PHYSICS LETTERS B

Letter

Final parsec problem of black hole mergers and ultralight dark matter

Hyeonmo Koo [a,1], Dongsu Bak [a,b,1], Inkyu Park [a,b], Sungwook E. Hong [c,d], Jae-Weon Lee [e],*

[a] *Physics Department, University of Seoul, Seoul 02504, Korea*
[b] *Natural Science Research Institute, University of Seoul, Seoul 02504, Korea*
[c] *Korea Astronomy and Space Science Institute, Daejeon 34055, Korea*
[d] *Astronomy Campus, University of Science and Technology, Daejeon 34055, Korea*
[e] *Department of Electrical and Electronic Engineering, Jungwon University, Chungbuk 28024, Korea*

ARTICLE INFO

Editor: P. Brax

ABSTRACT

When two galaxies merge, they often produce a supermassive black hole binary (SMBHB) at their center. Numerical simulations with stars and cold dark matter show that SMBHBs typically stall out at a distance of a few parsecs apart and take billions of years to coalesce. This is known as the final parsec problem. We suggest that ultralight dark matter (ULDM) halos around SMBHBs can generate dark matter waves due to dynamical friction. These waves can effectively carry away orbital energy from the black holes, rapidly driving them together. To test this hypothesis, we performed numerical simulations of black hole binaries inside ULDM halos. Due to gravitational cooling and quasi-normal modes, the loss-cone problem can be avoided. The decay time scale gives lower bounds on masses of the ULDM particles and SMBHBs comparable to observational data. Our results imply that ULDM waves can lead to the rapid orbital decay of black hole binaries.

1. Introduction

The mystery of supermassive black hole (SMBH) growth is one of the unsolved problems in astronomy. When two galaxies merge, they can form a supermassive black hole binary (SMBHB) at their center. However, numerical simulations show that SMBHBs typically become stuck at a distance of a few parsecs apart, and can take billions of years to merge. At this distance the density of the stars and gas near the SMBHB is too low for dynamical friction to be efficient, while the loss of orbital energy of the SMBHB due to gravitational waves is only efficient for distances less than $\mathcal{O}(10^{-2})$ pc [1,2]. Furthermore, the loss cone is depleted as the black holes (BHs) approach each other. This difficulty is known as the final parsec problem. The gravitational wave background recently observed by NANOGrav [3] is usually attributed to efficient SMBH mergers. This fact deepens the mystery. Proposed solutions to the final parsec problem often involve bringing in extra matter, such as additional stars or another BH interacting with the black hole binaries (BHBs) to help them merge.

Ultralight (fuzzy) dark matter (ULDM) [4–9] is a promising alternative to cold dark matter (CDM), as it has the potential to solve some of the small-scale issues of CDM such as the missing satellite problem, the plane of satellite galaxies problem, and the core-cusp problem [10–13]. In this model, the ULDM is in a Bose-Einstein condensate state of ultralight scalar particles with a typical mass $m \gtrsim 10^{-22}$eV. ULDM can be described with a macroscopic wave function ψ and the uncertainty principle suppresses too many small-scale structure formations. Beyond the galactic scale, the ULDM behaves like CDM and hence naturally solves the problems of CDM. This model has also been shown to be able to explain a wide range of astrophysical observations, including the rotation curves of galaxies [8,14,15], and the large-scale structures of the universe [16]. Recently, there has been a growing interest in the interactions between BHs and ULDM halos surrounding them [17], as these interactions could change the patterns of gravitational waves generated by BHBs.

In this letter, we suggest that, inside ULDM halos (for a review, see [18–23]), BHBs can generate dark matter (DM) waves due to dynamical friction (DF) and gravitational cooling (GC). The dynamical friction of ULDM [24] refers to the frictional force that arises from the gravitational interaction between a moving celestial object and ULDM wakes generated by the object. On the other hand, gravitational cooling [25] is a mechanism for relaxation by ejecting ULDM waves carrying out excessive kinetic energy and momentum. These waves can effectively carry away orbital energy

---

from the BHBs, rapidly driving them together and possibly solving the final parsec problem. To test this hypothesis, we perform numerical simulations of BHBs at $\mathcal{O}(\text{pc})$ separation in an ULDM soliton representing a central part of the galactic halos. Our results provide some evidence that DM waves could lead to the rapid orbital decay of the BHs.

Although collisionless CDM can temporally exert a larger DF force to SMBHBs compared to ULDM [26,27], its relaxation time to refill the phase space, given by

$$t_{\text{relax}}(r) = \frac{N_{CDM}}{8 \ln N_{CDM}} \left(\frac{r^3}{G M_{\text{tot}}(r)} \right)^{\frac{1}{2}}, \tag{1}$$

can easily exceed the age of the universe [28]. In this expression, N_{CDM} is the number of CDM particles, and $M_{\text{tot}}(r)$ is the total mass within a radius r from the center of a host galaxy, with $r \simeq 1\,\text{pc}$ in our study. As a result, the collisionless CDM model encounters the same loss-cone depletion problem as stars, making it difficult to resolve the final parsec problem. On the other hand, the collective motion of ULDM particles can help replenish the phase space after interacting with SMBHBs. There are several studies on the effects of DF by ULDM on SMBHBs. These studies typically focus either on the entrance states into DM halos (at $r \gtrsim \text{kpc}$) [27] or the final merging states (at $r \ll 0.01\,\text{pc}$) [29].

The aim of this paper is to numerically show fast orbital decays of SMBHBs at $r \simeq 1\,\text{pc}$ due to ULDM waves, whose dynamics is described by the Schrödinger-Poisson equations. The structure of this paper is organized as follows. In Sec. 2, we explain our numerical simulation for studying the orbital decay of BHBs. Then, Sec. 3 and 4 show the numerical results from Sec. 2 and their theoretical analyses, respectively. Finally, Sec. 5 summarizes our work.

2. Simulation

In our simulation, we shall treat BHs as point particles with the Newtonian gravity. The post-Newtonian correction to the gravitational potential of a BH at a distance of r from the BH is on the order of r_s/r, where $r_s = 2GM_{\text{bh}}/c^2$ is the Schwarzschild radius of the black hole. For $M_{\text{bh}} = 10^8 M_{\odot}$ and $r = 0.01\,\text{pc}$, this correction is $\mathcal{O}(10^{-3})$. Therefore, the Newtonian approximation of the gravitational field is good enough for our calculation.

The ULDM system, interacting gravitationally with N-particles (BHs), is governed by the Schrödinger-Poisson equations

$$i\hbar \frac{\partial}{\partial t} \psi(\mathbf{x},t) = -\frac{\hbar^2}{2m} \nabla^2 \psi(\mathbf{x},t) + m \left[V_{\text{U}}(\mathbf{x},t) + V_{\text{N}}(\mathbf{x},t) \right] \psi(\mathbf{x},t) \tag{2}$$

$$\nabla^2 V_{\text{U}}(\mathbf{x},t) = 4\pi G |\psi|^2(\mathbf{x},t) \tag{3}$$

with the mass m of ultralight DM particles and the Newton constant G. The wave function $\psi(\mathbf{x},t)$ for ULDM is a complex function of spacetime normalized as $\int d\mathbf{x} |\psi|^2 = M$ with M denoting the total ULDM mass such that the corresponding ULDM mass density is given by $\rho = |\psi|^2$. V_{U} is the potential from the ULDM mass distribution and V_{N} for the N-particles where k-th ($k = 1, 2, \cdots, N$) particle is located at $\mathbf{x}_k(t)$ with mass M_k. (In this work, $N = 2$.) Then, the N-body potential is given by $V_{\text{N}}(\mathbf{x},t) = \sum_{k=1}^{N} V_k(\mathbf{x},t)$ with $V_k(\mathbf{x},t) = -GM_k/|\mathbf{x} - \mathbf{x}_k(t)|$. With these potentials, the remaining N-body dynamics is described by

$$\ddot{\mathbf{x}}_k(t) = -\nabla V_{\text{U}}(\mathbf{x}_k(t),t) - \sum_{l \neq k} \nabla V_l(\mathbf{x}_k(t),t). \tag{4}$$

The total combined system enjoys two exact scaling relations. One is related to the scaling of the total ULDM mass M and the particle mass M_k: $[M, M_k, m, t, \mathbf{x}, V_{\text{U}}, V_k, \psi] \rightarrow [\kappa M, \kappa M_k, m, \kappa^{-2}t, \kappa^{-1}\mathbf{x}, \kappa^2 V_{\text{U}}, \kappa^2 V_k, \kappa^2 \psi]$. The other is the scaling of the ULDM particle mass m with the transformation $[M, M_k, m, t, \mathbf{x}, V_{\text{U}}, V_k, \psi] \rightarrow [M, M_k, \lambda m, \lambda^{-3}t, \lambda^{-2}\mathbf{x}, \lambda^2 V_{\text{U}}, \lambda^2 V_k, \lambda^3 \psi]$. Under these two transformations, the forms of the full equations (2), (3) and (4) remain invariant. These two scaling relations will be useful in the theoretical analysis in Sec. 4.

In our simulation, we use the Python pseudo-spectral solver, PyUltraLight2 [30], which is publicly available. This package, as in its prior version [31], solves the Schrödinger-Poisson equations but also incorporates the above-mentioned N-body dynamics in a full-fledged form. In this package, the system will be placed in a box of size L with a periodic boundary condition $x_i \sim x_i + L$. Below we take the box size to be $L = 40\text{pc}$, which is large enough to neglect the effect of the periodic boundary condition. The number of grid size in each direction is taken to be 500, and then the resulting spatial resolution becomes $\Delta x = 0.08\text{pc}$.

The default system value of the time resolution is set to be $\Delta t_d = 1.039\, m_{21}^{-1}\text{yr}$ with the ULDM particle mass ratio $m_{21} = m/(10^{-21}\text{eV})$, whose validity is rather well tested [30]. One may introduce the resolution factor (RF) defined by $\text{RF} = \Delta t/\Delta t_d$, where Δt denotes an actual choice of resolution time used for the simulation. Below we use the two choices: Δt_1 and Δt_2 respectively with $\text{RF1} \approx 0.9821$ and $\text{RF2} \approx 1.9249$. As will be discussed further below, RF1 is only used to see the validity of our simulation in Fig. 4. To obtain the time series data of separation and velocity of BHBs and others, we use RF2, which is mainly due to our limitation of computing resources. We have, however, checked that the relative difference of the separations always remains within 1% during the entire simulation time for one value of the BH mass. For the further discussions on the similar choices of time resolution, see [31] and also [10,32].

In our simulation below, we set the ULDM particle mass as $m = 10^{-21}\text{eV}$ ($m_{21} = 1$) and place a central ULDM halo of mass $M_s = 10^9 M_{\odot}$ (hence $M = M_s$) at the center of the box to represent a galactic DM core. This mass value is somewhat larger than the typical value ($m = 10^{-22}\text{eV}$), but there are recent observations favoring the heavier mass [33,34]. To clearly see the orbital decay, we chose a rather massive and compact DM soliton than usual one, mainly due to limited simulation resolution and computing resources. (This compact region can also be interpreted as the innermost part of the galactic dark matter cloud that surrounds BHs.) In Sec. 4, we theoretically show that a fast orbital decay can be achieved with a more realistic DM density. The central region will be modeled by the ground state (soliton) configuration of the Schrödinger-Poisson system in the absence of BHs or any other particles. The central DM density of the stable soliton core is given by [23]

$$\rho_0 = 7.05 \times 10^6 \left(\frac{m}{10^{-21}\text{ eV}} \right)^6 \left(\frac{M_s}{10^9 M_{\odot}} \right)^4 M_{\odot}\, \text{pc}^{-3}. \tag{5}$$

For $m = 10^{-21}$ eV and $M_s = 10^9 M_{\odot}$ we adopted, this equation gives a density larger than the typical central density of the local dwarf galaxies. However, almost all SMBHBs are found in massive galaxies, where the central dark matter density can be much higher than that of dwarf galaxies and often has a steeper profile. For example, the recent S2 orbital data from our galaxy provides an upper bound for the enclosed DM mass $M(r)$
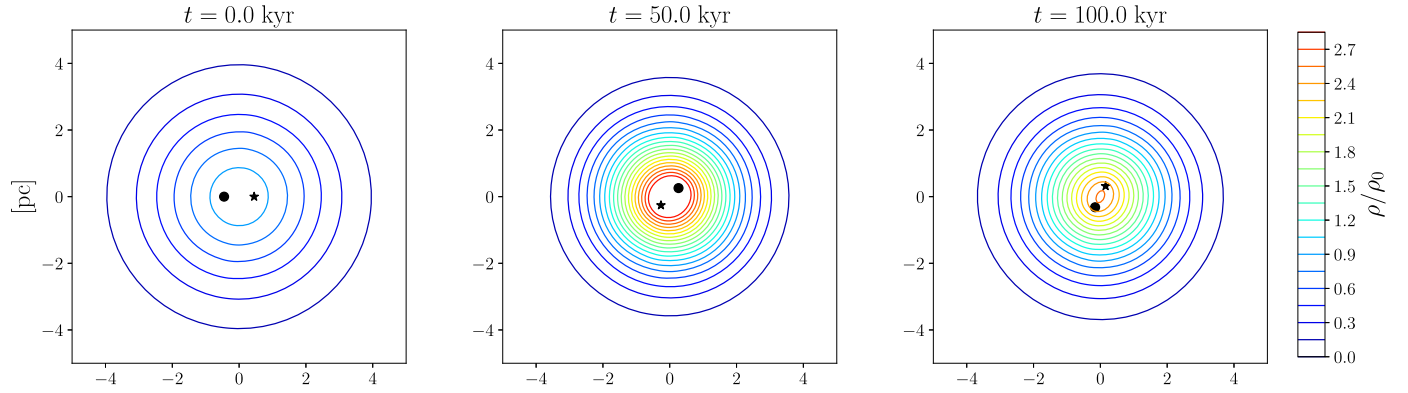
Fig. 1. Time evolution of the SMBHB of mass $M_{\rm bh} = 0.6 \times 10^8 {\rm M}_\odot$ inside ULDM halo of mass $M_s = 10^9 {\rm M}_\odot$. Colored lines represent the contour lines of ULDM density, where ρ_0 denotes the initial central density of an isolated ground-state ULDM halo system, as given in (5). The black dot and star in each diagram stand for our BHB embedded. See Fig. 2 for the orbital decay patterns. The BHs are revolving about 30 times within several periods of central density oscillations, whose details are not fully shown in this figure.
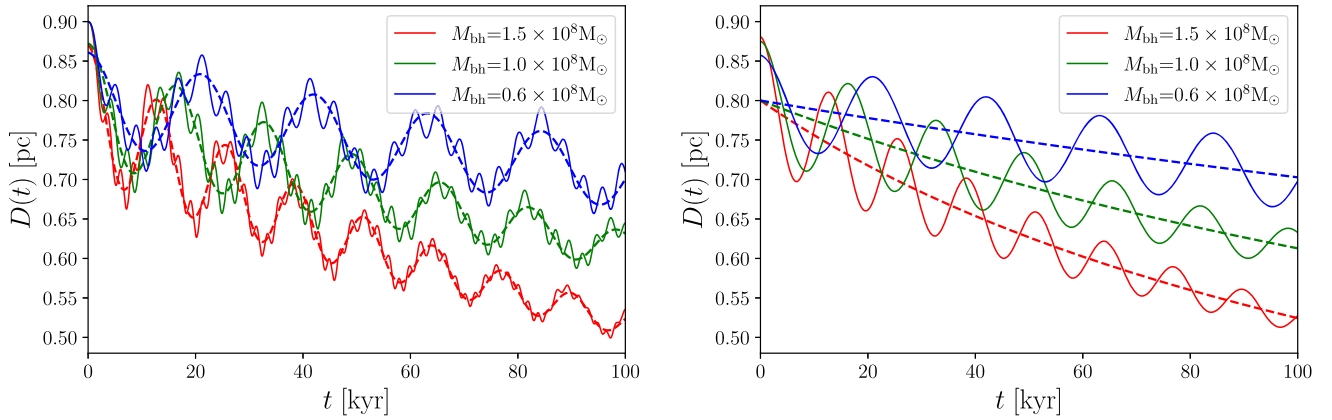

Fig. 2. Time evolution of BHB separation $D(t)$ for $M_{\rm bh} = (0.6, 1.0, 1.5) \times 10^8 {\rm M}_\odot$ (blue, green, and red lines). Left: simulated $D(t)$ (solid) and its full fit (dotted). Right: full fit of $D(t)$ (solid) and the *averaged* fit $\bar{D}(t)$ (dotted). See the texts for the definitions of full fit and averaged fit.

within the S2 orbit, which is about 0.1% of the SMBH mass ($M_{\rm bh} \simeq 4 \times 10^6 {\rm M}_\odot$) [35]. It can be translated into a bound $\rho_0 \lesssim 10^{13} {\rm M}_\odot {\rm pc}^{-3}$, which is unacceptably high for a core profile if we saturate the bound. Therefore, it is reasonable to expect that the actual DM density near central SMBHBs in a massive galaxy has a spike-like profile $\rho(r) \propto r^{-\gamma_{\rm sp}}$ with $1 < \gamma_{\rm sp} < 2$ rather than a wide and flat core profile [36,37]. Thus, we assume that our initial soliton DM distribution mimics a dense central region in a core rather than the core itself.

By analyzing the kinematics of a clockwise-rotating disk of stars and the S2 orbit in our galaxy the constraint $M(r = 0.3 {\rm pc}) \lesssim 10^6 {\rm M}_\odot$ is obtained for ULDM spikes [38]. A similar bound $M(r = 1 {\rm pc}) \lesssim 10^5 {\rm M}_\odot$ is obtained for collisionless DM models [39]. Therefore, for our theoretical analysis we will take $\rho(r = 1 {\rm pc}) \lesssim 10^5 {\rm M}_\odot {\rm pc}^{-3}$ as a conservative observational upper bound for the DM density at 1 pc in this paper. (See Fig. 5.)

We place two BHs of the same mass $M_{\rm bh}$ at $(\pm 0.45 {\rm pc}, 0, 0)$ respectively. To demonstrate the orbital decay of BHB, we generate 91 samples with uniformly varying $M_{\rm bh}$ from $0.6 \times 10^8 {\rm M}_\odot$ to $1.5 \times 10^8 {\rm M}_\odot$. In each sample, the initial velocity of each BH is chosen such that the corresponding initial orbit becomes nearly circular and the total simulation time set to be $T_{\rm sim} = 100 \, m_{21}^{-1}$ kyr.

3. Numerical results

Fig. 1 shows three snapshots of ULDM profiles with $M_{\rm bh} = 0.6 \times 10^8 {\rm M}_\odot$. The initial profile is depicted on the left panel, the $t = 50 {\rm kyr}$ on the middle, and the $t = 100 {\rm kyr}$ on the right. As the two BHs are revolving around each other, their movement stirs the surrounding ULDM halos and generates ULDM waves. The central region of the ULDM profile becomes elliptical, indicating a dipole-like perturbation by the SMBHB. Part of these waves escape the central region while carrying away the kinetic energy of the BHB, which is nothing but the gravitational cooling effect.

The time evolution of the binary separation $D(t)$ is our main concern here and depicted on the left panel of Fig. 2 for $M_{\rm bh} = (0.6, 1.0, 1.5) \times 10^8 {\rm M}_\odot$. Note that our initial ULDM halo configuration is not stationary since the extra interaction, due to the BHB, was initially not taken into account. The presence of BHB subsequently makes some ULDM halos attracted toward the BHs, and drives the ULDM halo system to oscillate radially together with the oscillation of the central ULDM density. This will be the major deriving force behind yet another gravitational cooling of the entire system.

As time progresses, the initial collapse will be quickly settled down with regular oscillations. The time evolution of $D(t)$ in general exhibits the slow and fast oscillations with a slowly decaying average $\bar{D}(t)$. Firstly, the fast oscillation is simply due to the slight eccentricity of the orbit and will be ignored in the following analysis. Next, the slow oscillation mainly follows from the global quasi-normal-mode [40] breathing of halo system together with the spherical central density oscillation, which is closely related to the gravitational cooling of the overall system. If there is no central BH, the time scale of this oscillation is of order of $\hbar^3/(m^3 G^2 M_s^2)$, where M_s is the mass of the soliton. In our case, SMBHB can slightly change the time scale ($T_{\rm slow}$ below). On the other hand, the decaying average $\bar{D}(t)$ seems to be related to another quasi-normal mode with a time scale τ_g from
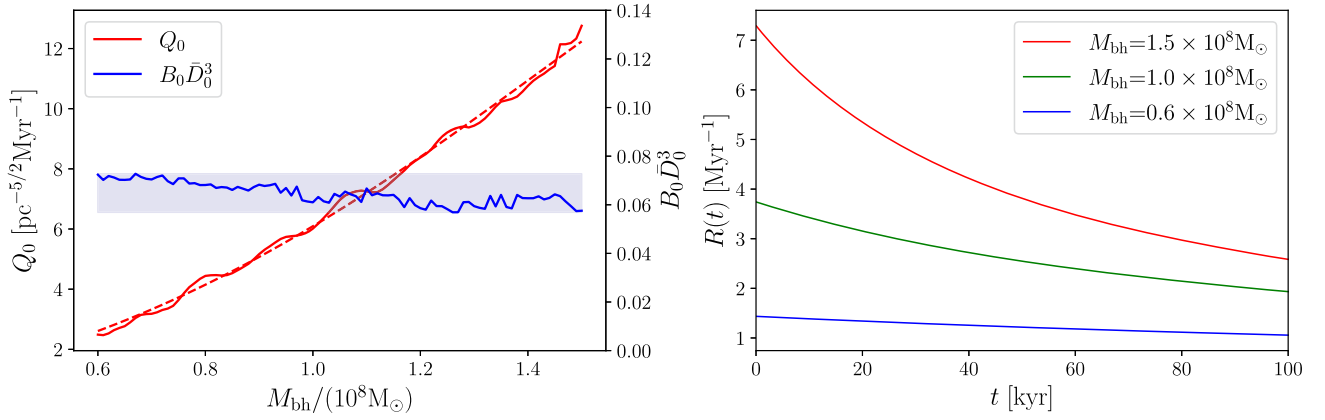
Fig. 3. Left: fitting parameters Q_0 (red line) and $B_0 \bar{D}_0^3$ (blue line) as a function of M_{bh} derived from the simulations for $M_{bh} = 0.6 - 1.5 \times 10^8 M_\odot$. Red dotted line: the best-fit of $Q_0(M_{bh})$ from the theoretical model described in Sec. 4. Light-blue shaded region emphasizes that $B_0 \bar{D}_0^3$ is nearly constant over M_{bh}. Right: time evolution of the decay rate of BHB separation $R(t)$ for $M_{bh} = (0.6, 1.0, 1.5) \times 10^8 M_\odot$ (blue, green, and red lines).

the local dipole perturbation by the rotating SMBHB at the center rather than by the global oscillation of the soliton. This mode is non-spherical and can carry away the angular momentum of the SMBHB (see Sec. 4 for details).

In the remainder of this section, we shall focus on extracting the *averaged* orbital decay function $\bar{D}(t)$, largely independent of the slow global oscillations. However, there is no well-defined procedure for such averaging, since the oscillations in Fig. 2 are in their nonlinear regime. Ignoring the fast oscillation, we try the following fit function

$$D(t) = \bar{D}(t) + A_s \, e^{-\alpha_s t} \cos(\omega t - \phi), \tag{6}$$

where the *averaged* function $\bar{D}(t)$ is defined by

$$\bar{D}(t) = \bar{D}_0 \left(1 + \frac{5}{2} Q_0 \bar{D}_0^{\frac{5}{2}} (t - t_0) \right)^{-\frac{2}{5}} \tag{7}$$

with the initial time $t_0 = 0$ in our case (see the next section for the derivation of equation (7)). It is clear that $1/(Q_0 \bar{D}_0^{\frac{5}{2}})$ gives an initial decay time scale.

We perform the $D(t)$ fitting with the fitting parameters, $(Q_0, \bar{D}_0, A_s, \alpha_s, \omega, \phi)$. For each of $M_{bh} = (0.6, 1.0, 1.5) \times 10^8 \, M_\odot$, we depict the fit function of $D(t)$ as a dotted curve on the left panel and as a solid curve on the right panel of Fig. 2. We have also added $\bar{D}(t)$ as a dotted curve for the comparison on the right panel. One finds $\bar{D}_0 \simeq 0.8$ pc there. The Q_0 as a function of M_{bh} is depicted on the left panel of Fig. 3. Using the parameters and the fit function $\bar{D}(t)$, one can also obtain the decay rate given by

$$R(t) \equiv -\frac{d \log \bar{D}}{dt} = \frac{Q_0 \bar{D}_0^{\frac{5}{2}}}{1 + \frac{5}{2} Q_0 \bar{D}_0^{\frac{5}{2}} (t - t_0)}, \tag{8}$$

which is plotted on the right panel of Fig. 3 for $M_{bh} = (0.6, 1.0, 1.5) \times 10^8 \, M_\odot$.

In the above fit, we also obtain the slow oscillation period $T_{slow} \equiv 2\pi/\omega$, which may be further fit with the formula

$$T_{slow} = \frac{q_0 \hbar^3}{m^3 G^2 (M_s + 2\gamma_s M_{bh})^2}, \tag{9}$$

where q_0 and γ_s are the fitting parameters. We numerically find that $q_0 = 86.53 \pm 0.29$ and $\gamma_s = 1.979 \pm 0.012$. Without the BHB, this time scale is well known (see for instance [32]). Here the main difference comes from the additional BH mass contribution weighted by the γ_s factor. In principle, the system may involve multiple quasi-normal modes, which are characterized by the parameters (q_0, γ_s). Especially, the quasi-normal mode caused by the orbital motion of BHB may involve different parameters as discussed in the next section.

Similarly to equation (6), the relative velocity of one BH with respect to the other can be fitted with the fit function given by

$$v_{rel}(t) = \bar{v}_{rel}(t) + \tilde{A}_s \, e^{-\tilde{\alpha}_s t} \cos(\tilde{\omega} t - \tilde{\phi}) \tag{10}$$

with $\bar{v}_{rel}(t) = 2\sqrt{GM_{bh}(1 + B_0 \bar{D}^3)/(2\bar{D})}$, which is the twice of the rotation velocity of each BH including the ULDM contribution parameterized by $M_{bh} B_0 \bar{D}^3 = 4\delta M_s$ (see the next section for the details). The fitting parameter $B_0 \bar{D}_0^3$ can be used for the consistency of our approximation in the next section and is depicted on the left panel of Fig. 3.

To demonstrate the stability of our simulation, we present the time evolution of ULDM energy, BH energy, and total energy in the left panel of Fig. 4. Additionally, we depict the relative variation of the ULDM mass, expressed as $\Delta M/M(0)$, where $M(0)$ denotes the initial ULDM mass. It is evident that the energy of BBHs is transferred to ULDM energy, while the total mass of ULDM remains conserved throughout the simulation. The right panel of the figure reveals that the virial ratio is about one half [40].
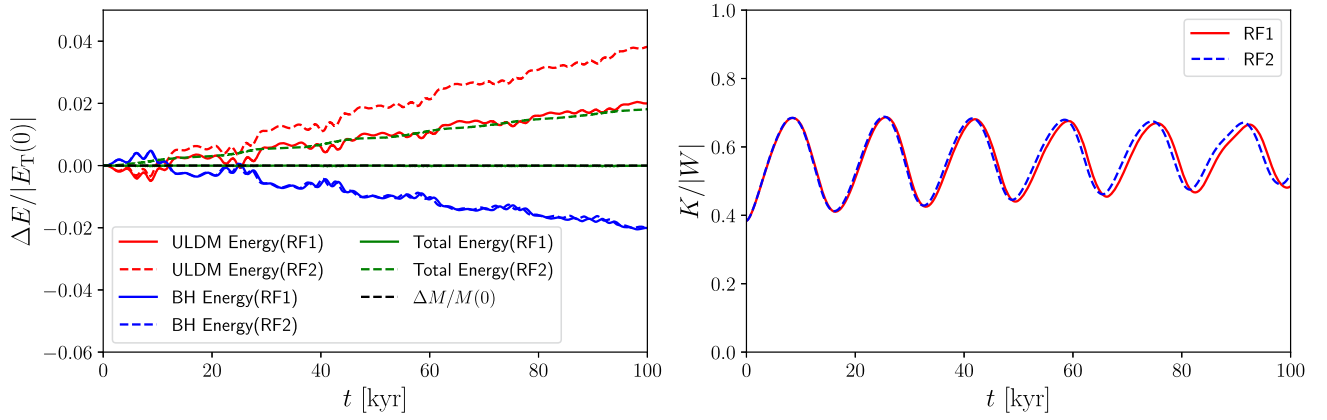
Fig. 4. Left: Time evolution of the ULDM energy, the BH energy, and the total energy for RF1 and RF2 (with $M_{\text{bh}} = 1.0 \times 10^8 M_\odot$). The black dashed line represents relative variation of the total mass of ULDM $\Delta M / M(0)$, which is negligible. Right: The ratio of the total kinetic energy and the absolute value of the total potential energy.

4. Theoretical analysis

In this section, we provide a rough estimate of the orbital decay timescale of a rotating SMBHB. This estimate is mainly based on dimensional analysis and intended to provide guidance for more accurate future calculations, because the system is highly nonlinear. Although it is difficult to clearly distinguish the contribution of dynamical friction from that of gravitational cooling, it turns out that, for our orbital decay problem, the dynamical friction (DF) is dominant over the gravitational cooling (GC) effect, especially in the small separation limit. This may be argued as follows.

We begin with the case of the DF contribution of a BH immersed in the ULDM halo given by

$$F_{\text{DF}} = 4\pi\rho \left(\frac{G M_{\text{bh}}}{v} \right)^2 C \,, \tag{11}$$

where ρ is the ULDM mass density and C is the DF coefficient. For the present context, this coefficient is rather well understood. See, for instance, equation (D8) of [23] for its details. There was introduced an effective cutoff radius scale r_{eff} of the surrounding spherical region over which the DF perturbation is effective. This scale may be taken as the smaller of the orbital radius $r = \bar{D}/2$ and the soliton size scale. In our case, since the orbital size is smaller than the soliton size, we take $r_{\text{eff}} = \alpha_{\text{DF}} \bar{D}/2$ where α_{DF} is an $\mathcal{O}(1)$ numerical constant which may be estimated by our simulation. When $k\, r_{\text{eff}} \ll 1$ with $k = mv/\hbar$, C can be approximated by [23,30]

$$C = \frac{1}{3}(k\, r_{\text{eff}})^2 = \frac{\alpha_{\text{DF}}^2}{12} \left(\frac{mv\bar{D}}{\hbar} \right)^2 \,. \tag{12}$$

It is well known that the value of C for ULDM is smaller than that for CDM. However, the wave nature (quasi-normal modes) of ULDM helps to replenish the loss cone. Figs. 1 and 2 illustrate rapid redistribution of central density of ULDM over time, unlike CDM. This might help the fast orbital decay with ULDM.

Due to the DF, the black holes lose their angular momentum $L_{\text{bh}} = M_{\text{bh}} r v$. From equations (11), (12), and $v \simeq \sqrt{G M_{\text{bh}}/2\bar{D}}$, we obtain the decay time scale [23,41],

$$\tau \equiv \frac{L_{\text{bh}}}{r F_{\text{DF}}} = \frac{1}{8\pi C \rho} \sqrt{\frac{M_{\text{bh}}}{2G \bar{D}^3}} = \frac{3\hbar^2}{\sqrt{2}\pi G^{3/2}} \frac{1}{\alpha_{\text{DF}}^2 m^2 \rho \sqrt{\bar{D}^5 M_{\text{bh}}}}$$

$$= \left(\frac{8.6 \times 10^{-3}}{\alpha_{\text{DF}}^2} \right) \left(\frac{m}{10^{-21}\,\text{eV}} \right)^{-2} \left(\frac{M_{\text{bh}}}{10^8 M_\odot} \right)^{-\frac{1}{2}} \left(\frac{\rho}{10^5 M_\odot/\text{pc}^3} \right)^{-1} \left(\frac{\bar{D}}{\text{pc}} \right)^{-\frac{5}{2}} \text{Gyr} \,, \tag{13}$$

where we later obtain $\alpha_{\text{DF}} \simeq 0.424$ from our simulation in Fig. 3 (see equation (17)) and ρ is the average DM density at $r = \bar{D}/2$.

The above equation shows that DF by ULDM can lead to an efficient orbital decay of BHBs within Gyr for moderate parameter ranges. To show that this rapid decay is not due to unnatural high dark matter density, we plot the allowed region in the parameter plane (m, ρ) for $\tau < 1$ Gyr in Fig. 5 with $\alpha_{\text{DF}} = 0.424$. From the graph one can see that the heavier m is, the lower ρ can be, and there are plausible parameter regions with $m \gtrsim 10^{-22}$ eV with relatively low DM density compatible with observations.

For the lowest BH mass observed ($M_{\text{bh}} \simeq 10^6 M_\odot$), which can lead to the longest τ, one can get $m > 6.9 \times 10^{-22}$ eV, if $\rho < 10^5 M_\odot/\text{pc}^3$. Conversely, if we choose $m < 6.9 \times 10^{-22}$ eV and $\rho < 10^5 M_\odot/\text{pc}^3$, we can find a bound for the black hole mass $M_{\text{bh}} > 10^6 M_\odot$. This means that BHBs with mass smaller than $10^6 M_\odot$ cannot merge efficiently to grow to supermassive black holes with $m \simeq 10^{-22}$ eV, which is often suggested to solve the small scale issues. This fact might explain why supermassive black holes typically have mass $M_{\text{bh}} > 10^6 M_\odot$.

There is another constraint on m if we use the soliton density $\tilde{M}_s/2/(4\pi r_h^3/3)$ for ρ. Here, $\tilde{M}_s = M_s + 2\gamma M_{\text{bh}}$ is an effective core mass determining the frequency of the quasi-normal mode from the ULDM halo mass combined with the BHB mass weighted by a factor γ and r_h is the half-mass radius of the ULDM halo given by $f_0 \hbar^2/(m^2 G \tilde{M}_s)$ with $f_0 = 3.925$ (see for instance [23]). Then the above equation can be rewritten as

$$\tau = \frac{4\sqrt{2} f_0^3 \hbar^8}{G^{9/2}} \frac{1}{\alpha_{\text{DF}}^2 m^8 \tilde{M}_s^4 \sqrt{M_{\text{bh}} \bar{D}^5}}$$
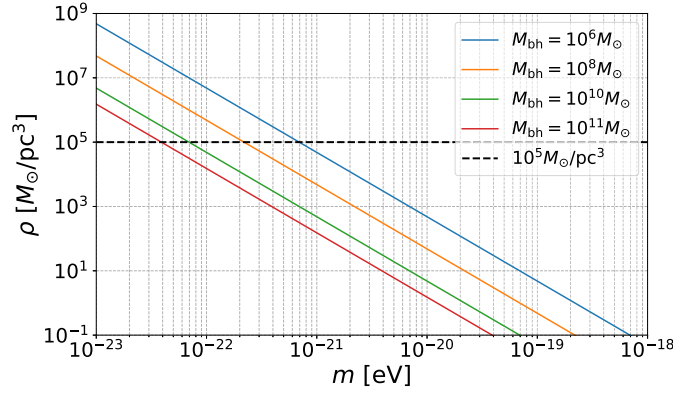
Fig. 5. The thick oblique lines represent the theoretical boundary where $\tau = 1$ Gyr for $M_{\mathrm{bh}} = (10^6, 10^8, 10^{10}, 10^{11})\mathrm{M}_\odot$, respectively, from the top to the bottom. (See equation (13).) The dashed horizontal line indicates the observational bound for the central dark matter density $\rho = 10^5 \mathrm{M}_\odot/\mathrm{pc}^3$ at $r \simeq 1$ pc. The separation is taken as $\bar{D} = 1$ pc and $\alpha_{\mathrm{DF}} = 0.424$. For each M_{bh}, only the parameter region below the horizontal line and over the corresponding thick line is allowed in our model.

$$= \left(\frac{2.72 \times 10^{-4}}{\alpha_{\mathrm{DF}}^2} \right) \left(\frac{m}{10^{-21}\ \mathrm{eV}} \right)^{-8} \left(\frac{M_{\mathrm{bh}}}{10^8 \mathrm{M}_\odot} \right)^{-\frac{1}{2}} \left(\frac{\tilde{M}_s}{10^9 \mathrm{M}_\odot} \right)^{-4} \left(\frac{\bar{D}}{\mathrm{pc}} \right)^{-\frac{5}{2}} \mathrm{Gyr}\,, \tag{14}$$

which is sensitive to m. The condition $\tau < 1$ Gyr now gives a bound $m > 9.1 \times 10^{-23}$ eV if we use the observational upper bounds $M_{\mathrm{bh}} \lesssim 10^{11} \mathrm{M}_\odot$ and $\tilde{M}_s \lesssim 10^{10} \mathrm{M}_\odot$ [42]. Remarkably, although our analysis about DF is not directly related to the small-scale problems, the bounds for m above are quite similar to the typical values that resolve the small-scale issues [9].

DF contribution may be compared to the one from the GC effect which is estimated below. As was mentioned already, the ULDM waves generated by the BHB carry away the momentum and orbital energy of the BHs. The corresponding GC timescale of the quasi-normal mode with SMBHB is given by $\tau_g = \hbar^3/(m^3 G^2 \tilde{M}_s^2)$ [40]. Since the perturbation has a local origin, this factor γ is in general slightly different from the global one (γ_s) in T_{slow} and will be determined separately below.

Each BH is carrying some extra amount of ULDM distribution that is gravitationally bound to the BH. We focus on the excess amount above the orbital average for a given orbital radius and call this mass δM bound to the BH. As we are considering a BHB with a separation $\bar{D}$, this density perturbation should be dipole-like in the small separation limit, and may be estimated as $\delta M = \alpha M_{\mathrm{bh}}(\pi \bar{D}/r_h)$ at the linear order, where $\alpha(M_{\mathrm{bh}}/M_s)$ is a dimensionless numerical coefficient.

We suggest that the magnitude of the momentum change of the ULDM halo should be balanced with that of δM, and then with the corresponding BH since δM is bound to the BH. We expect that mainly ULDM waves carry the increased momentum of the halo. For each revolution, if there were no dissipation, the BH would return to the original starting position and then the effective dissipative force, due to the GC effect, will be weighted by the cooling faction $\Delta\tau_c/\tau_g$ where $\Delta\tau_c \equiv \pi \bar{D}/v$ is approximately the orbital period of the BHs. This fraction can be also interpreted in the following way: We are looking to calculate the effective average force over the orbital period of the BHBs. Given that the time scale for gravitational cooling is about τ_g, we anticipate that a fraction, represented as $\Delta\tau_c/\tau_g$, of the soliton mass $\tilde{M}_s$ could be removed as ULDM waves due to the cooling throughout this period. Then, the effective force between the ULDM waves outgoing and δM is roughly $\bar{F}_{\mathrm{GC}} \propto G(\tilde{M}_s \Delta\tau_c/\tau_g)\delta M/r_h^2$. Note that we are estimating an average force for a specific moment in time, not solely for a single period. To determine the total momentum exchange, one must integrate the average force over the full duration of time. Thus, the effective force due to the GC effect is roughly

$$F_{\mathrm{GC}} = K_0 \left(\frac{\Delta\tau_c}{\tau_g} \right) \left(\frac{G\tilde{M}_s \delta M}{r_h^2} \right), \tag{15}$$

where K_0 is a numerical constant, and $\Delta\tau_c/\tau_g$ represents the above cooling fraction of the averaged force $\bar{F}_{\mathrm{GC}}$.

The ratio of F_{GC} over F_{DF} at given time may be straightforwardly evaluated as

$$\frac{F_{\mathrm{GC}}}{F_{\mathrm{DF}}} = \left(\frac{8\pi^2 \alpha K_0}{\alpha_{\mathrm{DF}}^2} \right) \sqrt{\frac{2 f_0 \bar{D}\, \tilde{M}_s^3}{r_h M_{\mathrm{bh}}^3}}, \tag{16}$$

where we again approximate the density of ULDM halo as $\rho = (\tilde{M}_s/2)/(4\pi r_h^3/3)$. Assuming the orbital decay in our simulation is solely from the DF effect, we find below that $\alpha_{\mathrm{DF}} \sim 0.424$ that is indeed of order one. This implies that, for our simulation, the GC contribution can be initially at most comparable to that of the DF. Also, in the small separation limit, the DF becomes dominant over the GC effect. Thus, from now on, we shall ignore the GC effect in our theoretical calculation and assume the orbital decay is mainly due to the DF. However, GC plays an important role during the orbital decay, as the spherical quasi-normal modes with the time scale T_{slow} can rapidly refill the phase space of ULDM that was removed by the SMBHB. (T_{slow} is much shorter than the decay time scale τ.) On the contrary, collisionless CDM experiences loss cone depletion due to its long relaxation time as mentioned earlier. It is one reason why ULDM better explains the rapid orbital decay, despite having a smaller C compared to CDM. Another problem of CDM is that the heat from DF of CDM can destroy CDM spikes [28].

Now, we compare our approximate calculation of τ with the numerical results. From equation (11), the infinitesimal orbital velocity change due to the DF is given by $dv = dt\,(F_{\mathrm{DF}}/M_{\mathrm{bh}})$. Assuming a circular orbit, the orbital velocity of BHs depends on both the BH mass (M_{bh}) and the ULDM mass (δM_s) enclosed within the orbit, with Kepler's law: $v = \sqrt{G\tilde{M}_{\mathrm{bh}}/2\bar{D}}$ where $\tilde{M}_{\mathrm{bh}} \equiv M_{\mathrm{bh}} + 4\delta M_s$. The enclosed mass within $\bar{D}/2$ is $\delta M_s = 4\pi \rho (\bar{D}/2)^3/3 = \eta\, m_{21}^6\, \tilde{M}_s^4 \bar{D}^3 = (B_0 \bar{D}^3)M_{\mathrm{bh}}/4$, where η is a constant and B_0 is introduced in the previous section. From now on, we assume

$B_0 \bar{D}^3 \ll 1$ so that its contribution in $\tilde{M}_{\rm bh}$ is ignored in the computation below. We also checked that our simulations in Sec. 3 indeed satisfy this condition. Then, $v \propto \bar{D}^{-\frac{1}{2}}$ and $\left| {\rm d}\bar{D}/\bar{D} \right| = 2 \left| {\rm d}v/v \right|$.

Finally, one obtains an approximate differential equation for $\bar{D}(t)$,

$$-\frac{{\rm d}\log\bar{D}}{{\rm d}t} = \frac{\alpha_{\rm DF}^2 m^8 G^{\frac{9}{2}} \tilde{M}_s^4 M_{\rm bh}^{\frac{1}{2}}}{2\sqrt{2} f_0^{\,3} \hbar^8} \bar{D}^{\frac{5}{2}} = Q_0 \bar{D}^{\frac{5}{2}}, \tag{17}$$

where $Q_0 \equiv \kappa\, m_{21}^8 \tilde{M}_s^4 M_{\rm bh}^{\frac{1}{2}}$ with a parameter $\kappa \equiv \alpha_{\rm DF}^2 m^8 G^{\frac{9}{2}}/(2\sqrt{2} f_0^{\,3} \hbar^8 m_{21}^8)$. This equation is perfectly consistent with the two scaling relations of our Schrödinger-Poisson system and may be integrated leading to the solution in equation (7).

Based on the above analysis, we fit Q_0 in Fig. 3 by assuming the quadratic form $\kappa = \kappa_0 \left[1 + 10\,\delta_1 (M_{\rm bh}/M_s) + 100\,\delta_2 (M_{\rm bh}/M_s)^2 \right]$ with fitting parameters $(\gamma, \kappa_0, \delta_1, \delta_2)$. This fit of Q_0 is depicted on the left panel of Fig. 3 by a dotted curve. One finds $\gamma = 2.192 \pm 0.034$, $\kappa_0 = (1.324 \pm 0.029) \times 10^{-43}\,{\rm kyr}^{-1}{\rm pc}^{-\frac{5}{2}}{\rm M}_\odot^{-\frac{9}{2}}$ with negligibly small δ_1 and δ_2. Then, $\kappa \simeq \kappa_0$ and it corresponds to $\alpha_{\rm DF} \simeq 0.424$.

We also obtain

$$Q_0 = 1.324 \left(\frac{\tilde{M}_s}{10^9 {\rm M}_\odot} \right)^4 \left(\frac{M_{\rm bh}}{10^8 {\rm M}_\odot} \right)^{1/2} \left(\frac{m}{10^{-21}{\rm eV}} \right)^8 {\rm Myr}^{-1}\,{\rm pc}^{-\frac{5}{2}}, \tag{18}$$

where the number 1.324 is from the fitting. The initial decay rate, given by $R(t_0) = \left| {\rm d}\log\bar{D}/{\rm d}t \right| \simeq Q_0 \bar{D}_0^{\frac{5}{2}}$ in equation (8), can be related to the decay scale τ in equation (14) as $\tau \simeq 2/R(t_0) = 2/(Q_0 \bar{D}_0^{\frac{5}{2}})$. For example, Fig. 3 shows that $2/R(t_0) \simeq 0.617$ Myr for $M_{\rm bh} = 10^8\,{\rm M}_\odot$ from our simulation, where $\tilde{M}_s = M_s + 2\gamma M_{\rm bh} \simeq 1.438 \times 10^9\,{\rm M}_\odot$. This time scale from simulation is comparable to the theoretical prediction $\tau = 0.47$ Myr in equation (14) with simple assumptions $\alpha_{\rm DF} = 1$, $\tilde{M}_s \simeq M_s = 10^9\,{\rm M}_\odot$ and $\bar{D}_0 = 0.8$ pc, thereby supporting our analysis. To check a full decay time, we need to do a simulation up to $\bar{D}(t) \sim 0.01$ pc, which is beyond the scope of this work.

5. Summary

In this letter, we demonstrate through numerical simulations that ULDM waves generated by rotating BHB in ULDM halos can cause rapid orbital decay and give a hint to the final parsec problem. This phenomenon is rather unexpected because the orbital decay of BHB in ULDM halos is generally considered to be slow due to the weak dissipation effects [43]. The orbital decay arises mainly from DF and GC of ULDM halos interacting with the BHB where DF becomes dominant over GC effect in the small separation limit of the BHB. One advantage of ULDM is that it is free from the loss cone problem, unlike collisionless CDM or stars. This is because ULDM has a wavelike nature, which helps it rapidly refill the phase space due to GC. Fig. 1 shows that the central density of ULDM may even increase as the separation of the black holes decreases unlike CDM. Investigating the influence of the ULDM waves on nearby stars might offer another potential solution to the loss-cone problem [44].

These findings have implications for future gravitational wave probes like LISA or NANOGrav, as ULDM waves can change the orbital properties of the SMBHBs and influence gravitational waves generated by them, especially when their separation is small. To further validate our results, higher-resolution simulations are planned for future studies.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Hyeonmo Koo reports financial support was provided by National Research Foundation of Korea. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

Acknowledgements

References

[1] M.C. Begelman, R.D. Blandford, M.J. Rees, Nature 287 (1980) 307.
[2] Miloš Milosavljević, David Merritt, AIP Conf. Proc. 686 (1) (2003).
[3] G. Agazie, et al., NANOGrav, Astrophys. J. Lett. 952 (2) (2023) L37.
[4] M.R. Baldeschi, G.B. Gelmini, R. Ruffini, Phys. Lett. B 122 (1983) 221.
[5] M. Membrado, A.F. Pacheco, J. Sañudo, Phys. Rev. A 39 (1989) 4207.
[6] S.-J. Sin, Phys. Rev. D 50 (1994) 3650.
[7] J.-W. Lee, I.-G. Koh, Phys. Rev. D 53 (1996) 2236.
[8] F.S. Guzman, T. Matos, Class. Quantum Gravity 17 (2000) L9.

[9] W. Hu, R. Barkana, A. Gruzinov, Phys. Rev. Lett. 85 (2000) 1158.
[10] S. Park, D. Bak, J.W. Lee, I. Park, J. Cosmol. Astropart. Phys. 12 (2022) 033.
[11] W.J.G. de Blok, A. Bosma, S.S. McGaugh, arXiv:astro-ph/0212102.
[12] A. Tasitsiomi, Int. J. Mod. Phys. D 12 (2003) 1157.
[13] T. Matos, D. Nunez, Rev. Mex. Astron. Astrofís. 44 (2008) 149.
[14] J.P. Mbelek, Astron. Astrophys. 424 (2004) 761–764.
[15] T.H. Lee, B.J. Lee, Phys. Rev. D 69 (2004) 127502.
[16] H.-Y. Schive, T. Chiueh, T. Broadhurst, Nat. Phys. 10 (2014) 496–499.
[17] T.H. Lee, B.J. Lee, Phys. Rev. Lett. 123 (2019) 021102.
[18] D.J.E. Marsh, Phys. Rep. 643 (2016) 1.
[19] J.-W. Lee, J. Korean Phys. Soc. 54 (2009) 2622.
[20] A. Suárez, V.H. Robles, T. Matos, Astrophys. Space Sci. Proc. 38 (2014) 107.
[21] T. Rindler-Daller, P.R. Shapiro, Mod. Phys. Lett. A 29 (2014) 1430002.
[22] P.-H. Chavanis, Phys. Rev. D 84 (2011) 043531.
[23] L. Hui, J.P. Ostriker, S. Tremaine, E. Witten, Phys. Rev. D 95 (2017) 043531.
[24] L. Lancaster, C. Giovanetti, P. Mocz, Y. Kahn, M. Lisanti, D.N. Spergel, J. Cosmol. Astropart. Phys. 01 (2020) 001.
[25] E. Seidel, W.M. Suen, Phys. Rev. D 42 (2) (1990) 384.
[26] A. Boudon, P. Brax, P. Valageas, Phys. Rev. D 106 (4) (2022) 043507.
[27] R. Boey, Y. Wang, E. Kendall, R. Easther, Phys. Rev. D 109 (10) (2024) 103526.
[28] J.I. Read, G. Gilmore, Mon. Not. R. Astron. Soc. 339 (2003) 949.
[29] J.C. Aurrekoetxea, K. Clough, J. Bamber, P.G. Ferreira, Phys. Rev. Lett. 132 (21) (2024) 211401.
[30] Y. Wang, R. Easther, arXiv:2110.03428.
[31] F. Edwards, E. Kendall, S. Hotchkiss, R. Easther, J. Cosmol. Astropart. Phys. 10 (2018) 027.
[32] D. Bak, J.-W. Lee, S. Park, J. Korean Phys. Soc. 79 (6) (2021) 582.
[33] T. Zimmermann, J. Alvey, D.J.E. Marsh, M. Fairbairn, J.I. Read, arXiv:2405.20374.
[34] E. Armengaud, N. Palanque-Delabrouille, C. Yèche, D.J.E. Marsh, J. Baur, Mon. Not. R. Astron. Soc. 471 (4) (2017) 4606–4614.
[35] G. Heißel, T. Paumard, G. Perrin, F. Vincent, Astron. Astrophys. 660 (2022) A13.
[36] P. Gondolo, J. Silk, Phys. Rev. Lett. 83 (1999) 1719–1722.
[37] A.A. Avilez, L.E. Padilla, T. Bernal-Marin, T. Matos, Mon. Not. R. Astron. Soc. 477 (3) (2018) 3257–3272.
[38] N. Bar, K. Blum, T. Lacroix, P. Panci, J. Cosmol. Astropart. Phys. 07 (2019) 045.
[39] T. Lacroix, Astron. Astrophys. 619 (2018) A46.
[40] F. Siddhartha Guzmán, L. Arturo Ureña-López, Phys. Rev. D 69 (2004) 124033.
[41] R. Buehler, V. Desjacques, Phys. Rev. D 107 (2) (2023) 023516.
[42] A. King, Mon. Not. R. Astron. Soc. 456 (2016) L109.
[43] R. Vicente, V. Cardoso, Phys. Rev. D 105 (2022) 083008.
[44] B.C. Bromley, P. Sandick, B. Shams Es Haghi, arXiv:2311.18013.