Multimessenger time delays from lensed gravitational waves

Tessa Baker[1,2,*] and Mark Trodden[2,†]

[1]*Denys Wilkinson Building, University of Oxford, Keble Road, Oxford OX1 3RH, United Kingdom*
[2]*Center for Particle Cosmology, Department of Physics and Astronomy, University of Pennsylvania,
Philadelphia, Pennsylvania 19104, USA*

We investigate the potential of high-energy astrophysical events, from which both massless and massive signals are detected, to probe fundamental physics. In particular, we consider how strong gravitational lensing can induce time delays in multimessenger signals from the same source. Obvious messenger examples are massless photons and gravitational waves, and massive neutrinos, although more exotic applications can also be imagined, such as to massive gravitons or axions. The different propagation times of the massive and massless particles can, in principle, place bounds on the total neutrino mass and probe cosmological parameters. Whilst measuring such an effect may pose a significant experimental challenge, we believe that the "massive time delay" represents an unexplored fundamental physics phenomenon.

I. INTRODUCTION

Photons are no longer our only window onto the universe. The recent detections of GW150914 and GW151226 [1,2] announced the arrival of the powerful new tool of gravitational wave astronomy. In addition, particle messengers—astrophysical neutrinos and cosmic rays—have been detected at Earth for some decades. However, modeling the production of gravitational waves (hereafter GWs) and particles during extreme astrophysical events is a challenging scientific problem, requiring advanced numerical work.

In contrast, the subsequent propagation of such signals across the universe is comparatively straightforward to describe. This prompts us to ask if the *propagation* of different multimessenger observables can be used as a new probe for cosmology and/or fundamental physics, independent of the complex details of particle or waveform generation.

In particular, small relativistic corrections that accumulate with propagation distance may become measurable for sources at high redshifts, revealing information about the difference between null and non-null geodesics of the intervening spacetime. This has the power to tell us about both the expansion rate of the universe, and also properties of the massive particles being used to trace the geodesics [3,4].

The simplest example that springs to mind is to compare the arrival times of photons and neutrinos from supernovae, first put forward by Zatsepin in 1968 [5]. However, in such a system astrophysical complexities are likely to dominate effects of interest to fundamental physics. For example, neutrinos from the famous supernovae 1987A arrived four hours *earlier* than the appearance of the optical counterpart, because of the prolonged escape time of photons from a dense supernova remnant [6].

Gravitational waves, on the other hand, suffer no such setbacks. Their minimal interaction with matter—and hence negligible scattering and absorption—makes them arguably a cleaner probe, if the source itself is not the chief object of interest. For example, one could ask the following, simplistic question: given that we know neutrinos have mass, whilst GWs are massless (in GR, at least), how much later would the neutrinos arrive at Earth—assuming the two were emitted simultaneously?

A simple calculation (presented in Appendix A) of propagation times shows that for a source in the redshift range $z = 0.5$–5, the difference in arrival times between a GW and a typical neutrino would be of order one second (see also [7]). Unfortunately, in a realistic scenario, there will be an additional contribution imprinted on this delay by the structure of the astrophysical source, i.e. the fact that emission of particles and GWs may not commence exactly simultaneously. This *intrinsic source delay* could be of order seconds or longer [8]. Without detailed knowledge and modeling of the source, it would be impossible to know how to split the measured difference in arrival time into its intrinsic and particle-mass contributions.

As we will show in this paper, the difficulty above can be resolved if the multimessenger signals encounter a gravitational lensing event(s) en route to Earth. Gravitational waves are subject to gravitational lensing in almost exactly the same manner as photons [9]. A key part of the derivations presented here is to develop a description of the lensing of massive particles, a topic that seems to be curiously absent from current literature. We will find that lensing imparts an additional delay in arrival times that is sensitive to the mass-squared of the messenger particles.

This *massive time delay* depends on a number of quantities of interest to both cosmology and particle

*tessa.baker@physics.ox.ac.uk
†trodden@physics.upenn.edu

physics, namely the mass of the particle involved, the redshift of the source, and the expansion history of the universe. Given how small neutrino masses (for example) are expected to be, it is clear that the massive time delay will remain small for them (< 1 second). The question we are interested in here is whether the massive time delay might nevertheless be, in some cases, large enough to be measured by a network of future gravitational wave and particle detectors.

The lensing of gravitational waves is developed in [9,10], and the detection rate of such events has been discussed for the AdLIGO [11], LISA [12,13] and Einstein Telescope [14] detector networks. Clearly, a confirmed single source of all three potential messengers—photons, gravitational waves, and massive particles—could offer even further possibilities. The identification of electromagnetic counterparts to GW sources is a major focus of current gravitational wave science, and is discussed extensively elsewhere [15,16].

Although the sources themselves are not the main focus of this paper, let us briefly enumerate some objects from which one might expect coincident emission of massive and massless particles.

(1) The most promising source, and the one we will generally focus on, is a merger between two neutron stars (hereafter NS). This can result in the formation of a hypermassive neutron star (HMNS) that is stable for timescales of order 10 ms. The HMNS continues to accrete material from surrounding debris; shocks associated to this high-energy accretion environment result in an outflow of neutrinos from the poles of the HMNS. This neutrino outflow is generally diffuse rather than highly beamed, with typical particle energies of order ~ 10 MeV [17,18].

(2) Other compact object mergers may produce particle emission in addition to their GW signals, although there is a higher degree of uncertainty here. A merger between a NS and a black hole (BH) has the necessary matter component, although without the formation of a long-lived HMNS the neutrino luminosity may be much lower [19]. A BH-BH merger could produce particle emission if there is an accretion disc close enough to be strongly affected by the merger [20].

(3) We have already mentioned supernovae above; if these are significantly asymmetric, they can produce GWs in addition to particle emissions [21–23]. For recent discussions of lensed extragalactic supernovae, see [24–26].

In this work we will treat a simplified scenario, considering the lensing of massless and massive relativistic particles by a single, isolated source—the strong lensing regime. In reality, our multimessengers are likely to experience many additional small deflections along their path, analogous to weak lensing in electromagnetic astronomy. We acknowledge from the start the existence of such complicating factors

in any realistic scenario; this work is intended to be a first step in fleshing out the key features of a hitherto unexplored phenomenon. We will discuss our omissions in Sec. IV, and leave a detailed comparison to projected experimental sensitivities for future investigation.

The structure of this paper is as follows: in Sec. II we derive the correction to the "flight time" of a massive particle, relative to a massless one, that encounters a strong gravitational lens. In Sec. III we explain how a strategy of differencing the massive and massless arrival times can ameliorate the unknown intrinsic delay between their emission. We then evaluate this differential massive time delay for some simple lens models: the singular isothermal sphere and the power-law lens. Section IV discusses some additional features of the phenomenon, which would likely complicate a measurement of the effects described here. We conclude in Sec. V. Several calculations tangential to our principal discussion are presented in the appendices.

II. THE MASSIVE TIME DELAY

A. Structure of the calculation

In Fig. 1 we show the basic features of the system under consideration. An energetic event at redshift z_S and conformal distance D_S releases both massless emissions (electromagnetic and/or gravitational radiation) and relativistic, massive particles of mass m. For argument's sake we will sometimes refer to these massive particles as neutrinos, though our formalism applies more generally. Note that the massless and massive fluxes will not typically commence exactly simultaneously—we discuss how to deal with this in Sec. III.

At redshift z_L and conformal distance D_L, both types of emission encounter a gravitational lens with three-dimensional density profile $\rho(\vec{r})$. Throughout this paper we will make use of the thin-lens approximation, which treats all the deflection as occurring instantaneously at a single plane. Quantities relating to the lens, such as its density and gravitational potential, will be projected onto a two-dimensional plane at redshift z_L.

The emissions travel on multiple paths i around this lens, with each path experiencing a total angular deflection α_i and finally being received by an observer O at angle θ_i to the optical axis OA (defined as the axis connecting the observer to the center of the lens). In optical strong lensing of extended sources, these multiple paths and detection angles correspond to multiple images of the same source, often distorted in an informative way that constrains the structure of lens.

The emissions we are interested in here originate from point sources. Hence, though we expect to detect multiple, identical point-sources of our massive and massless messengers, there will be no equivalent of the spatial image distortion seen in traditional optical lensing. However, in some cases it might be possible to associate the lensed point
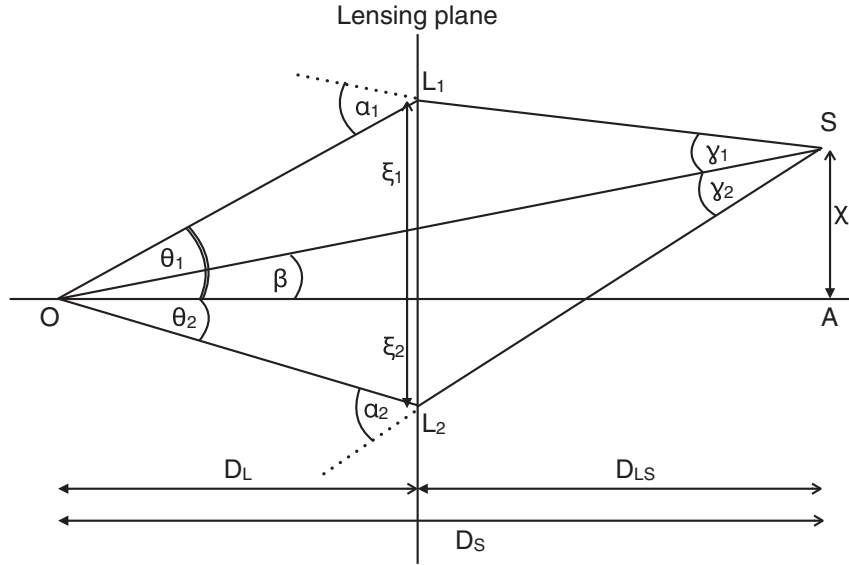
Lensing plane

FIG. 1. Diagram illustrating angles and conformal distances relevant for the derivation of Sec. II. Two lensed paths are shown (intersecting the lensing plane at L_1 and L_2), and also the undeflected path that the rays would all follow if the lens were absent, OS. ξ_1 and ξ_2 are two-dimensional position vectors of the images in the lensing plane, and χ is the position vector of the source in the plane AS. α_1 and α_2 are the deflection angles experienced by these two rays at the lens.

sources with lensed optical images of a host galaxy, although this will require significant advances in source localization of GW and neutrino detectors. That said, even if localization techniques do not improve sufficiently to allow the spatial resolution of lensed GW sources, emissions that have traveled along different lensed paths may still be *temporally* resolvable.

Within the simplified model outlined above, the total conformal travel time for a massive, relativistic particle has the structure:

$$\eta_{\text{total}}(\theta, m) = \eta_{\text{undeflected}}(m) + \eta_{\text{massless}}(\theta) + \eta_{\text{massive}}(\theta, m). \qquad (1)$$

The first term in this expression is the travel time from the source to the observer for a particle in the absence of any lens. It is the same for all paths (hence no dependence on θ), and provides the largest contribution to η_{total}; note however, that it will depend on the mass of the particle (Sec. II B). The remaining terms describe corrections to this minimum travel time induced by the presence of the lens; in the case of a massless particle only the first correction term in Eq. (1) exists.

In principle, a massive and massless particle will be deflected by slightly different amounts at the lens—the derivation of this effect is given in Appendix B. We would therefore expect there to be a small offset in the position of the neutrino and GW sources on-sky, as shown in Fig. 2. To account for this we compare flight times along two different but very close paths, by writing the reception angle of the massive particle as $\theta = \bar{\theta} + \delta\theta$, where $\bar{\theta}$ is the reception angle in the massless case [27].

Performing a Taylor expansion of (1), we then have:

$$\eta_{\text{total}}(\theta, m) \cong \eta_{\text{undeflected}}(m) + \eta_{\text{massless}}(\bar{\theta})$$
$$+ \left. \frac{\partial \eta_{\text{massless}}(\theta)}{\partial \theta} \right|_{\theta = \bar{\theta}} \delta\theta$$
$$+ \eta_{\text{massive}}(\bar{\theta}, m) + \mathcal{O}\left(\delta\theta^2, \frac{m^2}{p_0^2} \delta\theta\right), \qquad (2)$$

where the last term—using notation to be introduced shortly in Sec. II B—indicates that significantly suppressed terms have been neglected [28]. For this reason, it is sufficient to evaluate η_{massive} at $\bar{\theta}$.

The principle of least action (which for a massless particle, in the geometric optics limit, becomes Fermat's principle) tells us that a particle follows a path along which the travel time is stationary; thus the third term in Eq. (2) must vanish. Hence, to first order in $\delta\theta$, it is sufficient to evaluate the flight time of the both massive and massless particles at the same reception angle $\bar{\theta}$. Any difference in arrival times then arises purely from their null/non-null nature, rather than path differences. It is interesting to note, though, that this difference in deflection angle does exist in principle, and would need to be accounted for in high-accuracy calculations.

Now consider the conformal time delay between two emissions (either massive or massless) that travel about the lens on *widely*-separated paths, like those shown in Fig. 1. These are received at angles θ_1 and θ_2, corresponding to separate point source images on the sky. The conformal time interval between their arrivals is
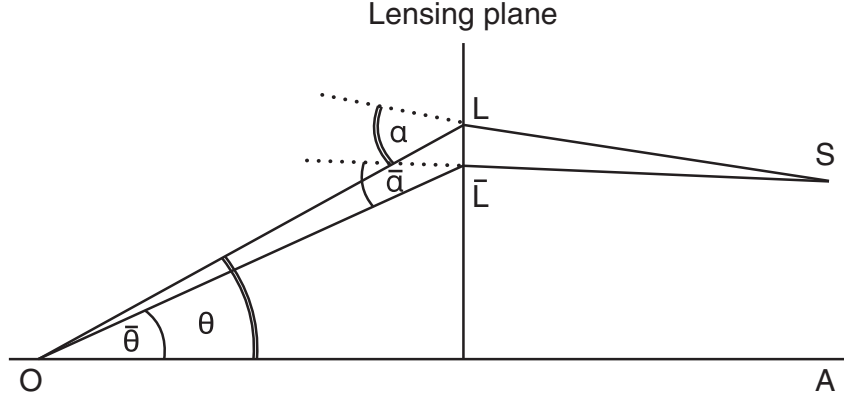
FIG. 2. As massive and massless particles are deflected by different amounts at the lens ($\bar{\alpha}$ and α respectively), one should evaluate their travel times along slightly different paths SLO and $S\bar{L}O$. However, this turns out to produce corrections that are second-order with respect to the differential massive time delay that is the main focus of this paper. Hence, in what follows, we can consider both massless and massive particles corresponding to the same image to have propagated along identical trajectories.

$$\Delta\eta(\theta_2, \theta_1) = \eta_{\text{total}}(\theta_2, m) - \eta_{\text{total}}(\theta_1, m). \quad (3)$$

The term $\eta_{\text{undeflected}}$ in Eq. (1) makes no contribution to $\Delta\eta(\theta_2, \theta_1)$, as it is the same for both paths and therefore cancels out. Hence, although η_{total} may be a significant fraction of the age of the universe, the *difference* between the (conformal) travel time of two rays received at $z = 0$ is much smaller than the Hubble time, typically between a few days and a few years. Since cosmological expansion is negligible over these timescales, the conformal time delay is equivalent to the physical time delay to an extremely good approximation [29], i.e.

$$\Delta t(\theta_2, \theta_1) = a(z = 0)\Delta\eta(\theta_2, \theta_1) = \Delta\eta(\theta_2, \theta_1). \quad (4)$$

Therefore, following [30], we can safely formulate our calculation in terms of conformal distances and conformal times.

B. Derivation

Our goal here is to calculate the travel time of a particle along the paths shown in Fig. 1; in the thin-lens approximation each of these consists of two straight-line segments, with a total deflection angle α_i incurred instantaneously at the lens. We use the letter i to label an unspecified number of different deflected paths; only two are shown in Fig. 1 for clarity.

We begin from the line element for a particle moving in a spacetime containing a single linear perturbation Φ in an otherwise homogeneous universe [31]:

$$ds^2 = -\epsilon c^2 d\lambda^2 = -c^2 a(\eta)^2 d\eta^2 \left(1 + \frac{2\Phi}{c^2}\right) + g_{ij} dx^i dx^j.$$

For a massive particle $d\lambda$ would be an element of proper time; we have chosen slightly unusual notation here to allow us to unify the massive and massless cases

(see below). Note that we have introduced a binary parameter in the first equality: $\epsilon = 1$ for a massive particle and $\epsilon = 0$ for a massless one. We have also chosen to keep the spatial part of the metric general for now. A little rearrangement of the second equality brings this line element to the form:

$$d\eta = \left(1 + \frac{2\Phi}{c^2}\right)^{-\frac{1}{2}} \frac{d\lambda}{a} \left[\epsilon + \frac{g_{ij}}{c^2} \frac{dx^i}{d\lambda} \frac{dx^j}{d\lambda}\right]^{\frac{1}{2}}. \quad (5)$$

In the massless case, the quantity λ in the expression above is simply an affine parameter (not proper time). Next, we Taylor-expand the first bracket (since Φ/c^2 is a small quantity) and introduce the spatial three-momentum magnitude:

$$p^2 = m^2 g_{ij} \frac{dx^i}{d\lambda} \frac{dx^j}{d\lambda}. \quad (6)$$

Using Eq. (6) in (5) then leads us to

$$d\eta \approx \left(1 - \frac{\Phi}{c^2}\right) \left[\epsilon + \frac{p^2}{m^2 c^2}\right]^{\frac{1}{2}} \frac{d\lambda}{a}. \quad (7)$$

To find the (conformal) travel time of the particle from source to observer, we need to integrate $d\eta$ along one of the paths SL_iO shown in Fig. 1. To do this, we first eliminate the affine parameter element $d\lambda$ in favor of a conformal distance element along the lensed path, $d\ell = \sqrt{\delta_{ij} dx^i dx^j}$ [32]. Expanding $g_{ij} = \delta_{ij} a^2 (1 - 2\Phi/c^2)$ and performing a Taylor expansion in Φ/c^2, Eq. (6) rearranges to become

$$d\lambda \approx \frac{ma}{p} \left(1 - \frac{\Phi}{c^2}\right) d\ell.$$

Substituting this into Eq. (7), we then obtain

$$dη \approx \left(1 - 2\frac{\Phi}{c^2}\right)\left[1 + \epsilon\frac{m^2 c^2}{p^2}\right]^{\frac{1}{2}}\frac{d\ell}{c}. \tag{8}$$

We can now write $d\ell = dy/\cos\gamma(y)$, where y measures distance along the *undeflected ray OS*, and $\gamma_i(x)$ is the angle the path element $d\ell$ makes with OS, giving

$$d\ell = \frac{dy}{\cos\gamma(y)} \approx \left(1 + \frac{\gamma(y)^2}{2}\right)dy, \tag{9}$$

where the large distances involved ensure that the small angle approximation used in the second equality is valid. Taking Eqs. (8) and (9) together, and neglecting terms that are second-order in the small quantities Φ and γ, we finally reach

$$dη \approx \left(1 + \frac{\gamma(y)^2}{2} - \frac{2\Phi}{c^2}\right)\left[1 + \frac{m^2 c^2}{p^2}\right]^{\frac{1}{2}}\frac{dy}{c}. \tag{10}$$

Note that now that the mass of the particle is explicitly present, the parameter ϵ is surplus to requirement and has been absorbed into m^2 in the line above.

Integrating Eq. (10) along the ray OS will yield the conformal time taken for the particle to travel the lensed path. In the thin-lens approximation, the integration path breaks into two stages, with the value of $\gamma(x)$ a constant along each (see Table I). However, when integrating Eq. (10) we must also remember that the spatial three-momentum p redshifts in proportion to $1/a$; this is true for both massive and massless particles [33]. Rearranging Eq. (10), multiplying out the first bracket and integrating, the conformal time taken to travel one of the lensed paths is given by

$$\eta(\theta, m) = \int_0^{D_S}\left[1 + \frac{m^2 c^2 a^2}{p_0^2}\right]^{\frac{1}{2}}\frac{dy}{c} + \int_0^{D_S}\left(\frac{\gamma^2}{2}\right)$$
$$\times\left[1 + \frac{m^2 c^2 a^2}{p_0^2}\right]^{\frac{1}{2}}\frac{dy}{c} + \int_{D_L-\delta}^{D_L+\delta}\left(-\frac{2\Phi}{c^2}\right)$$
$$\times\left[1 + \frac{m^2 c^2 a^2}{p_0^2}\right]^{\frac{1}{2}}\frac{dy}{c} \tag{11}$$

$$\approx \int_0^{D_S}\left[1 + \frac{m^2 c^2 a^2}{p_0^2}\right]^{\frac{1}{2}}\frac{dy}{c} + \frac{\gamma_{OL}^2}{2}\int_0^{D_L}\frac{dy}{c}$$
$$+ \frac{\gamma_{LS}^2}{2}\int_{D_L}^{D_S}\frac{dy}{c} - \int_{D_L-\delta}^{D_L+\delta}\frac{2\Phi}{c^2}\frac{dy}{c} + \frac{1}{2}\left(\frac{mc}{p_0}\right)^2$$
$$\times\left\{\frac{\gamma_{OL}^2}{2}\int_0^{D_L}a^2\frac{dy}{c} + \frac{\gamma_{LS}^2}{2}\int_{D_L}^{D_S}a^2\frac{dy}{c}\right.$$
$$\left. - \int_{D_L-\delta}^{D_L+\delta}\frac{2\Phi}{c^2}a^2\frac{dy}{c}\right\}. \tag{12}$$

In the first line above we have written the three-momentum as $p = p_0/a$, where p_0 is the value at redshift zero. For most

TABLE I. Values of γ, the angle between a deflected ray and the observer-source axis OS, for the two paths shown in Fig. 1. The second column shows the value on the path segment between the source and the lens; the third column is the value on the segment between the lens and observer.

Path	γ_{LS}	γ_{OL}
1 (upper)	$\alpha_1 - \theta_1 + \beta$	$\theta_1 - \beta$
2 (lower)	$\alpha_2 - \theta_2 - \beta$	$\theta_2 + \beta$

of the results in Sec. III we will used a fixed value p_0, so we have not included it as an explicit argument of η here.

The third integral has a restricted integration range, since the integrand is only nonzero in a small region of size 2δ near the potential well Φ. In moving to the second line we have performed a Taylor expansion in the small quantity mc/p_0 ($mc \ll p_0$ since we are dealing with relativistic particles), and have broken the second integral of Eq. (11) into the two sections OL and LS indicated in Fig. 1. Since γ is a constant along these sections, it can be factored out of the integrals. Note that the integrands in the final line above pick up a factor of a^2 from the redshifting of the three-momentum.

Equation (12) is valid for any lensed path in the thin-lens approximation. However, for the rest of this paper we will specialize to the two-image case illustrated in Fig. 1. For now, let us evaluate the travel time along the upper path shown in Fig. 1, using the values of γ given in Table I. We identify the first term of Eq. (12) as the unlensed travel time of a particle, i.e. $\eta_{\text{undeflected}}$ in Eq. (1). Proceeding to evaluate the remaining integrals, we obtain

$$\eta(\theta_1, m) = \eta_{\text{undeflected}}(m) + \frac{(\theta_1 - \beta)^2}{2}\frac{D_L}{c}$$
$$+ \frac{(\alpha_1 - \theta_1 + \beta)^2}{2}\frac{(D_S - D_L)}{c} - \frac{D_L D_S}{c D_{LS}}\psi(\theta_1)$$
$$+ \frac{1}{2}\left(\frac{mc}{p_0}\right)^2\left\{\frac{(\theta_1 - \beta)^2}{2}\int_1^{a_L}\frac{da}{H(a)}\right.$$
$$\left. + \frac{(\alpha_1 - \theta_1 + \beta)^2}{2}\int_{a_L}^{a_S}\frac{da}{H(a)} - \frac{D_L D_S}{c D_{LS}}\psi(\theta_1)a_L^2\right\}, \tag{13}$$

where we have changed the integration variables for the integrals in the second line, and in the final term have taken the limit $\delta \ll D_S$ implied by the thin-lens approximation. We have also defined the two-dimensional projected potential as

$$\psi(\theta) = \frac{D_{LS}}{D_L D_S}\int_{D_L-\delta}^{D_L+\delta}\frac{2\Phi(\theta, y)}{c^2}dy. \tag{14}$$

Let us briefly focus on the lensing contribution to the travel time that exists for both massive and massless particles, i.e. the second, third and fourth terms of Eq. (13). We make use of the simply-named *lens equation* [34]

$$\vec{\alpha}_{\text{scal}} = \frac{D_{LS}}{D_S}\vec{\alpha} = \vec{\theta} - \vec{\beta}, \qquad (15)$$

where the first equality defines the *scaled* deflection angle. The second equality is a standard relation that can be derived by consideration of equivalent triangles in Fig. 1. Notice that it is important to consider the direction of deflection here. If we let $\hat{\phi}$ denote a unit vector in the clockwise direction with respect to OA, then (referring to Fig. 1) $\vec{\gamma}_1 = \gamma_1\hat{\phi}$, but $\vec{\gamma}_2 = -\gamma_2\hat{\phi}$. This results in the sign differences seen in Table I for the two paths.

Isolating the non-mass-dependent contribution to the time delay and substituting in Eq. (15) we obtain

$$\eta_{\text{massless}}(\theta_1, m) = \frac{(\theta_1 - \beta)^2}{2}\frac{D_L}{c} + \frac{1}{2}\left[\left(\frac{D_S}{D_{LS}} - 1\right)(\theta_1 - \beta)\right]^2$$
$$\times \frac{D_{LS}}{c} - \frac{D_L D_S}{c D_{LS}}\psi(\theta_1) \qquad (16)$$

$$= \frac{(\theta_1 - \beta)^2}{2}\frac{D_L}{c}\left[1 + \frac{D_L}{D_{LS}}\right] - \frac{D_L D_S}{c D_{LS}}\psi(\theta_1) \qquad (17)$$

$$= \frac{D_L D_S}{c D_{LS}}\left[\frac{1}{2}(\theta_1 - \beta)^2 - \psi(\theta_1)\right]. \qquad (18)$$

Equation (18), is equivalent to the standard expression for the time delay of lensed photons, often expressed in terms of the Fermat potential [35]. There are two effects that contribute to the lensed travel time: a *geometric delay* that arises purely from the increased path length [first term in Eq. (18)] and a *Shapiro delay* that occurs as particles pass through a gravitational potential well (second term). Note that the Shapiro delay is incurred at a single redshift, z_L. We note in passing that our derivation of this expression—beginning from a line element—differs substantially from the most widely used presentation, which involves deducing the geometric and Shapiro terms from Fermat's principle.

For a massive particle, the second line in Eq. (13) also comes into play. Using the lens equation once more, the massive correction to the travel time along the upper path of Fig. 1 can be written as

$$\eta_{\text{massive}}(\theta_1, m) = \frac{1}{2}\left(\frac{mc}{p_0}\right)^2\left\{\frac{1}{2}(\theta_1 - \beta)^2\left[\int_1^{a_L}\frac{da}{H(a)}\right.\right.$$
$$\left.\left. + \frac{D_L^2}{D_{LS}^2}\int_{a_L}^{a_S}\frac{da}{H(a)}\right] - a_L^2\frac{D_L D_S}{c D_{LS}}\psi(\theta_1)\right\}. \qquad (19)$$

This expression merits a few comments. First, note that the correction to the travel time of a massive, relativistic particle has an overall prefactor of $(mc/p_0)^2$, as might be intuited from special relativistic considerations. For all the scenarios discussed in this paper, the initial energy, E_0, of the massive particle is substantially greater than its rest mass. Hence in what follows we will implicitly take $(mc/p_0)^2 = m^2c^4/(E_0^2 - m^2c^4) \approx (mc^2/E_0)^2$. In Sec. III, where we evaluate this correction numerically, we will see that this ratio is the single most important factor controlling the magnitude of the effects derived here.

Second, distinct geometric (first two terms) and Shapiro (last term) contributions are still identifiable in Eq. (19), even though the final form is not as elegantly compact as Eq. (18). We note that the Shapiro-like correction for a massive particle in a Schwarzchild metric is derived in [36].

Third, η_{massive} depends on the cosmological expansion history in a more complicated manner than its massless counterpart; note that the expansion history of the universe only enters Eq. (18) via the overall prefactor $D_L D_S/D_{LS}$. This difference occurs because the redshifting three-momentum of massive particles affects their propagation speed, and hence their travel time. Whilst clearly massless particles experience energy-momentum redshifting as well, it does not alter their propagation speed and hence does not affect the massless time delay in such an intricate way.

Finally, given the complicated dependence of Eq. (19) on D_L, D_S and D_{LS}, the redshift-dependence of η_{massive} is not easy to predict. In particular, it may not necessarily peak when the lens is halfway between the observer and source, as typically occurs for lensing kernels. We will study this further in Sec. III.

We have found expressions for the three contributions to the travel time identified in Eq. (1): $\eta_{\text{undeflected}}(m)$, $\eta_{\text{massless}}(\theta)$ and $\eta_{\text{massive}}(\theta, m)$. We now have all the tools to calculate the relative time delay between the arrival of two massive particles that have traveled on different paths around a lens, or between a massive and massless particle traveling the same path.

III. APPLICATION & LENS MODELS

A. Differencing strategy

For convenience we summarize here the results of the previous section, now generalized to apply to both paths in Fig. 1:

$$\eta_{\text{total}}(\theta) = \eta_{\text{undeflected}}(m) + \eta_{\text{massless}}(\theta)$$
$$+ \eta_{\text{massive}}(\theta, m) + \mathcal{O}\left(\left[\frac{mc}{p_0}\right]^4\right) \qquad (20)$$

$$\eta_{\text{undeflected}}(m) = \int_0^{D_S}\left[1 + \frac{m^2c^2a^2}{p_0^2}\right]^{\frac{1}{2}}\frac{dy}{c} \qquad (21)$$

$$\eta_{\text{massless}}(\theta) = \frac{D_L D_S}{c D_{LS}}\left[\frac{1}{2}(\theta \pm \beta)^2 - \psi(\theta)\right] \qquad (22)$$

$$\eta_{\text{massive}}(\theta, m)) = \frac{1}{2}\left(\frac{mc}{p_0}\right)^2 \left\{ \frac{1}{2}(\theta \pm \beta)^2 \right.$$
$$\times \left[\int_1^{a_L} \frac{da}{H(a)} + \frac{D_L^2}{D_{LS}^2} \int_{a_L}^{a_S} \frac{da}{H(a)} \right]$$
$$\left. - a_L^2 \frac{D_L D_S}{c D_{LS}} \psi(\theta) \right\}. \qquad (23)$$

The $\pm$ signs correspond to the paths below (+, $i = 2$) and above (−, $i = 1$) the optical axis shown in Fig. 1. The simple lens models considered in this paper produce only two images, such that we can always orient the system as shown in the figure. We leave the treatment of the more realistic system—which, famously, can only produce odd numbers of images [37]—to a future work.

Let us explain here the most sensible way to combine these particle flight times. As we described in the introduction, a major systematic error when trying to measure the delay in arrival between massive and massless particles would be the unknown relative emission time. So far we have implicitly assumed that all our messenger particles set off exactly simultaneously, but this is unlikely for any realistic source. For example, in a supernova the neutrino diffusion timescale in a collapsing stellar core is of order a second [38]. In the case of a binary BH system, there will be a similar light-speed propagation time for information about the merger to reach the surrounding accretion disc. Hence there is an intrinsic component of the time delay set by the details of a high-energy astrophysics event, which is unknowable without detailed numerical modeling of the event and high levels of certainty for the source parameters.

Fortunately, this is where our strong lensing formalism can help. Consider a futuristic experimental scenario which detects the following four events, all confirmed as originating from the same sky location:

(i) t_a: time in a massless signal (e.g. a GW waveform) identified as the merger event.

(ii) t_b: peak flux in the accompanying massive particle signal.

(iii) t_c: merger time in a massless signal, with the same structure as the previous massless signal ("massless echo").

(iv) t_d: peak flux in a second massive particle signal, with the same flux variations as the previous massive particle signal ("massive echo").

The first two events here correspond to messengers arriving from the same image, which have traveled along the same lensed path. The latter two events correspond to messengers arriving from the second image, having traversed a different lensed path. For example, messengers traveling the upper path in Fig. 1 could give rise to t_a and t_b, and their lensed echoes traveling along the lower path give rise to t_c and t_d.

We focus on the following time intervals:

$$t_b - t_a = \delta\eta_{\text{intrinsic}} + \eta_{\text{total}}(\theta_1, m \neq 0) - \eta_{\text{total}}(\theta_1, m = 0) \qquad (24)$$

$$t_d - t_c = \delta\eta_{\text{intrinsic}} + \eta_{\text{total}}(\theta_2, m \neq 0) - \eta_{\text{total}}(\theta_2, m = 0) \qquad (25)$$

$$\mathcal{T} = (t_d - t_c) - (t_b - t_a)$$
$$= \eta_{\text{massive}}(\theta_2, m) - \eta_{\text{massive}}(\theta_1, m) \qquad (26)$$

where $\delta\eta_{\text{intrinsic}}$ represents the delay in emission between massive and massless messengers that is intrinsic to the source, as described above. The first equality of Eq. (26) defines the quantity $\mathcal{T}$, and the second equality uses Eq. (21). The intervals $t_d - t_c$ and $t_b - t_a$ correspond to emissions arriving from the same image of the source; they are expected to be small compared to $t_d - t_b$ or $t_c - t_a$, which correspond to emissions of the same kind arriving from different lensed images.

The intrinsic delay between the emission of massive and massless particles is a property of the source, and is not affected by the subsequent strong lensing. Hence it is the same for both lensed images, and therefore can be canceled out by the differencing strategy outlined above. However, this strategy assumes an idealistic situation in which the flux variations of both the massive and massless signals are well-sampled. In a real-world scenario this may not be possible—we will discuss this issue further in Sec. IV.

With a major source of error thus circumvented, we can now proceed to estimate the magnitude of the quantity $\mathcal{T}$, which we will term the *differential massive time delay*. In particular, we are interested to study the sensitivity of $\mathcal{T}$ to neutrino mass and the late-time cosmological acceleration (for those sources at cosmological redshifts). In the rest of this section we do this using two simple lens models. Although these may not be realistic as a global description of (say) galaxy clusters, our work here only requires a sufficient description of the innermost region of the lens.

In Appendix C we give a few relevant formulas for describing the properties of a lens model. These belong to the standard formalism of strong lensing and can be found in many introductory texts.

B. The singular isothermal sphere lens

The singular isothermal sphere (SIS) is one of the most commonly used toy lens mass models, and is a good approximation for the central regions of early-type galaxies [39]. It has the spherically symmetric three-dimensional density profile

$$\rho(r) = \frac{\sigma_v^2}{2\pi G r^2}, \qquad (27)$$

where the particles that constitute the lens have a Maxwellian velocity distribution with one-dimensional velocity dispersion σ_v. The SIS has some unusual properties: it has a constant 3D gravitational potential throughout, a density singularity at $r \to 0$ and infinite mass as $r \to \infty$.

These pathologies do not usually cause problems when we consider particles passing the lens at intermediate distances; however, it is the presence of the central singularity that allows SIS lenses to evade Burke's odd number theorem [37].

The constant potential of the SIS results in further interesting features. In particular, one finds that all rays reaching the lens undergo the same deflection toward the lens center [34]. That is, $\alpha(\theta)_{\text{scal}}$ is a constant:

$$\alpha_{\text{scal}}^{\text{SIS}}(\bar{\theta}) = \theta_E^{\text{SIS}} = 4\pi \left(\frac{\sigma_v}{c}\right)^2 \frac{D_{LS}}{D_S}, \qquad (28)$$

where θ_E^{SIS} is the Einstein radius of the lens; we will drop the label "SIS" for the remainder of this subsection. Using the formulas of Appendix C, it is fairly easy to derive the projected mean surface density and projected (i.e. 2D) gravitational potential of this model:

$$\bar{\kappa}(\theta) = \frac{\theta_E}{\theta} \qquad \psi(\theta) = \theta_E \theta. \qquad (29)$$

If the source angular position (β) lies within the Einstein radius, the SIS lens forms two images on opposite sides of the lens. Using Eq. (15), these are located at angular radii

$$\theta_{\pm} = \theta_E \pm \beta, \qquad (30)$$

where for convenience we have labeled the two images $\{\theta_+, \theta_-\}$ instead of $\{\theta_1, \theta_2\}$ [note that the "plus" path in Eq. (30) corresponds to the upper path in Fig. 1, which actually incurs the minus signs in Table I]. For the SIS lens only, the following relations then hold:

$$2\theta_E = \theta_+ + \theta_- \qquad 2\beta = \theta_+ - \theta_-. \qquad (31)$$

For a real strongly lensed system the image angular positions θ_+ and θ_- can be measured. In order to progress with our theoretical calculation, we will introduce an asymmetry parameter, $\mathcal{A}$, that quantifies the offset of the source position from θ_E, as follows:

$$\theta_- = \mathcal{A}\theta_E, \quad \text{where } 0 \leq \mathcal{A} \leq 1 \qquad (32)$$

$$\Rightarrow \theta_+ = \theta_E(2 - \mathcal{A}), \qquad (33)$$

where $\mathcal{A} = 1$ would imply a perfect Einstein ring system, for which the two images merge.

We can now evaluate Eq. (26) for the SIS model, using the results summarized above. The terms $\eta_{\text{undeflected}}$ and $\eta_{\text{massless}}(\bar{\theta})$ cancel out when we difference the time delays of massive and massless particles, as expected. The only term that contributes to the differential massive time delay in the SIS case is then η_{massive}. We obtain

$$\mathcal{T} = \frac{1}{2}\left(\frac{mc}{p_0}\right)^2 a_L^2 \frac{D_L D_S}{c D_{LS}} \theta_E(\theta_+ - \theta_-) \qquad (34)$$

$$= \left(\frac{mc}{p_0}\right)^2 a_L^2 \frac{D_L D_S}{c D_{LS}} \theta_E^2 (1 - \mathcal{A}) \qquad (35)$$

$$= \left(\frac{mc}{p_0}\right)^2 a_L^2 \frac{D_L D_{LS}}{c D_S} \left(4\pi \frac{\sigma_v^2}{c^2}\right)^2 (1 - \mathcal{A}), \qquad (36)$$

where in fact even the geometric contribution to the differential massive time delay has canceled, and we are left with only a pure Shapiro-like contribution. This vanishing of the geometric-like term is a unique feature of the SIS lens, due to its constant deflection angle; it does not occur for other lens models.

The left panel of Fig. 3 shows the evaluation of $\mathcal{T}$ for parameter values $m = 0.3$ eV, $p_0 = 10$ MeV, $\mathcal{A} = 0.75$ and standard ΛCDM cosmological parameters. We use the value $m = 0.3$ eV as a conservative upper bound on the neutrino mass, based on the constraint $\sum m_\nu = 0.23$ eV from the Planck satellite [40]. An energy of 10 MeV is consistent with the typical neutrino energies produced by accretion onto hypermassive neutron stars after a NS-NS merger [17,18] and in supernovae.

We note that a realistic situation would likely involve some spread in particle emission energies, and hence a dispersion in arrival times. In a similar vein, neutrino oscillations will ensure that even neutrinos emitted as an instantaneous burst are received with a spread of arrival times. The size of this dispersion could well be comparable or larger than the differential massive time delay we are pursuing. However, our quantity $\mathcal{T}$ is defined as a *difference* of multiple event timings, all of which will be dispersed in the same manner. Therefore, as long as the massive and massless fluxes are well-sampled—so that the peak of a dispersed signal can be located—our calculation remains unaffected [41]. We assume a futuristic scenario where such sampling is possible; of course, this may not be achievable, see Sec. IV. We do not consider here any effects relating to the structure of the neutrino hierarchy; see [22] for a discussion.

For comparison, the right-hand panel of Fig. 3 shows the standard, massless part of the time delay. For the parameter values under consideration here, this ranges between tens of days and tens of years. Note that, in order to maximize the small corrections of interest, we are considering higher source and lens redshifts than most optical strong lensing studies. This is the cause of some of our unusually large massless time delay values.

We see that, irrespective of the source redshift, the differential massive time delay peaks when the lens is located at redshifts around 0.2–0.5. The shape of the curves in the left-hand panel can be understood using Eq. (35) as follows: the prefactor of $D_L D_S/D_{LS}$ is shared with the massless time delay [see Eq. (22)], and imparts the broad,
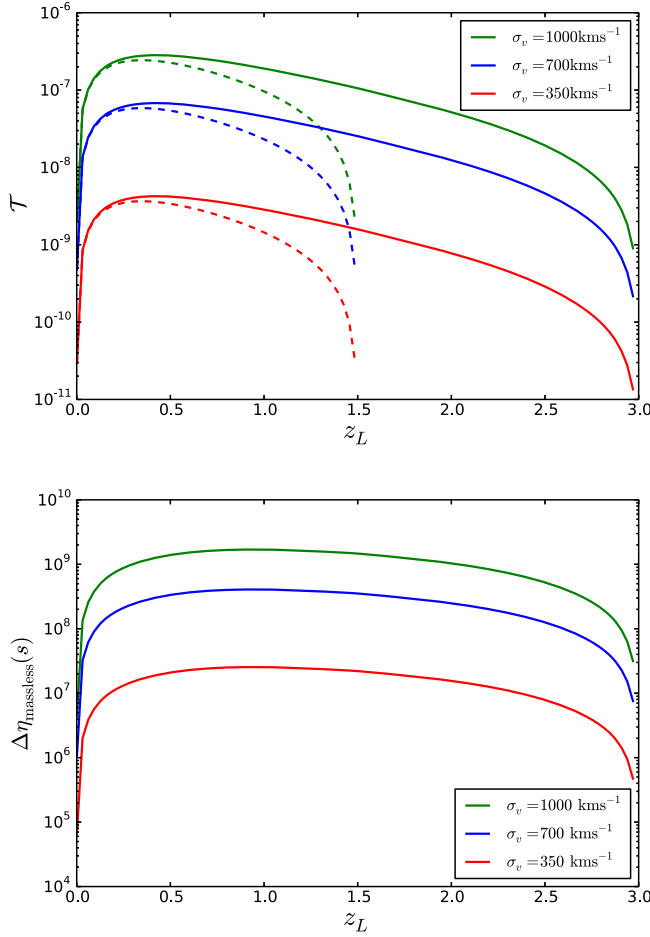
FIG. 4. The dependence of the differential massive time delay on the late-time expansion history, controlled via the energy density Ω_Λ. A flat cosmology is assumed in all cases. Solid lines represent a system with $z_S = 3.0$, whilst dashed lines are for $z_S = 1.5$. Particle properties are the same as in Fig. 3; the lens velocity dispersion used is $\sigma_v = 930 \text{ km s}^{-1}$ and the asymmetry parameter is $\mathcal{A} = 0.75$.

FIG. 3. Left: Differential massive time delay $\mathcal{T}$ [Eq. (26)] for a SIS lens. Solid curves are for a source galaxy at $z_S = 3.0$, whilst dashed curves are for $z_S = 1.5$. Note that z_L, the redshift of the lens, cannot exceed z_S. All curves are evaluated for parameters $m = 0.3$ eV, $\mathcal{A} = 0.75$, $p_0 = 10$ MeV. Planck 2015 cosmological parameters are used. Right: The massless part of the time delay, as would be measured in standard strong lensing studies. The velocity dispersion σ_v acts as a proxy for the mass of the lens. Only curves for $z_S = 3$ are shown.

flat shape seen in the right panel of Fig. 3. However, this shape is modulated by the appearance of θ_E^2 in Eq. (35): for a fixed source redshift, the Einstein radius—being an angular scale measured by the observer—decreases as the lens is moved to higher redshifts. This decline is responsible for the skew toward low z_L in the left panel of Fig. 3.

The differential massive time delay is also somewhat sensitive to cosmological parameters, as shown in Fig. 4. For a fixed source redshift, an increase in Ω_Λ boosts the conformal distances appearing in the numerator of Eq. (36). In the ideal scenario of having multiple well-understood multimessenger lensing systems with $z_L > 0.5$, the differential massive time delay could provide a new method to probe the equation of state of the dark energy sector (if assumptions are made about the
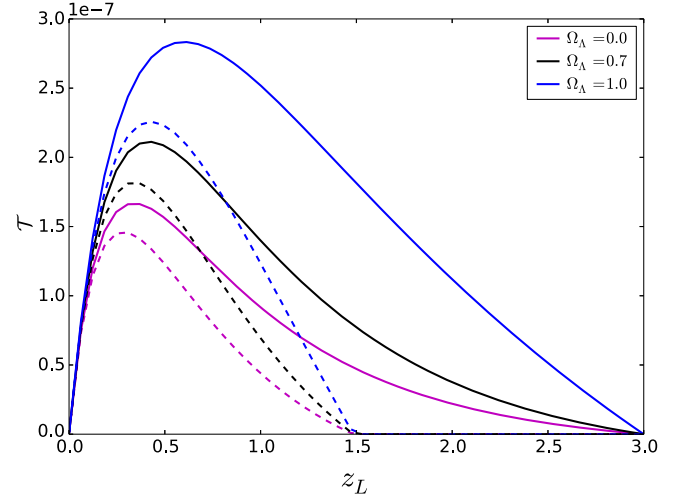
neutrino mass). This complements existing cosmological parameter constraints made using the massless time delay of photons [42,43]. There is also the possibility of testing for novel effects such as the violation of C, P and CP symmetries in gravity [44], which we will treat in a future investigation.

The velocity dispersion, σ_v has a very strong influence on the magnitude of $\mathcal{T}$—note that it appears to the fourth power in Eq. (36). In the SIS case σ_v acts as a proxy for the mass of the lens, suggesting that lensing by galaxy clusters (which generally have larger σ_v than individual galaxies) may be a more promising, albeit still challenging, candidate for a measurable differential massive time delay.

However, the selection of systems with a high velocity dispersion or mass must be balanced against the corresponding interval between the massless and massive echoes [i.e. the interval $t_c - t_b$ in Eq. (25)]. The lowest curve in the right-hand panel of Fig. 3 ($\sigma_v = 350 \text{ km s}^{-1}$) has a window of a few months between echoes, whilst for the uppermost curve ($\sigma_v = 1000 \text{ km s}^{-1}$) it is of order thirty years. We note that similar precision timing experiments spanning decades have already been carried out, for example, monitoring the inspiral rate of the Hulse-Taylor pulsar [45]. Unquestionably, though, this makes for an inconveniently slow experiment.

One can speculate on more exotic scenarios: if neutral particles heavier than neutrinos were emitted in conjunction with GWs or photons, then the effects discussed here could be orders of magnitude larger. As an illustrative example, consider a situation in which particles with the mass of a nucleon are produced during an event with energy similar
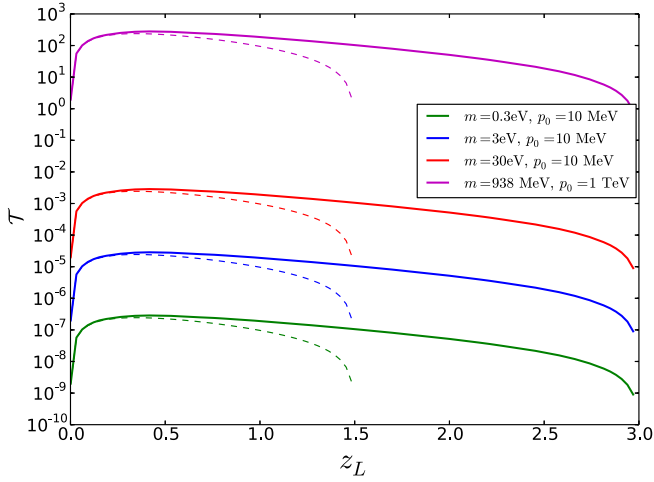
FIG. 5. The differential massive time delay for particles of varying mass and energy. The lowest curve corresponds to our fiducial case of $m = 0.3$ eV, $p_0 = 10$ MeV, representing a typical neutrino from a NS-NS merger. The uppermost curve corresponds to the case of a high-energy, neutral particle with the mass of a nucleon. Two intermediate cases are also shown. Solid and dashed curves are the same as previous figures. All curves use $\sigma_v = 930$ km s^{-1} and Planck cosmological parameters.

to that of a gamma-ray burst (GRB). Taking $m = 938$ MeV and $p_0 \sim 1$ TeV, the prefactor $(mc/p_0)^2$ is boosted by a factor of 10^8 and differential massive time delays of order tens of seconds become possible, see Fig. 5. Note that charged particles would be deflected by both Galactic and intergalactic magnetic fields, destroying the signals we are interested in here. See [46] for a discussion of neutral cosmic ray candidates.

Another hypothetical scenario would be to consider the time delays experienced by WIMPs such as axions, theoretical particles hypothesized to solve the strong CP problem of QCD, and appearing generically in string theory [47]. Most attention focuses on ultralight axions (10^{-33} eV–10^{-18} eV) as dark matter candidates, but heavier axions are possible (not as dark matter) and could be produced in high-energy events such as supernovae and NS mergers [48]. Axions with a mass of order 1 keV would experience a differential massive time delay of order seconds; however, we will not pursue such exotic scenarios further here.

We note that shortly after the present work appeared online, a letter by Fan *et al.* [49] was released, applying similar concepts to constraining the speed of propagation of gravitational waves.

C. The power-law lens

The first step in complexity beyond the SIS is the power-law (PL) lens model. This has the spherically symmetric density profile

$$\rho(r) = \rho_0 \left(\frac{r_0}{r}\right)^n, \qquad (37)$$

where the SIS lens is recovered for $n = 2$. Like the SIS, the PL lens has an infinite central cusp that is not problematic for our current work. This can be alleviated, if desired, by the use of softened power-laws such as $\rho \propto (r^2 + s^2)^{\frac{n}{2}}$, where s is a constant [50].

For some values of n, the PL lens can produce more than two images, e.g. for $n = 1.5$, three images are present when the source lies inside the tangential critical curve of the lens plane. To facilitate discussion with the SIS case, in this paper we will use only two of these images (specifically, the two at greatest radial distance from the lens center). We note, though, that a third image—if resolvable—offers further differencing possibilities that could be used either as a check of nuisance parameters, or to provide multiple measurements of the differential massive time delay.

Referring to the standard lensing definitions given in Appendix C, the PL lens has the following potential, scaled deflection and convergence profiles [51]:

$$\alpha_{\rm scal}(\theta) = \theta_E \left(\frac{\theta}{\theta_E}\right)^{2-n} \qquad \psi(\theta) = \frac{\theta_E^2}{3-n} \left(\frac{\theta}{\theta_E}\right)^{3-n} \qquad (38)$$

$$\kappa(\theta) = \frac{3-n}{2} \left(\frac{\theta}{\theta_E}\right)^{1-n} \qquad \bar{\kappa}(\theta) = \left(\frac{\theta}{\theta_E}\right)^{1-n}. \qquad (39)$$

Equations (31) no longer hold, but we will define analogous quantities (though note a factor of 2 difference in the second definition):

$$2\langle\theta\rangle = \theta_+ + \theta_- \qquad \Delta\theta = \theta_+ - \theta_-. \qquad (40)$$

In what follows, we will assume that the annulus enclosed by the two images is narrow compared to their offset from the lens center, i.e. $\Delta\theta \ll \theta_+, \theta_-$. This is reasonable, since highly asymmetric lensing systems usually have at least one strongly demagnified image, and are therefore less likely to be identified.

The Einstein radius is now given by [51]

$$\theta_E = \left(\frac{\theta_+ + \theta_-}{\theta_+^{2-n} + \theta_-^{2-n}}\right)^{\frac{1}{n-1}}. \qquad (41)$$

As we did for the SIS lens, we will parametrize one of the image positions in terms of the Einstein radius as $\theta_- = \mathcal{A}\theta_E$. The narrow-annulus approximation above then allows us to expand in the quantity $\Delta\theta/\theta_E$ (which will also be small for values of $\mathcal{A}$ close to 1. See Sec. 4.3 of [51] for a similar expansion). A little algebra with Eq. (41) leads to the following relations:

$$2\langle\theta\rangle = \mathcal{A}^{2-n}\theta_E \left[2 + (2-n)\frac{\Delta\theta}{\mathcal{A}\theta_E}\right] \qquad (42)$$

$$\Delta\theta = \frac{2\mathcal{A}\theta_E[\mathcal{A}^{1-n} - 1]}{1 - \mathcal{A}^{2-n}(2-n)} \qquad (43)$$

$$\mathcal{T} \equiv \mathcal{T}_{\text{geo}} + \mathcal{T}_{\text{Shap}} \tag{44}$$

$$= \frac{1}{2} \frac{m^2 c^2}{p_0^2} \Delta\theta \theta_E \mathcal{A}^{2-n} \left\{ (n-2)\mathcal{A}^{1-n} \left[\int_1^{a_L} \frac{1}{H} da \right. \right.$$
$$\left. \left. + \frac{D_L^2}{D_{LS}^2} \int_{a_L}^{a_S} \frac{1}{H} da \right] + a_L^2 \frac{D_L D_S}{c D_{LS}} \right\}. \tag{45}$$

We see that the geometric contribution to $\mathcal{T}$ vanishes for $n = 2$, in agreement with Sec. III B.

Figure 6 shows the influence of the power-law index, n, on the geometric and Shapiro-like contributions to the differential massive time delay. The Shapiro contribution always remains positive, whilst the geometric contribution switches sign about its vanishing point at $n = 2$. For $n < 2$ there can be a significant degree of cancellation between the two contributions, whilst for $n > 2.5$ the geometric contribution is dominant.

Physically, the sign change in the geometric contribution arises because of the behavior of $\alpha(\theta)$ in Eq. (38). Returning to the simple two-image picture of Fig. 1, one image will sit inside the Einstein radius and one exterior to it. For $n > 2$, rays from the image inside the Einstein radius will have experienced the greatest deflection at the lens. This corresponds to the intuitive picture that a ray passing close to the center of the lens, where the density is highest, will be more strongly deflected than one passing "further out." We then expect signals from the θ_- image (lower path in Fig. 1) to arrive after those from the θ_+ image.

In the case of $n < 2$ the mass of spherical shells increases with radii. Because the majority of the lens mass is now situated at large radii from its center, the image appearing at $\theta > \theta_E$ experiences the greater deflection [see the first of Eqs. (38)]. The $n = 1$ case corresponds to a uniform critical sheet, whilst the central regions of galaxies are sometimes modeled using $1 \lesssim n \lesssim 2$ [51]. Based on purely geometrical considerations, one would now expect signals from the θ_- image to arrive first; hence $\mathcal{T}_{\text{geo}}$ changes sign. However, the Shapiro-like contribution to the differential massive time delay has the potential to contradict this intuition, if it is large enough to outweigh the geometrical term.

D. Magnifications

As well as deflecting emissions onto multiple paths, gravitational lenses are able to focus (or sometimes defocus) a bundle of rays en route to the observer. For extended electromagnetic sources, the focusing of rays results in a decreased image area and hence, due to the conservation of surface brightness, a boost in flux. Magnification occurs similarly for point-like sources, though the situation is slightly different (as clearly there can be no change in image area): the brief explanation is that rays which otherwise would not intersect the observer now do so, due to their deflection at the lens.
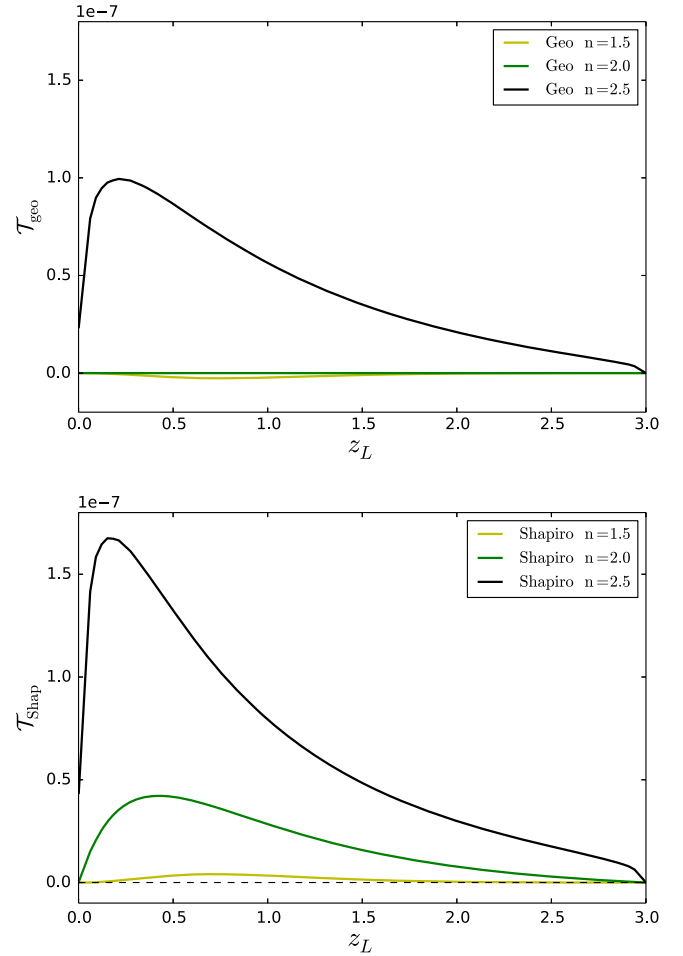

FIG. 6. Contributions to the relative time delay for a power-law lens, with density profile $\rho \propto r^{-n}$. Both panels have particle parameters $p_0 = 10$ MeV, $m = 0.3$ eV, and lens parameters $z_S = 3.0$, $\mathcal{A} = 0.95$, $\rho_0 = 2 \times 10^{14} \, M_\odot \, \text{Mpc}^{-3}$ and $r_0 = 0.4$ Mpc. Left: The geometric contribution, i.e. the first two terms of eq. (46). This vanishes at $n = 2$, the SIS case, and changes sign either side of this value. Right: The Shapiro contribution, i.e. the third term of eq. (46). The difference between the $n = 2$ case here and in Fig. 3 is due to the different values of $\mathcal{A}$ used. Note that in the case of $n = 1.5$ a third image is also present, and here we are only considering the differential massive time delay between two of them—see text.

The magnification is defined as the ratio of the lensed to the unlensed flux. Although the source flux is clearly a function of frequency, the spectral shape is preserved by lensing and hence μ is independent of frequency. As detailed in standard lensing texts, if the mapping of a point from the image plane to the source plane is given by the 2D vector function $\vec{\beta}(\vec{\theta})$, then the magnification factor is

$$\mu = \frac{1}{|\det \mathbf{A}|}, \quad \text{where } \mathbf{A} = \frac{\partial \vec{\beta}}{\partial \vec{\theta}}. \tag{46}$$

For an axisymmetric source we have [34]:

$$\det \mathbf{A} = (1 - \bar{\kappa})(1 + \bar{\kappa} - 2\kappa), \qquad (47)$$

where the dimensionless surface densities (equivalent to convergence), κ and $\bar{\kappa}$, are defined in Appendix C.

Since they follow the same null geodesics as photons, these expressions should apply equally well to GWs. The only two requirements are that the geometric optics approximation remains valid, and that we avoid the exceptional case of a perfect Einstein ring system, for which $\kappa \to 1$, and hence the magnification above formally diverges. Takahashi [52] estimates that for a GW of characteristic frequency f, the geometric optics approximation holds for lens masses above $10^5 \ \mathrm{M}_\odot \ (f/\mathrm{Hz})^{-1}$; since we are using galaxy clusters $(\sim 10^{14} - 10^{15} \ \mathrm{M}_\odot)$ as our lenses and NS binaries $(f \sim 10\text{–}1000 \ \mathrm{Hz})$ as sources, we are always safely in this regime.

The only exclusion from our calculations, then, is the finely tuned case of $\theta = \theta_E$ exactly; this requires a full wave optics treatment to remove the apparent divergence [53]. We will not make this digression here, but one can rest easy that the singularities seen in Fig. 7 below remain finite in a fully correct treatment.

Taking Eqs. (38), (39) and (47) together, we will calculate the magnification of the images produced by a power-law lens. We will continue to use the notation of Sec. III, that is, we write $\theta_2 = \mathcal{A}\theta_E$ (note that the asymmetry parameter $\mathcal{A}$ should not be confused with the Hessian matrix $\mathbf{A}$). Formally, for a given value of $\mathcal{A}$, the location of the other image can be determined (in terms of θ_E) by solving Eq. (41). In practice, this is awkward to do analytically except in special cases such as $n = 2, 3$, etc. Hence we shall make the same restrictions and approximations as used in Sec. III C, and study systems for which the images are separated by a narrow annulus such

that $\Delta\theta = \theta_+ - \theta_- \ll \theta_+, \theta_-$. Under these conditions, we quickly arrive at:

$$\det \mathbf{A}|_{\theta_+} \approx \left[1 - \mathcal{A}^{(1-n)}\left\{1 + (1-n)\frac{\Delta\theta}{\mathcal{A}\theta_E}\right\}\right]$$
$$\times \left[1 + (n-2)\mathcal{A}^{(1-n)}\left\{1 + (1-n)\frac{\Delta\theta}{\mathcal{A}\theta_E}\right\}\right]$$
$$(48)$$

$$\det \mathbf{A}|_{\theta_-} \approx [1 - \mathcal{A}^{(1-n)}][1 + (n-2)\mathcal{A}^{(1-n)}], \qquad (49)$$

where $\Delta\theta$ is given by eq. (43). Note that the above two lines then depend solely on the alignment of the lensing system and the density profile of the lens. As usual, we can see that the SIS case, $n = 2$, simplifies the above expressions considerably.

Figure 7 shows the total magnification of the power-law lens, which sums over the individual magnification of all images:

$$\mu_{\mathrm{Tot}} = \sum_i (|\det \mathbf{A}|_{\theta_i}|)^{-1} \qquad (50)$$

$$\equiv (|\det \mathbf{A}|_{\theta_+}|)^{-1} + (|\det \mathbf{A}|_{\theta_-}|)^{-1}. \qquad (51)$$

For our fiducial case of $n = 2$ and $\mathcal{A} = 0.95$ the total magnification is just under 40. This is split roughly evenly between the two images, though the image inside the Einstein radius is slightly brighter: $\mu_+ \simeq 17.2$ and $\mu_- \simeq 19$. In general we cannot simply extended our formulas through $\mathcal{A} = 1$, as we expect the number of images to change at critical curves in the image plane (such as $\theta_E = 1$). However, the constraints of Eqs. (31) imply that in the SIS case the two images are equally displaced from θ_E, so we can think of $A > 1$ as simply a choice to parametrize the outermost image instead of the innermost one. The corresponding magnification plot would then just be a reflection of Fig. 7 about the axis $\mathcal{A} = 1$. We note in passing that the magnification values discussed here are comparable to the recent detection of SN iPTF16geu, the first multiply-imaged Type Ia SN, with a total magnification $\mu \sim 56$ [54].

For massive particles, the relevant observable is the specific particle intensity, measured in units of $\mathrm{m}^{-2} \ \mathrm{s}^{-1} \ \mathrm{J}^{-1} \ \mathrm{sr}^{-1}$ (though the sr^{-1} is irrelevant for an effective point source), or more conveniently the flux. This too will be boosted by a magnification factor very similar to that of the GWs and photons. Any small differences in the magnification factor of non-null particles compared to null ones are likely to be dominated by uncertainties in the particle luminosity of the source; hence, to a first approximation, we can assume massless and massive particles to experience the same magnification factor.

Dai *et al.* [55] have emphasized that for LIGO there exists a degeneracy between a lensed GW from a high-redshift, low-mass source and an unlensed GW from a low-redshift, high-mass source. They argue that in the
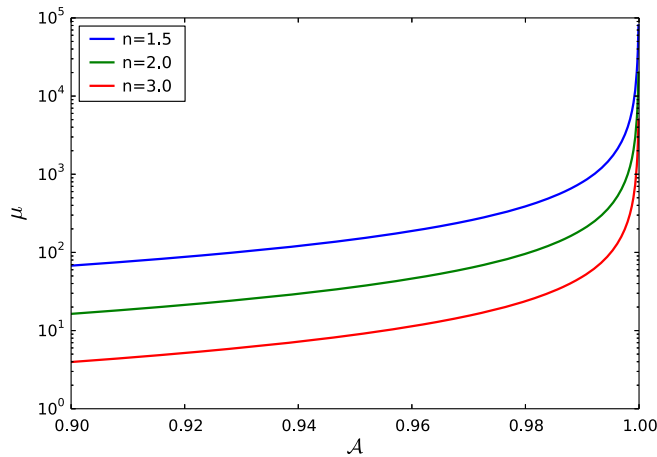

FIG. 7. The total magnification (summed over both images) for the power-law lens. $\mathcal{A} = 1$ corresponds to a perfect Einstein ring system; in the geometric optics limit, the magnification of a point source then becomes infinite. This breakdown signals the need for a wave optics treatment. Note that the magnification factor depends solely on $\mathcal{A}$ and n.

lensed case the GW echo will be registered as a separate GW event, and hence will not be of use in breaking the degeneracy. We note that, in the more futuristic scenario considered here, detection of the corresponding massive echo could in principle help to confirm candidate lensed GWs and hence break this degeneracy.

IV. SYSTEMATICS AND SUBTLETIES

Thus far, our calculations and discussion have been based around idealized toy models. We stress that this paper is intended to be a largely theoretical discussion of an interesting phenomenon in fundamental physics, and is not an observational call to arms. Nevertheless, in this section we will discuss some confounding factors that would require careful treatment in a futuristic attempt to measure the differential massive time delay.

A. Source redshifts

Our expression for the differential massive time delay [Eq. (19)] depends on the redshifts of the source and lens. The lens redshift is expected to be measurable from an electromagnetic counterpart (e.g. a massive galaxy cluster at low redshift), but the same is not necessarily true of the source redshift. In the case of an asymmetric supernovae this *is* likely to be possible, though the electromagnetic counterpart would only become visible some time (hours to days) after the GW and neutrino signals, due to the photon escape time of the supernovae remnant.

NS mergers are a candidate for the production of short gamma-ray bursts (GRBs), and so too may have an electromagnetic counterpart in a few select cases. However, GRBs are thought to be beamed in a dipolar fashion, whilst GWs in GR are of course quadrupolar. This implies that only a fraction $\sim 10^{-3}$ of the NS-NS merger GW signals we receive should be accompanied by a GRB [56].

In the absence of an electromagnetic counterpart, and with current ground-based detectors, the mass and redshift of a compact merger are famously degenerate. Specifically, the waveform constrains only the redshifted mass combination $M_z = M(1 + z)$. As we discussed in Sec. III D, this degeneracy is preserved even in the case of gravitationally lensed GWs.

However, the authors of [56] have identified a method to break this degeneracy that may be achievable by future GW detectors, at least in the NS-NS merger case. In brief, the method relies on studying two effects in the GW waveform: (i) corrections to the orbital phase due to tidal effects during the inspiral stage, and (ii) prominent spectral features during the post-merger HMNS phase. These two effects have different dependencies on the *true* total mass of the binary and its redshift, leading to near-orthogonal contours in the (M, z) parameter plane (see Fig. 1 of [56]).

Assuming progress in determining the NS equation of state, the authors of [56] report a mass determination within

1% accuracy for all the cases they considered. It is reasonable to speculate that when/if experiments become sufficiently mature to measure the differential massive time delay, either (a) the redshift of NS-NS mergers will be measurable from GWs alone, or (b) optical counterparts will be frequently available.

B. Lens identification

Though it may not be necessary to have an electromagnetic counterpart for the source, it is essential for the lens. Inferring the lens mass distribution from sheared optical images is now a mature field [57], and is crucial to step beyond the simple analytic density prescriptions used in the present work.

Furthermore, the optical lens image will help with sky localization of the lensed GWs and neutrinos. For example, the Large Synoptic Survey Telescope (LSST [58]) is expected to detect cluster-scale strong lenses at roughly 0.04–0.1 sq $\cdot$ deg^{-2} [59], whilst a future GW detector network of LIGO+VIRGO+KAGRA [60,61] should be able to localize GW sources to less than 10 sq. degrees [62]. Hence the combined observations should be able to pin down a cluster-scale lens to within a handful of candidates. In the case of more than one candidate lens, estimates of their redshifts and masses will be necessary to further ascertain which belongs to the system of interest. However, once the correct lens has been identified, the source of the lensed GWs and neutrinos can then be localized to broadly lie within the area occupied by the optical images, at most tens of arcseconds. Triangulation using a network of neutrino detectors is not expected to provide significant enhancements over the localization capabilities of a single neutrino detector [63], making the process of lens identification particularly crucial.

The influence of multiple lensing events would also need to be modeled. In this paper we have considered a situation in which the multimessenger signals encounter only a signal, large lens during their propagation. In reality, they are likely to additionally encounter many smaller weak lensing events by intervening matter structures [7,30,64], and some of these will vary over the spatial region spanned by our effective images. This will add some intrinsic scatter to the arrival time of the multimessenger echoes.

In addition, one should also model the Shapiro delay induced by the gravitational potential well of the Galaxy [65,66]. Whether uncertainties in these effects eradicate the differential massive time delay signal requires a detailed study [4].

C. Detectors and count rates

Solar and atmospheric neutrino backgrounds are an obvious systematic for detection of the lensed neutrino signals discussed here; for example, the dominant solar neutrino background at ~ 10 MeV is a flux of roughly

5×10^6 cm^{-2} s^{-1} from the decay of Boron-8. Fortunately, extracting neutrino candidates of interest from underneath these backgrounds is a well-studied topic [67]. The solar neutrino background can be minimized by concentrating on neutrino events with energies $\gtrsim$20 MeV, though this comes at the expense of a lowered count rate from our sources of interest. We note in passing that the SuperKamiokande collaboration were able to identify and dismiss four neutrino events that were candidates for association with the gravitational wave events GW150914 and GW151226 [68].

Our method for cancelling out the unknown intrinsic source delay (Sec. III A) should work well in the case of well-sampled fluxes, e.g. if the emissions occurs as short, intense bursts. If the emissions occur over a longer time-scale and is poorly sampled, there will be difficulties in recovering the matching shapes of the original signal and its echo(es). Not only does this make successful identification of the massive echo less likely, but it will also introduce an additional source of error in the measurements of times t_b and t_d, and hence into $\mathcal{T}$.

Undoubtedly low neutrino flux counts are a major obstacle for the extragalactic sources we have talked about in this paper. It may be that the differential massive time delay will only ever be measurable for Galactic sources, for which up to $\sim 10^5$ neutrino events are expected with future kiloton-scale detectors. Local-group sources should also be within reach: a future neutrino experiment of several hundred kilotons is expected to detect a few dozen neutrinos from an event in the Andromeda galaxy [69]. For these sources there is no possibility of constraining cosmological parameters, but arguably any bounds on the neutrino mass should be cleaner without such degeneracies. Beyond the local group begins a battle between the flux scaling as $1/$distance^2 and the number of sources increasing as approximately distance^3. We will not attempt a prediction of the expected outcome in the present work.

Note that the typical neutrino energies considered in this paper ($\sim$10 MeV) are below the detection threshold of some current major detectors such as IceCube and ANTARES [70,71], though SuperKamiokande, KamLAND and the Sudbury Neutrino Observatory all have thresholds of order a few MeV [72–74]. Whilst GRBs associated to compact object mergers may produce high-energy neutrinos ($\sim$TeV–PeV) detectable by IceCube and ANTARES, the differential massive time delay associated to these will be miniscule. Still, detection of such high-energy neutrinos may help with on-sky source localization for the lower-energy neutrino counterparts [75].

Given the magnitude of the differential massive time delay ($\sim$0.1 μs), it will be necessary to know the distance between all detectors involved in the measurement to within a metre or so (simply considering the magnitude of $c\mathcal{T}$). Based on current and improving GPS sensitivities, this should not prove problematic.

V. DISCUSSION

We have begun here an exploration of what might be learned from extreme but rare astrophysical events by future observatories detecting massless carriers, such as photons or gravitational waves, as well as massive ones, such as neutrinos or other neutral particles. In particular, we have considered the information provided by strong gravitational lensing of such signals by large massive bodies close to the line of sight to such events.

We can draw an analogy between the present status of GW astronomy and the early days of CMB detection. In the year 2000, shortly after the first CMB acoustic peak was detected by BOOMERANG and MAXIMA [76,77], it would have seemed absurdly optimistic to consider measuring the eighth acoustic peak with the precision now achieved by the Planck satellite. Yet, thanks to continual improvements in detector technology and data analysis techniques, intervening experiments (WMAP [78]), and increasingly sophisticated understanding of systematics, such exquisite measurements are now possible. Similarly, though measuring the effects discussed in this paper is unfeasible with present understanding and experiments, we envision an equally rapid progression for the forthcoming decade(s) of multimessenger astronomy. One of our goals here is to stimulate forward-thinking about the novel science that might be possible in future.

To this end we have derived an expression for the differential arrival time of massive and massless particles with a common origin. The resulting expression is sensitive to particle properties, cosmological parameters, and the masses and separations of elements in the lensing system. Though we have only evaluated the magnitude of this correction for simplified lens models, it could be applied to real lensing systems whose mass distribution is relatively well-constrained. In the examples studied here, the differential time delay is found to have a magnitude of order 0.1 μs. We note a powerfully general paper by Fleury [79] offers explanation as to why time delay effects remain small, even when the angular deflection and magnification of messengers is substantial.

Neutrino detectors are already capable of measuring intervals of order 0.1 μs, having at present a time resolution down to 100 ns. GW detectors lag behind somewhat—the timing resolution of the LIGO detectors is currently at the order of 100 μs [80,81]. However, with the broader frequency ranges proposed for next-generation detectors like the Einstein telescope, and improvements in detector technology, it seems reasonable to speculate that the necessary precision might be available to future GW experiments.

In addition, we note that a key feature of the differential massive time delay is a near-coincident feature in multiple media and multiple detectors. Although clearly this requires a global coordination of experiments, the simultaneous nature of events should assist with the selection and rejection of candidate detections.

For ease of discussion, we have generally referred to extragalactic merging neutron stars as a candidate source. However, we remind the reader that—under the right conditions—NS-BH mergers, BH-BH mergers, and asymmetric supernovae are all potential alternative candidates. Likewise, we have generally focused on GWs as the relevant massless messenger; photons can also take this role (and would arguably be easier to work with), if one is certain that prompt emission is being detected.

We have not attempted here a forecast of the constraints attainable on neutrino masses or cosmological parameters from measurements of the differential massive time delay. To do so would require a detailed description of the experiments involved; we suspect that instruments with the required sensitivity are not even at the blueprint stage yet. Although designs for the LISA observatory [82] are progressing rapidly, the most promising sources listed above fall outside of its frequency range.

Even if the differential massive time delay is not within foreseeable experimental reach for any source types, or in fact will always be lost to source uncertainties and systematics, the lensing of massive particles seems an inherently interesting and underexplored counterpart to the extensive research in optical lensing (and the much less-studied field of GW lensing).

Several logical extensions of the present paper would be

(i) A detailed study of the joint redshift distributions of candidate sources and massive galaxy clusters, resulting in an estimate for the number of simultaneous GW-neutrino lensing systems in principle detectable at Earth;

(ii) Further investigation into the projected sensitivities of future GW and neutrino detectors, their timing resolutions and sky localization errors, particularly when operating as a network;

(iii) A more sophisticated treatment of the massive particle flux expected from the sources listed in the introduction, accounting for the energy spectrum of different species and the effects of poor sampling of the burst at detection;

(iv) The possibility that stacking signals from separate events could alleviate the issue of low count rates from extragalactic sources. For example, one might imagine stacking signals from all systems that have a source and lens in the same redshift bin. This might allow a first measurement of the extragalactic differential massive time delay, even if the stacking technique means that constraining fundamental parameters is not possible;

(v) A consideration of the further differencing opportunities presented by lensing systems with more than two images (Sec. III C);

(vi) Investigation into the potential of the differential massive time delay and related phenomena to constrain the violation of C, P and CP symmetries in gravity.

We hope to take up some of these questions in future work.

APPENDIX A: UNLENSED MASSIVE TIME DELAY

Here we provide a basic calculation of the difference between the propagation time of massive and massless particles *in the absence* of any lensing effects. See [7] for a calculation that incorporates the effects of cosmological perturbations. We begin from the familiar flat FRW line element:

$$ds^2 = -\mathrm{d}t^2 + a(t)^2[\mathrm{d}r^2 + r^2(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\phi^2)]. \quad (A1)$$

For pure radial motion, the timelike component of the geodesic equation is

$$\frac{\mathrm{d}^2t}{\mathrm{d}\lambda^2} + a\dot{a}\left(\frac{\mathrm{d}r}{\mathrm{d}\lambda}\right)^2 = 0, \quad (A2)$$

where λ is an affine parameter. The normalization of the four-velocity for a massive particle gives us

$$u^\mu u_\mu = -\left(\frac{\mathrm{d}t}{\mathrm{d}\lambda}\right)^2 + a^2\left(\frac{\mathrm{d}r}{\mathrm{d}\lambda}\right)^2 = -1, \quad (A3)$$

and combining the two equations above to eliminate $\mathrm{d}r/\mathrm{d}\lambda$, we obtain

$$\frac{\mathrm{d}^2t}{\mathrm{d}\lambda^2} + \frac{\dot{a}}{a}\left[\left(\frac{\mathrm{d}t}{\mathrm{d}\lambda}\right)^2 - 1\right] = 0. \quad (A4)$$

Integrating this leads to (where C is a constant):

$$\left(\frac{\mathrm{d}t}{\mathrm{d}\lambda}\right)^2 - 1 = \frac{C}{a^2}. \quad (A5)$$

We multiply this by m and use the definition of four-momentum $P^\mu = mu^\mu = m\mathrm{d}x^\mu/\mathrm{d}\lambda$, $P^0 = E$ to yield

$$E^2 - m^2 \equiv g_{ij}P^iP^j = \frac{Cm^2}{a^2}, \qquad \text{(A6)}$$

from which we see that the magnitude of the spatial three-momentum redshifts as $1/a$:

$$p = \sqrt{g_{ij}P^iP^j} = \sqrt{C}\frac{m}{a}. \qquad \text{(A7)}$$

We will need this result shortly. Now, to find the time taken for a neutrino to travel a cosmological distance, Eq. (A3) can be rewritten as:

$$1 - \left(\frac{d\lambda}{dt}\right)^2 = 1 - \frac{m^2}{E^2} = a^2\left(\frac{dr}{dt}\right)^2 \qquad \text{(A8)}$$

$$\Rightarrow \frac{dr}{dt} = \frac{\sqrt{E^2 - m^2}}{aE}. \qquad \text{(A9)}$$

Using our result from Eq. (A7), we can write $p = a_0 p_0/a$ where a_0 and p_0 are defined at some fixed time, so that

$$\frac{dr}{dt} = \frac{p}{a\sqrt{m^2 + p^2}} = \frac{a_0 p_0}{a^2\sqrt{m^2 + \frac{a_0^2}{a^2}p_0^2}}. \qquad \text{(A10)}$$

For convenience we define $y_0 = a_0 p_0$, in terms of which the above expression then rearranges as:

$$dr = \left[a\sqrt{1 + \frac{a^2 m^2}{y_0^2}}\right]^{-1} dt. \qquad \text{(A11)}$$

Integrating this expression would give us the conformal distance traveled by a massive particle with initial spatial momentum p_0 in a time interval t.

Now we relate this to the analogous, simpler expression for a massless particle; for linguistic convenience we will refer to a photon, but our results apply equally to GWs. We know that the total *conformal* distances traveled by the photon and the neutrino are the same. In principle, the physical distances traveled by the photon and the massive particle are different, since the universe continues to expand during the small time interval between their arrival at Earth. Equating the conformal distances, then:

$$r = \int_{t_0}^{t_\nu} \frac{1}{\sqrt{1 + \frac{a^2 m^2}{y_0^2}}} \frac{dt}{a} \equiv \int_{t_0}^{t_\gamma} \frac{dt}{a}, \qquad \text{(A12)}$$

where t_0 is the (idealized) simultaneous time of emission, t_γ is the time the photon arrives at Earth, t_ν is the time of arrival of the massive particle, and all times correspond to those measured by a comoving observer.

For all the examples discussed in this paper, the mass of the particle is much smaller than its initial energy. This correspondingly implies $m^2 \ll y_0^2$, such that:
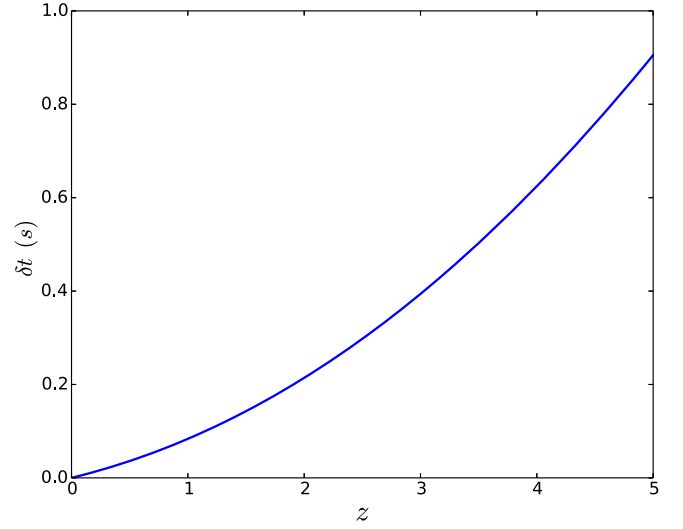

FIG. 8. The difference in arrival time between an unlensed massless particle (photon or gravitational wave) and unlensed massive particle with $m = 0.3$ eV, $E_0 = 10$ MeV, as a function of source redshift. Even for high-redshift sources, the difference in arrival times remains less than a second.

$$\left[1 + \left(\frac{am}{y_0}\right)^2\right]^{-\frac{1}{2}} \approx 1 - \frac{1}{2}\left(\frac{a}{a_0}\frac{m}{p_0}\right)^2 + \mathcal{O}\left[\frac{m^4}{p_0^4}\right]. \qquad \text{(A13)}$$

Substituting this into eq. (A12) yields

$$r = \int_{t_0}^{t_\nu} \left[1 - \frac{1}{2}\left(\frac{a}{a_0}\frac{m}{p_0}\right)^2\right] \frac{dt}{a} \equiv \int_{t_0}^{t_\gamma} \frac{dt}{a}. \qquad \text{(A14)}$$

This integral can of course be evaluated exactly. However, for analytic purposes it is helpful to use the fact that $t_\gamma \simeq t_\nu - \delta t$. Then the right-hand side can be written:

$$\int_{t_0}^{t_\nu - \delta t} \frac{dt}{a} = \int_{t_0}^{t_\nu} \frac{dt}{a} - \int_{t_\nu - \delta t}^{t_\nu} \frac{dt}{a} \qquad \text{(A15)}$$

$$\approx \int_{t_0}^{t_\nu} \frac{dt}{a} - \frac{1}{a(t_\nu)}\delta t, \qquad \text{(A16)}$$

where in the second line we have used the fact that the time interval δt is very small compared to the cosmological expansion time. Using this in Eq. (A14) and cancelling terms on either side:

$$\int_{t_0}^{t_\nu} \left[-\frac{1}{2}\left(\frac{a}{a_0}\frac{m}{p_0}\right)^2\right] \frac{dt}{a} = -\frac{1}{a(t_\nu)}\delta t. \qquad \text{(A17)}$$

Normalizing the scale factor such that $a(t_\nu) = 1$ today and converting the integral to be with respect to redshift, we obtain

$$\delta t = \frac{1}{2}\left(\frac{m}{p_0}\right)^2 \int_0^{z_0} \left(\frac{1+z_0}{1+z}\right)^2 \frac{dz}{H(z)}, \qquad (A18)$$

where z_0 is the source redshift. To the accuracy that we are working here, we can take $p_0 \approx E_0$ in the above expression, where E_0 is the initial energy of the particle.

The resulting difference in arrival time between a massless particle and one with $m = 0.3$ eV (the maximal neutrino mass) and $E_0 = 10$ MeV is shown in Fig. 8. As can be seen, even for high-redshift sources the difference in arrival times remains of order a second.

APPENDIX B: LENSING OF MASSIVE PARTICLES BY A POINT MASS

In this appendix we derive the modification to the well-known formula for the lensing of a massless particle by a point mass M, i.e. $\alpha = \frac{4GM}{c^2}$, for a massive particle. The classic derivation for the massless case can be found in many introductory GR texts, e.g. [83]. Thanks to the two-dimensional equivalent of Birkhoff's theorem, the result extends to any axially symmetric mass distribution interior to the trajectory of the lensed particle. Extended lenses that encompass the particle trajectory (e.g. a galaxy cluster) require a more sophisticated treatment, though the essential conclusions of this section remain the same.

As in Sec. II, we can account for both massless and massive particle cases by writing the normalization of the four-velocity as:

$$u^\mu u_\mu = -\epsilon \qquad (B1)$$

and specifying $\epsilon = 1$ for a massive particle, $\epsilon = 0$ for a photon, say. Expanding the above expression in a Schwarzschild metric and setting $G = c = 1$:

$$-\left(1 - \frac{2M}{r}\right)\left(\frac{dt}{d\lambda}\right)^2 + \left(1 - \frac{2M}{r}\right)^{-1}\left(\frac{dr}{d\lambda}\right)^2$$
$$+ r^2\left(\frac{d\phi}{d\lambda}\right)^2 = -\epsilon, \qquad (B2)$$

where λ is an affine parameter and we have chosen the motion of our lensed particle to be in the $\theta = \pi/2$ plane. We use the two Killing vectors of the Schwarzschild metric, denoted here as $\vec{\xi}$ and $\vec{\chi}$ to find the usual energy and angular momentum conserved quantities:

$$\vec{\xi} = \{1, 0, 0, 0\} \qquad \vec{\chi} = \{0, 0, 0, 1\} \qquad (B3)$$

$$e = -\vec{\xi} \cdot \vec{u} = \left(1 - \frac{2M}{r}\right)\frac{dt}{d\lambda} \qquad (B4)$$

$$\ell = \vec{\chi} \cdot \vec{u} = r^2 \frac{d\phi}{d\lambda}. \qquad (B5)$$

For a massive particle, $\vec{u}$ is the four-velocity. For a photon, one can choose the affine parameter such that $\vec{u}$ coincides with the four-momentum of the photon. Substituting these conserved quantities into Eq. (B2) and rearranging, we obtain:

$$\left(\frac{dr}{d\lambda}\right)^2 = e^2 - \left(1 - \frac{2M}{r}\right)\left(\frac{\ell^2}{r^2} + \epsilon\right). \qquad (B6)$$

Let us replace e by something directly measurable (in principle). When the particle is at its closest approach to the lensing object we have $dr/d\lambda = 0$, leading to:

$$e^2 = \left(1 - \frac{2M}{R_{\min}}\right)\left(\frac{\ell^2}{R_{\min}^2} + \epsilon\right), \qquad (B7)$$

where $R_{\min}$ is the distance of closest approach. Note that when $\epsilon = 0$ we recover the fact that only the ratio e/ℓ is measurable for a photon. Substituting the above expression back into Eq. (B6), and dividing the result by Eq. (B5):

$$\frac{1}{r^4}\left(\frac{dr}{d\phi}\right)^2 = \frac{1}{\ell^2}\left[\left(1 - \frac{2M}{R_{\min}}\right)\left(\frac{\ell^2}{R^2} + \epsilon\right)\right.$$
$$\left. - \left(1 - \frac{2M}{r}\right)\left(\frac{\ell^2}{r^2} + \epsilon\right)\right]. \qquad (B8)$$

Changing variables to $u = R_{\min}/r$, inverting and rearranging gives

$$\frac{d\phi}{du} = \left[\left(1 - \frac{2M}{R_{\min}}\right)\left(1 + \epsilon\frac{R_{\min}^2}{\ell^2}\right)\right.$$
$$\left. - \left(1 - \frac{2Mu}{R_{\min}}\right)\left(u^2 + \epsilon\frac{R_{\min}^2}{\ell^2}\right)\right]^{-\frac{1}{2}} \qquad (B9)$$

$$= \left[1 - u^2 - \frac{2M}{R}\left(1 + \epsilon\frac{R_{\min}^2}{\ell^2}(1 - u) - u^3\right)\right]^{-\frac{1}{2}} \qquad (B10)$$

$$= (1 - u^2)^{-\frac{1}{2}}\left[1 - \frac{2M}{R_{\min}}(1 - u^2)^{-1}\right.$$
$$\left. \times \left(1 + \epsilon\frac{R_{\min}^2}{\ell^2}(1 - u) - u^3\right)\right]^{-\frac{1}{2}}. \qquad (B11)$$

Another change of variables, $u = \cos\alpha$, then yields

$$d\phi = -d\alpha\left[1 - \frac{2M}{R_{\min}}\frac{1}{(1 - \cos^2\alpha)}\right.$$
$$\left. \times \left(1 + \epsilon\frac{R_{\min}^2}{\ell^2}(1 - \cos\alpha) - \cos^3\alpha\right)\right]^{-\frac{1}{2}}. \qquad (B12)$$

Using the identity

$$\frac{1-\cos^3\alpha}{1-\cos^2\alpha} \equiv \frac{(1-\cos\alpha)(1+\cos\alpha+\cos^2\alpha)}{(1-\cos\alpha)(1+\cos\alpha)} \quad (B13)$$

$$\equiv \frac{1}{1+\cos\alpha} + \cos\alpha \quad (B14)$$

to simplify the integrand, Eq. (B12) then becomes:

$$d\phi = -d\alpha\left[1 - \frac{2M}{R_{\min}}\left(\cos\alpha + \frac{(1+\epsilon\frac{R_{\min}^2}{\ell^2})}{1+\cos\alpha}\right)\right]^{-\frac{1}{2}}. \quad (B15)$$

For all the situations discussed in this paper, the lensed particles remain far from the Schwarzschild radius of the lens. Therefore we can perform a Taylor expansion in the small parameter $2M/R_{\min} \equiv r_S/R_{\min} \ll 1$ (where r_S is the Schwarzschild radius), obtaining

$$d\phi \approx -d\alpha\left[1 + \frac{M}{R_{\min}}\left(\cos\alpha + \frac{(1+\epsilon\frac{R_{\min}^2}{\ell^2})}{1+\cos\alpha}\right)\right]$$
$$+ \mathcal{O}\left(\frac{M^2}{R_{\min}^2}\right). \quad (B16)$$

In fact, in the case of the massive particle, an additional assumption is needed for the Taylor expansion performed above to remain valid: that the ratio $R_{\min}/\ell$ does not grow very large. In flat space ℓ has the interpretation of the angular momentum per unit rest mass. Hence can roughly estimate $\ell \sim dv$, where d is the impact parameter between the particle and the lens and v is the initial 3-velocity of the particle. Temporarily reinstating factors of c for dimensional clarity, we require:

$$\frac{R_{\min}^2 c^2}{\ell^2} \sim \frac{R_{\min}^2}{d^2}\frac{c^2}{v^2} \lesssim 1. \quad (B17)$$

For small deflection events, the impact parameter b and distance of closest approach $R_{\min}$ are comparable in magnitude. So our condition for the Taylor expansion to be valid then reduces to $c^2/v^2 \lesssim 1$. Since $c/v < 1$ is forbidden, we will have to enforce the condition $v \sim c$. Note this somewhat unusual situation—the results that follow here are *only* valid for particles that are at least moderately relativistic.

Integrating from $\alpha = \pi/2$ to $\alpha = 0$ corresponds to moving along the particle trajectory from infinity to its closest approach to the central mass. By the symmetry of the particle's approach and retreat, the deflection angle α will then be $2\phi_* - \pi$ (see Fig. 9)
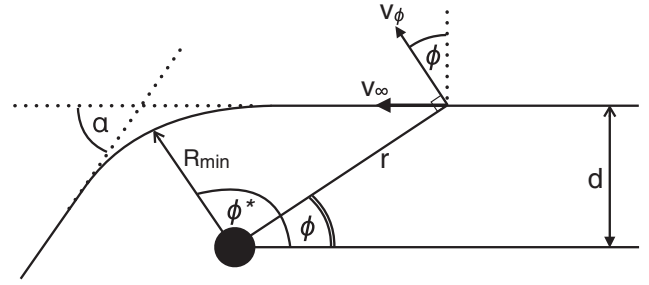

FIG. 9. Diagram indicating the quantities needed for the calculation of Appendix B.

$$\phi_* \approx \int_0^{\frac{\pi}{2}} d\alpha\left[1 + \frac{M}{R}\left(\cos\alpha + \frac{(1+\epsilon\frac{R^2}{\ell^2})}{1+\cos\alpha}\right)\right]$$
$$+ \mathcal{O}\left(\frac{M^2}{R^2}\right) \quad (B18)$$

$$= \left[\alpha + \frac{M}{R}\sin\alpha + \frac{M}{R}\left(1+\epsilon\frac{R^2}{\ell^2}\right)\tan\left(\frac{\alpha}{2}\right)\right]_0^{\frac{\pi}{2}} \quad (B19)$$

$$= \frac{\pi}{2} + \frac{M}{R}\left(2 + \epsilon\frac{R^2}{\ell^2}\right), \quad (B20)$$

$$\Rightarrow \alpha = 4\frac{M}{R}\left(1 + \frac{\epsilon}{2}\frac{R^2}{\ell^2}\right) + \mathcal{O}\left(\frac{M^2}{R^2}\right). \quad (B21)$$

Recall the definition of ℓ [Eq. (B5)] is in terms of the affine parameter λ, which is equivalent to the proper time τ for a massive particle. Since ℓ is a constant, we can choose to evaluate it anywhere along the particle trajectory. For convenience, we choose to do this at a location that is (a) sufficiently far from the lens that we can neglect the potential well Φ at leading order, but (b) close enough not to be separated from the lens by a cosmological distance (i.e. $\sim$Gpc). We approximate this intermediate regime as Minkowski space, and use it to link the particle motion in the large-scale FRW space to its local motion near the lens.

In this Minkowski patch, the particle proper time and the time measured by an observer at rest with respect to the lens are related by $d\tau = dt/\gamma(v)$, where v is the particle velocity (constant in the patch). Then we have (with reference to Fig. 9):

$$\ell = r^2\frac{d\phi}{d\tau} = \gamma r^2\frac{d\phi}{dt} = \gamma r v_\phi = \gamma r v\sin\phi \quad (B22)$$

$$= \gamma r v\left(\frac{d}{r}\right) = \gamma v d, \quad (B23)$$

where v_ϕ is the velocity component in the azimuthal direction. Now using this in Eq. (B21):

$$\alpha = \frac{4GM}{R_{\min}c^2}\left[1 + \frac{\epsilon}{2}\frac{R_{\min}^2}{d^2}\left(\frac{c^2}{v^2} - 1\right)\right], \quad (B24)$$

where we have reinstated factors of G and c. One can show fairly easily (we do not do so here for brevity) that the difference between $R_{\min}$ and d is a number of order $r_S/R_{\min}$, and hence, to the order at which we are working, we can rewrite the above as:

$$\alpha = \frac{4GM}{dc^2}\left[1 + \frac{1}{2}\left(\frac{c^2}{v^2} - 1\right)\right] + \mathcal{O}\left(\frac{r_S^2}{R_{\min}^2}\right). \qquad \text{(B25)}$$

Note that we recover the standard result for a massless particle in the limit $v \to c$ (so we do not need the ϵ parameter any more). We remind the reader that we specialized to relativistic particles in Eq. (B17), so this expression is not valid in the limit $v \to 0$.

The analogous derivation for an extended lensing mass (which may not be entirely interior to the particle trajectory) follows by straightforward integration over a distribution of point masses. The resulting deflection (scaled) angle is

$$\vec{\alpha}(\vec{\theta}) = \frac{1}{\pi}\int d^2\theta'\kappa(\vec{\theta}')\left[1 + \frac{1}{2}\left(\frac{c^2}{v^2} - 1\right)\right]\left(\frac{\vec{\theta} - \vec{\theta}'}{|\vec{\theta} - \vec{\theta}'|^2}\right),$$
$$\text{(B26)}$$

where $\vec{\theta}$ is the angular position in the lens plane, the integral is taken over the entire lens, and $\kappa(\vec{\theta})$ is the dimensionless surface mass density, to be defined in the next appendix.

From Eqs. (B25) and (B26), we can see that the standard formulas for deflection of photons incur a small correction sensitive to the velocity (equivalently, the mass and momentum) of a massive particle. One therefore might expect null and non-null effective images of the same source to be slightly misaligned in the sky. However, as explained in Sec. II, the effect of this misalignment on the differential massive time delay can be neglected to the accuracy used throughout this paper.

APPENDIX C: LENSING BY A MASS DISTRIBUTION

Some readers of this paper may be unfamiliar with the formalism of strong gravitational lensing; here we provide a brief summary of some of the standard definitions. Further details and excellent pedagogical introductions may be found in [34,51,53].

The expressions here make use of the thin-lens approximation. To simplify their formulation, we will assume axial symmetry about the optical axis (the line connecting the observer to the center of the lens). More general, vectorial versions can be found in the references above.

As discussed in Sec. II A and derived in Appendix B, in general the deflection angle experienced by a massive particle is slightly different to that experienced by a massless particle. However, this correction only becomes

relevant at order $(mc/p_0)^4$, and so is not needed for the present work. Hence all expressions in this appendix relate to lensing of massless particles.

We start with a lens model with three-dimensional density profile $\rho(\vec{r})$. Under the thin-lens approximation we project this onto a surface at $z = z_L$ [84]. The projected surface mass density is

$$\Sigma(\vec{\xi}) = \int dr_3\rho(\vec{r}), \qquad \text{(C1)}$$

where $\vec{r} = \{\vec{\xi}, r_3\}$ is a 3D position vector centred on the lens that can be decomposed into a component r_3 along the optical axis, and a 2D position vector in the lensing plane, $\vec{\xi}$ (see Fig. 1).

One can easily show that the vectorial deflection resulting from a 2D distribution of mass elements is [34]

$$\vec{\alpha}(\vec{\xi}) = \frac{4G}{c^2}\frac{\vec{\xi}}{\xi^2}\left[2\pi\int_0^\xi d\xi'\Sigma(\xi')\xi'\right], \qquad \text{(C2)}$$

where, for example, ξ is the magnitude of $\vec{\xi}$. The square bracket gives the mass contained within a radius ξ in the lensing plane. Equation (C2) is loosely comparable to the standard formula for deflection by a point mass with impact parameter b; $\alpha = 4GM/bc^2$. The prefactor of $\vec{\xi}/\xi^2$ is analogous to the factor $1/b$, but also indicates that the deflection is toward the center of the lens.

The quantity appearing in the lensing equation (15) is in fact the "scaled deflection angle," $(D_{LS}/D_S)\alpha$. To obtain this, we take the magnitude of the equation above and replace the 2D position vectors by angular positions using $\xi = D_L\theta$:

$$\alpha_{\text{scal}}(\theta) = \frac{D_{LS}}{D_S}\alpha(\theta) \qquad \text{(C3)}$$

$$= \frac{D_{LS}D_L}{D_S}\frac{4\pi G}{c^2}\frac{1}{\theta}\left[2\int_0^\theta d\theta'\Sigma(\theta')\theta'\right] \qquad \text{(C4)}$$

$$= \frac{1}{\theta}\left[2\int_0^\theta d\theta'\kappa(\theta')\theta'\right], \qquad \text{(C5)}$$

where the convergence $\kappa(\theta)$ and critical density Σ_{cr} are defined as

$$\kappa(\theta) = \frac{\Sigma(\theta)}{\Sigma_{\text{cr}}}$$

$$\Sigma_{\text{cr}} = \frac{c^2}{4\pi G}\frac{D_S}{D_L D_{LS}}. \qquad \text{(C6)}$$

One final simplification is helpful. The second square bracket in Eq. (C5) is, up to a factor of π, the dimensionless mass of the lens contained within angular radius θ

(the dimensions having been removed by Σ_{cr} in the denominator of κ). Defining the dimensionless mean surface mass density by

$$\bar{\kappa}(\theta) = \frac{M(<\theta)}{\pi\theta^2}, \qquad (C7)$$

Eq. (C5) then becomes

$$\alpha_{scal}(\theta) = \bar{\kappa}(\theta)\theta. \qquad (C8)$$

Finally, recall that the scalar form of the lens equation is

$$\alpha_{scal}(\theta) = \theta \pm \beta. \qquad (C9)$$

In this way the factors of $\theta - \beta$ (and similar) that appear in our calculations can be calculated from $\rho(r)$. Note also that the projected 2D potential $\psi(\theta)$ can be related to the density profile via $\vec{\alpha}_{scal}(\theta) = \tilde{\nabla}_\theta \psi(\theta)$, where $\tilde{\nabla}_\theta$ is a derivative in the lensing plane.

[1] B. P. Abbott *et al.*, Phys. Rev. Lett. **116**, 061102 (2016).
[2] B. P. Abbott *et al.*, Phys. Rev. Lett. **116**, 241103 (2016).
[3] L. Stodolsky, Phys. Lett. B **473**, 61 (2000).
[4] L. Stodolsky, in *Neutrino Physics—Its Impact on Particle Physics, Astrophysics and Cosmology*, edited by J. Bahcall, W. Haxton, K. Kubodera, and C. Poole (World Scientific, Singapore, 2001), p. 14.
[5] G. I. Zatsepin, J. Exp. Theor. Phys. Lett. **8**, 205 (1968).
[6] K. Hirata *et al.*, Phys. Rev. Lett. **58**, 1490 (1987).
[7] G. Fanizza, M. Gasperini, G. Marozzi, and G. Veneziano, Phys. Lett. B **757**, 505 (2016).
[8] X. Li, Y.-M. Hu, Y.-Z. Fan, and D.-M. Wei, Astrophys. J. **827**, 75 (2016).
[9] K. S. Thorne, in *Three Hundred Years of Gravitation*, edited by S. W. Hawking and W. Israel (Cambridge University Press, Cambridge, 1987), p. 330.
[10] A. A. Ruffa, Astrophys. J. **517**, L31 (1999).
[11] Y. Wang, A. Stebbins, and E. L. Turner, Phys. Rev. Lett. **77**, 2875 (1996).
[12] M. Sereno, P. Jetzer, A. Sesana, and M. Volonteri, Mon. Not. R. Astron. Soc. **415**, 2773 (2011).
[13] M. Sereno, A. Sesana, A. Bleuler, P. Jetzer, M. Volonteri, and M. C. Begelman, Phys. Rev. Lett. **105**, 251101 (2010).
[14] A. Piórkowska, M. Biesiada, and Z.-H. Zhu, J. Cosmol. Astropart. Phys. 10 (2013) 022.
[15] S. Nissanke, M. Kasliwal, and A. Georgieva, Astrophys. J. **767**, 124 (2013).
[16] J. Aasi, J. Abadie, B. P. Abbott, R. Abbott, T. Abbott, M. R. Abernathy, T. Accadia, F. Acernese, C. Adams, T. Adams *et al.*, Astrophys. J. Suppl. Ser. **211**, 7 (2014).
[17] Y. Sekiguchi, K. Kiuchi, K. Kyutoku, and M. Shibata, Phys. Rev. Lett. **107**, 051102 (2011).
[18] F. Foucart, E. O'Connor, L. Roberts, L. E. Kidder, H. P. Pfeiffer, and M. A. Scheel, Phys. Rev. D **94**, 123016 (2016).
[19] O. L. Caballero, G. C. McLaughlin, and R. Surman, Phys. Rev. D **80**, 123004 (2009).
[20] O. L. Caballero, T. Zielinski, G. C. McLaughlin, and R. Surman, Phys. Rev. D **93**, 123015 (2016).
[21] K. N. Yakunin, P. Marronetti, A. Mezzacappa, S. W. Bruenn, C.-T. Lee, M. A. Chertkow, W. R. Hix, J. M. Blondin, E. J. Lentz, O. E. B. Messer, and S. Yoshida, Classical Quantum Gravity **27**, 194005 (2010).
[22] C. D. Ott, E. P. O'Connor, S. Gossan, E. Abdikamalov, U. C. T. Gamma, and S. Drasco, Nucl. Phys. B, Proc. Suppl. **235–236**, 381 (2013).
[23] H. Andresen, B. Mueller, E. Mueller, and H.-T. Janka, arXiv:1607.05199.
[24] R. M. Quimby, M. C. Werner, M. Oguri, S. More, A. More, M. Tanaka, K. Nomoto, T. J. Moriya, G. Folatelli, K. Maeda, and M. Bersten, Astrophys. J. Lett. **768**, L20 (2013).
[25] R. M. Quimby, M. Oguri, A. More, S. More, T. J. Moriya, M. C. Werner, M. Tanaka, G. Folatelli, M. C. Bersten, K. Maeda, and K. Nomoto, Science **344**, 396 (2014).
[26] P. L. Kelly *et al.*, Science **347**, 1123 (2015).
[27] The Newtonian intuition that a massive particle should experience greater deflection that a massless one—because its slower speed implies less momentum to "escape"—turns out to be correct.
[28] We will see in Sec. II B that we are effectively doing a doing a double Taylor expansion here, in the small parameters $\delta\theta$ and m/p_0.
[29] This may seem counterintuitive initially. If the source is at a cosmological redshift, surely the expansion of the universe has a significant affect on its travel time? How, then, can the conformal time delay be equivalent to the physical one? The nuance here is that cosmological expansion affects all lensed paths in the same way, so does not contribute to their relative difference at $z = 0$. That is, $t_{total}(\theta) \neq \eta_{total}(\theta)$, but $\Delta t(\theta_2, \theta_1) \simeq \Delta\eta(\theta_2, \theta_1)$ to a high degree of accuracy.
[30] U. Seljak, Astrophys. J. **436**, 509 (1994).
[31] Note that a more correct way to describe this situation would be to use the McVittie metric instead. However, in the regime we are considering—far outside the gravitational radius of the lens—the McVittie metric reduces to a form equivalent to Eq. (II B).
[32] This Euclidean notion of distance is justified since we are working in the thin-lens approximation and on a conformal grid.
[33] D. Baumann, Cosmology lecture notes, Cambridge mathematical tripos, http://www.damtp.cam.ac.uk/user/db275/Cosmology/Lectures.pdf.
[34] P. Schneider, arXiv:astro-ph/0509252.
[35] R. Blandford and R. Narayan, Astrophys. J. **310**, 568 (1986).

[36] S. K. Bose and W. D. McGlinn, Phys. Rev. D **38**, 2335 (1988).

[37] W. L. Burke, Astrophys. J. **244**, L1 (1981).

[38] W. K. Rose, *Advanced Stellar Astrophysics* (Cambridge University Press, Cambridge, England, 1998), p. 351, ISBN 0521581885.

[39] L. V. E. Koopmans, A. Bolton, T. Treu, O. Czoske, M. W. Auger, M. Barnabè, S. Vegetti, R. Gavazzi, L. A. Moustakas, and S. Burles, Astrophys. J. Lett. **703**, L51 (2009).

[40] P. A. R. Ade, N. Aghanim, M. Arnaud, M. Ashdown, J. Aumont, C. Baccigalupi, A. J. Banday, R. B. Barreiro, J. G. Bartlett *et al.* (Planck Collaboration), Astron. Astrophys. **594**, A13 (2016).

[41] Expressed another way, neutrino oscillations mean that a delta-function emission burst will be received with some arrival time distribution. The same is true for the lensed echoes of the original signal. Our calculation essentially relies on accurately differencing these distributions; their width is unimportant.

[42] T. Treu and P. J. Marshall, Astron. Astrophys. Rev. **24**, 11 (2016).

[43] V. Bonvin, F. Courbin, S. H. Suyu, P. J. Marshall, C. E. Rusu, D. Sluse, M. Tewes, K. C. Wong, T. Collett, C. D. Fassnacht, T. Treu, M. W. Auger, S. Hilbert, L. V. E. Koopmans, G. Meylan, N. Rumbaugh, A. Sonnenfeld, and C. Spiniello, arXiv:1607.01790.

[44] J. Klein and S. Thorsett, Phys. Lett. A **145**, 79 (1990).

[45] I. H. Stairs, Living Rev. Relativ. **6** 5 (2003).

[46] P. G. Tinyakov and I. I. Tkachev, J. Exp. Theor. Phys. **106**, 481 (2008).

[47] D. J. E. Marsh, Phys. Rep. **643**, 1 (2016).

[48] G. G. Raffelt, in *Axions, Lecture Notes in Physics*, edited by M. Kuster, G. Raffelt, and B. Beltrán (Springer Verlag, Berlin, 2008), Vol. 741, p. 51.

[49] X.-L. Fan, K. Liao, M. Biesiada, A. Piorkowska-Kurpas, and Z.-H. Zhu, Phys. Rev. Lett. **118**, 091102 (2017).

[50] R. Barkana, Astrophys. J. **502**, 531 (1998).

[51] P. Schneider, *Saas-Fee Advanced Course 33: Gravitational Lensing: Strong, Weak and Micro*, edited by G. Meylan, P. Jetzer, P. North, P. Schneider, C. S. Kochanek, and J. Wambsganss (Springer-Verlag, Berlin, 2006).

[52] R. Takahashi, Astrophys. J. **835**, 103 (2017).

[53] P. Schneider, J. Ehlers, and E. E. Falco, *Gravitational Lenses, XIV* (Springer-Verlag, Berlin, 1992), p. 112.

[54] A. Goobar *et al.*, arXiv:1611.00014.

[55] L. Dai, T. Venumadhav, and K. Sigurdson, Phys. Rev. D **95**, 044011 (2017).

[56] C. Messenger, K. Takami, S. Gossan, L. Rezzolla, and B. S. Sathyaprakash, Phys. Rev. X **4**, 041004 (2014).

[57] M. Kilbinger, Rep. Prog. Phys. **78**, 086901 (2015).

[58] https://www.lsst.org.

[59] P. A. Abell, J. Allison, S. F. Anderson, J. R. Andrew, J. R. P. Angel, L. Armus, D. Arnett, S. J. Asztalos, T. S. Axelrod *et al.* (LSST Science Collaboration), arXiv:0912.0201.

[60] http://www.virgo-gw.eu.

[61] http://gwcenter.icrr.u-tokyo.ac.jp/en/.

[62] C. Moore (private communication).

[63] J. F. Beacom and P. Vogel, Phys. Rev. D **60**, 033007 (1999).

[64] A. More, S. H. Suyu, M. Oguri, S. More, and C.-H. Lee, Astrophys. J. **835**, L25 (2017).

[65] M. J. Longo, Phys. Rev. Lett. **60**, 173 (1988).

[66] E. O. Kahya and S. Desai, Phys. Lett. B **756**, 265 (2016).

[67] J. A. Formaggio and C. Martoff, Annu. Rev. Nucl. Part. Sci. **54**, 361 (2004).

[68] K. Abe, K. Haga, Y. Hayato, and the Super-Kamiokande Collaboration, Astrophys. J. Lett. **830**, L11 (2016).

[69] K. Scholberg, Annu. Rev. Nucl. Part. Sci. **62**, 81 (2012).

[70] F. Halzen and S. R. Klein, Rev. Sci. Instrum. **81**, 081101 (2010).

[71] M. Ageron, J. A. Aguilar, I. Al Samarai, A. Albert, F. Ameli, M. André, M. Anghinolfi, G. Anton, S. Anvar, M. Ardid *et al.*, Nucl. Instrum. Methods Phys. Res., Sect. A **656**, 11 (2011).

[72] E. Thrane (Super-Kamiokande Collaboration), Astrophys. J. **704**, 503 (2009).

[73] K. Asakura (KamLAND Collaboration), Astrophys. J. **818**, 91 (2016).

[74] SNO+ Collaboration, Adv. High Energy Phys. **2016**, 6194250 (2016).

[75] S. Fukuda *et al.*, Astrophys. J. **578**, 317 (2002).

[76] W. C. Jones *et al.*, Astrophys. J. **647**, 823 (2006).

[77] R. Stompor, S. Hanany, M. E. Abroe, J. Borrill, P. G. Ferreira, A. H. Jaffe, B. Johnson, A. T. Lee, B. Rabii, P. L. Richards, G. Smoot, C. Winant, and J. H. P. Wu, arXiv:astro-ph/0309409.

[78] C. L. Bennett *et al.*, Astrophys. J. Suppl. Ser. **208**, 20 (2013).

[79] P. Fleury, Phys. Lett. B **760**, 350 (2016).

[80] S. Fairhurst, New J. Phys. **11**, 123006 (2009).

[81] Y. Aso, E. Goetz, P. Kalmus, L. Matone, S. Márka, J. Myers, B. O'Reilly, R. Savage, P. Schwinberg, X. Siemens, D. Sigg, and N. Smith, Classical Quantum Gravity **26**, 055010 (2009).

[82] https://www.elisascience.org.

[83] J. B. Hartle, *Gravity* (Addison Wesley, San Francisco, CA, USA, 2003), p. 582, ISBN 0-8053-8662-9.

[84] Note that for a sufficiently large lens, the flat-sky approximation does not hold. A surface of fixed redshift is then curved instead of a plane.