. M. Faber, J. R. Primack, and M. J. Rees, "Formation of Galaxies and Large Scale Structure with Cold Dark Matter," *Nature* **311** (1984) 517–525.

[42] C. Lacey and S. Cole, "Merger rates in hierarchical models of galaxy formation," *Mon. Not. Roy. Astron. Soc.* **262** no. 3, (1993) 627–649.

[43] Q.-j. Yu and S. Tremaine, "Observational constraints on growth of massive black holes," *Mon. Not. Roy. Astron. Soc.* **335** (2002) 965–976, `arXiv:astro-ph/0203082`.

[44] Y. Dubois, C. Pichon, J. Devriendt, J. Silk, M. Haehnelt, T. Kimm, and A. Slyz, "Blowing cold flows away: the impact of early AGN activity on the formation of a brightest cluster galaxy progenitor," *Mon. Not. Roy. Astron. Soc.* **428** (2013) 2885, arXiv:1206.5838 [astro-ph.CO].

[45] A. Kulier, J. P. Ostriker, P. Natarajan, C. N. Lackner, and R. Cen, "Understanding black hole mass assembly via accretion and mergers at late times in cosmological simulations," *Astrophys. J.* **799** no. 2, (2015) 178, arXiv:1307.3684 [astro-ph.CO].

[46] Y. Dubois, M. Volonteri, and J. Silk, "Black hole evolution – III. Statistical properties of mass growth and spin evolution using large-scale hydrodynamical cosmological simulations," *Mon. Not. Roy. Astron. Soc.* **440** no. 2, (2014) 1590–1606, arXiv:1304.4583 [astro-ph.CO].

[47] M. C. Begelman, M. Volonteri, and M. J. Rees, "Formation of supermassive black holes by direct collapse in pregalactic halos," *Mon. Not. Roy. Astron. Soc.* **370** (2006) 289–298, arXiv:astro-ph/0602363.

[48] L. Mayer, S. Kazantzidis, P. Madau, M. Colpi, T. R. Quinn, and J. Wadsley, "Rapid Formation of Supermassive Black Hole Binaries in Galaxy Mergers with Gas," *Science* **316** (2007) 1874–1877, arXiv:0706.1562 [astro-ph].

[49] L. Mayer, S. Kazantzidis, A. Escala, and S. Callegari, "Direct Formation of Supermassive Black Holes via Multi-Scale Gas Inflows in Galaxy Mergers," *Nature* **466** (2010) 1082, arXiv:0912.4262 [astro-ph.CO].

[50] M. Volonteri, M. Habouzit, and M. Colpi, "The origins of massive black holes," *Nature Rev. Phys.* **3** no. 11, (2021) 732–743, arXiv:2110.10175 [astro-ph.GA].

[51] P. Madau and M. J. Rees, "Massive black holes as Population III remnants," *Astrophys. J. Lett.* **551** (2001) L27–L30, arXiv:astro-ph/0101223.

[52] S. F. Portegies Zwart, H. Baumgardt, P. Hut, J. Makino, and S. L. W. McMillan, "The Formation of massive black holes through collision runaway in dense young star clusters," *Nature* **428** (2004) 724, arXiv:astro-ph/0402622.

[53] M. Freitag, M. A. Gurkan, and F. A. Rasio, "Runaway collisions in young star clusters. 2. Numerical results," *Mon. Not. Roy. Astron. Soc.* **368** (2006) 141–161, arXiv:astro-ph/0503130.

[54] B. Carr and J. Silk, "Primordial Black Holes as Generators of Cosmic Structures," *Mon. Not. Roy. Astron. Soc.* **478** no. 3, (2018) 3756–3775, `arXiv:1801.00672 [astro-ph.CO]`.

[55] S. F. Bramberger, R. H. Brandenberger, P. Jreidini, and J. Quintin, "Cosmic String Loops as the Seeds of Super-Massive Black Holes," *JCAP* **06** (2015) 007, `arXiv:1503.02317 [astro-ph.CO]`.

[56] A. De Rosa *et al.*, "The quest for dual and binary supermassive black holes: A multi-messenger view," *New Astron. Rev.* **86** (2019) 101525, `arXiv:2001.06293 [astro-ph.GA]`.

[57] M. Liska, C. Hesp, A. Tchekhovskoy, A. Ingram, M. van der Klis, and S. Markoff, "Formation of Precessing Jets by Tilted Black-hole Discs in 3D General Relativistic MHD Simulations," *Mon. Not. Roy. Astron. Soc.* **474** no. 1, (2018) L81–L85, `arXiv:1707.06619 [astro-ph.HE]`.

[58] L. Dey *et al.*, "The Unique Blazar OJ 287 and its Massive Binary Black Hole Central Engine," *Universe* **5** no. 5, (2019) 108, `arXiv:1905.02689 [astro-ph.GA]`.

[59] G. D. Quinlan, "The dynamical evolution of massive black hole binaries - I. hardening in a fixed stellar background," *New Astron.* **1** (1996) 35–56, `arXiv:astro-ph/9601092`.

[60] D. Merritt, "Loss-cone Dynamics," *Class. Quant. Grav.* **30** (2013) 244005, `arXiv:1307.3268 [astro-ph.GA]`.

[61] F. M. Khan, K. Holley-Bockelmann, P. Berczik, and A. Just, "Supermassive Black Hole Binary Evolution in Axisymmetric Galaxies: The final parsec problem is not a problem," *Astrophys. J.* **773** (2013) 100, `arXiv:1302.1871 [astro-ph.GA]`.

[62] E. Vasiliev, F. Antonini, and D. Merritt, "The final-parsec problem in nonspherical galaxies revisited," *Astrophys. J.* **785** (2014) 163, `arXiv:1311.1167 [astro-ph.GA]`.

[63] B. Kocsis and A. Sesana, "Gas driven massive black hole binaries: signatures in the nHz gravitational wave background," *Mon. Not. Roy. Astron. Soc.* **411** (2011) 1467, `arXiv:1002.0584 [astro-ph.CO]`.

[64] F. G. Goicovic, A. Sesana, J. Cuadra, and F. Stasyszyn, "Infalling clouds on to supermassive black hole binaries – II. Binary evolution and the final parsec problem,"

Mon. Not. Roy. Astron. Soc. **472** no. 1, (2017) 514–531, arXiv:1602.01966 [astro-ph.HE].

[65] F. G. Goicovic, C. Maureira-Fredes, A. Sesana, P. Amaro-Seoane, and J. Cuadra, "Accretion of clumpy cold gas onto massive black hole binaries: a possible fast route to binary coalescence," *Mon. Not. Roy. Astron. Soc.* **479** no. 3, (2018) 3438–3455, arXiv:1801.04937 [astro-ph.HE].

[66] L. Z. Kelley, L. Blecha, and L. Hernquist, "Massive Black Hole Binary Mergers in Dynamical Galactic Environments," *Mon. Not. Roy. Astron. Soc.* **464** no. 3, (2017) 3131–3157, arXiv:1606.01900 [astro-ph.HE].

[67] M. Bonetti, A. Sesana, E. Barausse, and F. Haardt, "Post-Newtonian evolution of massive black hole triplets in galactic nuclei – III. A robust lower limit to the nHz stochastic background of gravitational waves," *Mon. Not. Roy. Astron. Soc.* **477** no. 2, (2018) 2599–2612, arXiv:1709.06095 [astro-ph.GA].

[68] T. Ryu, R. Perna, Z. Haiman, J. P. Ostriker, and N. C. Stone, "Interactions between multiple supermassive black holes in galactic nuclei: a solution to the final parsec problem," *Mon. Not. Roy. Astron. Soc.* **473** no. 3, (2018) 3410–3433, arXiv:1709.06501.

[69] S. Chen, A. Sesana, and C. J. Conselice, "Constraining astrophysical observables of Galaxy and Supermassive Black Hole Binary Mergers using Pulsar Timing Arrays," *Mon. Not. Roy. Astron. Soc.* **488** no. 1, (2019) 401–418, arXiv:1810.04184 [astro-ph.GA].

[70] G. Girelli, L. Pozzetti, M. Bolzonella, C. Giocoli, F. Marulli, and M. Baldi, "The stellar-to-halo mass relation over the past 12 Gyr: I. Standard ΛCDM model," *Astron. Astrophys.* **634** (2020) A135, arXiv:2001.02230 [astro-ph.CO].

[71] F. Zwicky, "Die Rotverschiebung von extragalaktischen Nebeln," *Helv. Phys. Acta* **6** (1933) 110–127.

[72] V. C. Rubin and W. K. Ford, Jr., "Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions," *Astrophys. J.* **159** (1970) 379–403.

[73] **Planck** Collaboration, N. Aghanim *et al.*, "Planck 2018 results. VI. Cosmological parameters," *Astron. Astrophys.* **641** (2020) A6, arXiv:1807.06209 [astro-ph.CO]. [Erratum: Astron.Astrophys. 652, C4 (2021)].

[74] S. Tulin and H.-B. Yu, "Dark Matter Self-interactions and Small Scale Structure," *Phys. Rept.* **730** (2018) 1–57, `arXiv:1705.02358 [hep-ph]`.

[75] M. Kaplinghat, S. Tulin, and H.-B. Yu, "Dark Matter Halos as Particle Colliders: Unified Solution to Small-Scale Structure Puzzles from Dwarfs to Clusters," *Phys. Rev. Lett.* **116** no. 4, (2016) 041302, `arXiv:1508.03339 [astro-ph.CO]`.

[76] A. Klypin, G. Yepes, S. Gottlober, F. Prada, and S. Hess, "MultiDark simulations: the story of dark matter halo concentrations and density profiles," *Mon. Not. Roy. Astron. Soc.* **457** no. 4, (2016) 4340–4359, `arXiv:1411.4001 [astro-ph.CO]`.

[77] S. L. Shapiro and V. Paschalidis, "Self-interacting dark matter cusps around massive black holes," *Phys. Rev. D* **89** no. 2, (2014) 023506, `arXiv:1402.0005 [astro-ph.CO]`.

[78] D. Merritt, "Single and binary black holes and their influence on nuclear structure," in *Carnegie Observatories Centennial Symposium. 1. Coevolution of Black Holes and Galaxies*. 1, 2003. `arXiv:astro-ph/0301257`.

[79] T. Nakano and J. Makino, "On the cusp around central black holes in luminous elliptical galaxies," *Astrophys. J. Lett.* **525** (1999) L77, `arXiv:astro-ph/9906131`.

[80] P. Ullio, H. Zhao, and M. Kamionkowski, "A Dark matter spike at the galactic center?," *Phys. Rev. D* **64** (2001) 043504, `arXiv:astro-ph/0101481`.

[81] Y. N. Eroshenko, "Dark matter density spikes around primordial black holes," *Astron. Lett.* **42** no. 6, (2016) 347–356, `arXiv:1607.00612 [astro-ph.HE]`.

[82] L. Sadeghian, F. Ferrer, and C. M. Will, "Dark matter distributions around massive black holes: A general relativistic analysis," *Phys. Rev. D* **88** no. 6, (2013) 063522, `arXiv:1305.2619 [astro-ph.GA]`.

[83] D. Morin, *Classical Mechanics With Problems and Solutions*. Cambridge, 2004.

[84] P. C. Peters and J. Mathews, "Gravitational radiation from point masses in a Keplerian orbit," *Phys. Rev.* **131** (1963) 435–439.

[85] P. S. Cole, A. Coogan, B. J. Kavanagh, and G. Bertone, "Measuring dark matter spikes around primordial black holes with Einstein Telescope and Cosmic Explorer," *Phys. Rev. D* **107** no. 8, (2023) 083006, `arXiv:2207.07576 [astro-ph.CO]`.

[86] P. Jangra, B. J. Kavanagh, and J. M. Diego, "Impact of dark matter spikes on the merger rates of Primordial Black Holes," *JCAP* **11** (2023) 069, `arXiv:2304.05892 [astro-ph.CO]`.

[87] P. C. Peters, "Gravitational Radiation and the Motion of Two Point Masses," *Phys. Rev.* **136** (1964) B1224–B1232.

[88] J. Binney and S. Tremaine, *Galactic Dynamics*. Princeton University Press, 2008.

[89] S. Burke-Spolaor *et al.*, "The Astrophysics of Nanohertz Gravitational Waves," *Astron. Astrophys. Rev.* **27** no. 1, (2019) 5, `arXiv:1811.08826 [astro-ph.HE]`.

[90] A. Gualandris and D. Merritt, "Ejection of Supermassive Black Holes from Galaxy Cores," *Astrophys. J.* **678** (2008) 780, `arXiv:0708.0771 [astro-ph]`.

[91] R. A. Hulse, "The discovery of the binary pulsar," *Rev. Mod. Phys.* **66** (1994) 699–710.

[92] L. S. Finn and K. S. Thorne, "Gravitational waves from a compact star in a circular, inspiral orbit, in the equatorial plane of a massive, spinning black hole, as observed by LISA," *Phys. Rev. D* **62** (2000) 124021, `arXiv:gr-qc/0007074`.

[93] **LISA** Collaboration, P. A. Seoane *et al.*, "Astrophysics with the Laser Interferometer Space Antenna," *Living Rev. Rel.* **26** no. 1, (2023) 2, `arXiv:2203.06016 [gr-qc]`.