

# Discovering gravitationally lensed gravitational waves: predicted rates, candidate selection, and localization with the Vera Rubin Observatory

Graham P. Smith <sup>1,★</sup>, Andrew Robertson <sup>2</sup>, Guillaume Mahler <sup>3,4</sup>, Matt Nicholl <sup>1,5</sup>,  
Dan Ryczanowski <sup>1</sup>, Matteo Bianconi<sup>1</sup>, Keren Sharon <sup>6</sup>, Richard Massey <sup>4</sup>, Johan Richard <sup>7</sup>  
and Mathilde Jauzac <sup>3,4,8,9</sup>

<sup>1</sup>*School of Physics and Astronomy, University of Birmingham, Edgbaston B15 2TT, UK*

<sup>2</sup>*Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Dr., Pasadena, CA 91109, USA*

<sup>3</sup>*Centre for Extragalactic Astronomy, Department of Physics, Durham University, South Road, Durham DH1 3LE, UK*

<sup>4</sup>*Institute for Computational Cosmology, Department of Physics, Durham University, South Road, Durham DH1 3LE, UK*

<sup>5</sup>*Institute for Gravitational Wave Astronomy, University of Birmingham, Birmingham B15 2TT, UK*

<sup>6</sup>*Department of Astronomy, University of Michigan, 1085 S. University Ave, Ann Arbor, MI 48109, USA*

<sup>7</sup>*Univ Lyon, Univ Lyon1, Ens de Lyon, CNRS, Centre de Recherche Astrophysique de Lyon UMR5574, F-69230 Saint-Genis-Laval, France*

<sup>8</sup>*Astrophysics Research Centre, University of KwaZulu-Natal, Westville Campus, Durban 4041, South Africa*

<sup>9</sup>*School of Mathematics, Statistics and Computer Science, University of KwaZulu-Natal, Westville Campus, Durban 4041, South Africa*

Accepted 2023 January 9. Received 2022 December 8; in original form 2022 April 28

## ABSTRACT

Secure confirmation that a gravitational wave (GW) has been gravitationally lensed would bring together these two pillars of General Relativity for the first time. This breakthrough is challenging for many reasons, including: GW sky localization uncertainties dwarf the angular scale of gravitational lensing, the mass and structure of gravitational lenses is diverse, the mass function of stellar remnant compact objects is not yet well constrained, and GW detectors do not operate continuously. We introduce a new approach that is agnostic to the mass and structure of the lenses, compare the efficiency of different methods for lensed GW discovery, and explore detection of lensed kilonova counterparts as a direct method for localizing candidates. Our main conclusions are: (1) lensed neutron star mergers (NS–NS) are magnified into the ‘mass gap’ between NS and black holes, therefore selecting candidates from public GW alerts with high mass gap probability is efficient, (2) the rate of detectable lensed NS–NS will approach one per year in the mid-2020s, (3) the arrival time difference between lensed NS–NS images is  $1\text{ s} \lesssim \Delta t \lesssim 1\text{ yr}$ , and thus well-matched to the operations of GW detectors and optical telescopes, (4) lensed kilonova counterparts are faint at peak (e.g.  $r_{\text{AB}} \simeq 24\text{--}26$  in the mid-2020s), fade quickly ( $d < 2\text{ d}$ ), and are detectable with target of opportunity observations with large wide-field telescopes. For example, just  $\lesssim 0.25$  per cent of Vera C. Rubin Observatory’s observing time will be sufficient to follow up one well-localized candidate per year. Our predictions also provide a physically well-defined basis for exploring electromagnetically the exciting new ‘mass gap’ discovery space.

**Key words:** gravitational lensing; strong – gravitational waves.

## 1 INTRODUCTION

Gravitational waves (GWs) are magnified, deflected, and delayed as they traverse gravitational fields, just like the more familiar gravitational lensing of light from distant galaxies and quasars (Ohanian 1974; Wang, Stebbins & Turner 1996; Nakamura 1998), pointing to exciting opportunities to detect multiple gravitationally magnified images of a distant GW source. Detection of gravitationally lensed GWs will be the first opportunity to combine these two pillars of Einstein’s General Relativity in a single experiment, leading to novel tests of General Relativity and general metric gravity theories (e.g. Mukherjee, Wandelt & Silk 2020a, b; Finke et al. 2021; Goyal et al. 2021), measurements of the Hubble constant and tests of the

Friedmann–LeMaître–Robertson–Walker metric (e.g. Sereno et al. 2011; Liao et al. 2017; Cao et al. 2019; Li, Fan & Gou 2019; Hannuksela et al. 2020; Hou et al. 2020), testing the speed of GW propagation (e.g. Baker & Trodden 2017; Collett & Bacon 2017; Fan et al. 2017), and constraints on intermediate mass black holes (BH) and primordial BHs via their microlensing signal (Lai et al. 2018; Diego 2020; Oguri & Takahashi 2020; Wang, Herrera-Martín & Hu 2021; Urrutia & Vaskonen 2022). These studies set a broad expectation that gravitationally lensed GWs is the science of the late 2030s and 2040s when the third generation of GW detectors will begin operations, generally focus on lensed binary black hole (BH–BH) mergers over lensed binary neutron star (NS–NS) mergers, and tend to assume that the only relevant population of lenses is massive early-type galaxies. We discuss these three themes in turn below.

In the more familiar realm of gravitational lensing of electromagnetic (EM) radiation, early detections of gravitationally lensed high

\* E-mail: [gps@star.sr.bham.ac.uk](mailto:gps@star.sr.bham.ac.uk)

redshift galaxies, some of which were the redshift record breakers of their day (e.g. Soucail et al. 1987; Lynds & Petrosian 1989; Mellier et al. 1991; Franx et al. 1997; Ellis et al. 2001; Kneib et al. 2004), pre-dated systematic studies of large samples of lensed galaxies by several decades (e.g. Atek et al. 2015; McLeod, McLure & Dunlop 2016; Ishigaki et al. 2018; Bhatawdekar et al. 2019). Current progress towards the first detection and early science of gravitationally lensed GWs therefore resonates with the status of lensed high redshift galaxies in the mid-1980s and detection of lensed quasars a little earlier than that (Walsh, Carswell & Weymann 1979) – i.e. the years leading up to the first detection. Likewise, systematic study of large samples of lensed GWs with next-generation GW instruments in the 2030s and 2040s will resemble the status of lensed galaxy and lensed quasar studies of the last decade.

A growing suite of predictions are consistent with a rate of detection of lensed GWs of around one per year in the coming years with the current generation of GW detectors (Li et al. 2018; Ng et al. 2018; Oguri 2018, 2019; Smith et al. 2019a; Wierda et al. 2021; Xu, Ezquiaga & Holz 2022). Echoing the progress driven by the emergence of charge coupled devices (CCDs) in optical astronomy in the 1980s, several groups have therefore been considering whether any of the gravitationally lensed GWs detected by current instruments are lensed (e.g. Broadhurst, Diego & Smoot 2018; Smith et al. 2018, 2019b; Hannuksela et al. 2019; Singer, Goldstein & Bloom 2019; Abbott et al. 2020; Dai et al. 2020; McIsaac et al. 2020; Diego, Broadhurst & Smoot 2021; Liu, Magaña Hernandez & Creighton 2021; The LIGO Scientific Collaboration 2021; Bianconi et al. 2022). In turn, this is motivating development of techniques to identify candidate lensed GWs (e.g. Li et al. 2018; Ng et al. 2018; Oguri 2018; Smith et al. 2019a; Hannuksela et al. 2020; Robertson et al. 2020; Ryczanowski et al. 2020, 2022; Wierda et al. 2021; Wempe et al. 2022; Xu et al. 2022; Yang et al. 2022), and joint investigations of lensed GWs and the stochastic GW background (Buscicchio et al. 2020a, b; Mukherjee & Silk 2021).

It is inarguable that multimessenger detection of a lensed GW will be transformational, i.e. detection of a lensed NS–NS merger and its lensed EM counterpart. This is because the EM detection will localize the GW source to  $\lesssim 1$  arcsec accuracy with ground-based data, and potentially  $\simeq 10$  milliarcsec accuracy with space-based observations. This will be the lensing analogue of the localization of GW170817 to a specific location in the outskirts of its host galaxy NGC4993 via the detection of its optical/infrared and short gamma-ray burst (SGRB) counterpart (Abbott et al. 2017b; Margutti & Chornock 2021, and references therein). Such accurate localization would eliminate reliance on statistical arguments for or against several poorly localized BH–BH detections being images of the same lensed source and several different galaxies being the lens and/or the host (e.g. Hannuksela et al. 2020; Wempe et al. 2022). Moreover, despite intriguing candidates (e.g. Graham et al. 2020), there are no confirmed EM counterparts to BH–BH mergers available to help with localization. Whilst the detection rates of GW signals from lensed NS–NS mergers are expected to be lower than for lensed BH–BH mergers, preliminary estimates indicate that multimessenger detection of lensed NS–NS mergers will be feasible in the 2020s (Oguri 2019; Smith et al. 2019a).

Detection of multiple images of the lensed EM counterpart will also expand the science of lensed GWs, for example into the chemical enrichment history of the universe and the physics of kilonovae. For example, detection of the trailing image of the lensed kilonova will potentially probe the rising portion of the kilonova in the rest-frame ultraviolet (UV). This would shed light on arguably the biggest unsolved mystery of GW170817, namely the physics responsible for

the very bright blue early component of its emission (e.g. Arcavi 2018; Piro & Kollmeier 2018; Nicholl et al. 2021). The feasibility of such measurements will depend on the arrival time difference between the images and the reliability of any predictions that are required, as highlighted by resolved and confirmed detections of strongly lensed supernovae (Kelly et al. 2015; Jauzac et al. 2016; Treu et al. 2016; Goobar et al. 2017, 2022; More et al. 2017).

The optical depth to strong lensing (defined here as large lens magnification,  $\mu > 10$ ) spans galaxies, groups, and clusters of galaxies (Hilbert et al. 2008), with the latest calibration being consistent with lenses in each decade of mass spanning  $10^{12} \leq M_{200} \leq 10^{15} M_{\odot}$  contributing approximately equally to the optical depth (Robertson et al. 2020). The scaling of lensing probability with magnification in the strong lensing regime is also well defined analytically as  $dp/d\mu \propto \mu^{-3}$  and – importantly – independent of the lens mass (Blandford & Narayan 1986). Consequently, Hilbert et al. and Robertson et al.’s results persist at all magnifications of  $\mu > 2$ .

The relationship between lens magnification and image multiplicity is less well defined, as it depends on the relationship between the mass and the density profile slope of lenses. The basic idea was sketched by Turner, Ostriker & Gott (1984) and is still valid today (e.g. Smith et al. 2001; Gavazzi et al. 2003; Sand et al. 2004; Smith et al. 2009; Richard et al. 2010; Umetsu et al. 2016; Fox et al. 2022). Briefly, galaxy clusters tend to have shallower density profiles than individual galaxies and therefore the former are less efficient than the latter at forming multiple images of background objects at low magnification, typically  $2 < \mu \lesssim 10$ . However, clusters still contribute their ‘fair share’ of the optical depth at low magnification, they just do so by forming a single image. This is important in the context of predicting rates of and searching for gravitationally lensed GWs. In particular, a currently un-constrained and potentially significant fraction of the optical depth to lens magnifications of  $2 < \mu \lesssim 10$  does not contribute to multiple image formation. Conversely, as highlighted by resolved confirmed strongly lensed supernova discoveries (Kelly et al. 2015; Goobar et al. 2017, 2022), searching for high magnification strong lensing requires consideration of lenses spanning galaxies, groups, and clusters and not just a focus on clusters.

Decisive progress towards robust and unchallenged detection of a gravitationally lensed GW rests on identifying candidates and localizing them to a lensed host galaxy. Searches for candidates have so far identified either pairs of GW detections with sky localizations and waveforms that are consistent with being lensed images of the same source, and/or individual GW sources that appear to have an anomalous mass relative to an assumed underlying mass function (Smith et al. 2018, 2019b; Hannuksela et al. 2019; Singer et al. 2019; Abbott et al. 2020; Dai et al. 2020; McIsaac et al. 2020; Diego et al. 2021; Liu et al. 2021; The LIGO Scientific Collaboration 2021; Bianconi et al. 2022). The former approach relies on the GW detectors collecting data of the required quality when  $\geq 2$  of the lensed images arrived at Earth, and the latter approach is sensitive to the assumed mass function. For example, Broadhurst et al. (2018), Broadhurst, Diego & Smoot (2019), Broadhurst, Diego & Smoot (2022) argue that if the true stellar remnant BH mass function in the universe peaks at  $M_{\text{BH}} \simeq 10 M_{\odot}$  (i.e. well-matched to the BHs thus far detected in the Milky Way), then the GW detections apparently from systems comprising BH in the range  $M_{\text{BH}} \gtrsim 20 M_{\odot}$  are dominated by lensed events.

Efficient selection of candidate lensed GWs therefore faces several important challenges. First, how to define anomalous when the underlying mass function of sources is not known a priori? Secondly, how efficient is searching for multiple detections in a rapidly growing

catalogue of GW detections from the LIGO, Virgo, and KAGRA (LVK) detectors that are dominated by sources with 90 per cent credible sky localization uncertainties of  $\Omega_{90} > 100 \text{ degree}^2$  (Petrov et al. 2022)? Moreover, the GW detectors rarely operate uninterrupted for more than a few hours and are offline for technology upgrades for time-scales of  $\gtrsim 1 \text{ yr}$ .

In this article, we develop a new magnification-based approach to predicting rates of gravitationally lensed GWs that naturally integrates strong lenses of all masses and thus allows us to calibrate both the rates of lensed GW detections, the typical magnifications that they suffer, and their location in the mass–distance plane. We apply this approach to lensing of both BH–BH and NS–NS mergers and combine our magnification-based predictions with time delay theory to identify the typical arrival time differences for these sources and lenses of different mass and structure. This enables us to combine our predictions in the mass–distance plane with expected arrival time differences to develop efficient strategies for identifying and localizing candidate lensed GWs. This leads us to predicting the light curves of lensed kilonova counterparts of lensed NS–NS mergers and considering their detectability with the Vera C. Rubin Observatory (Rubin).

We review the theory and phenomenology of strong lenses and develop our new magnification-based approach in Section 2, then describe the GW source populations that we consider in Section 3, before applying the former to the latter and making our predictions in Section 4. We combine our lensed GW predictions with Nicholl et al.’s (2021) kilonova light curve models to predict lensed kilonova light curves and assess the feasibility of detection with Rubin Target of Opportunity observations in Section 5, and close with a summary in Section 6. We assume a flat cosmology with  $H_0 = 67.9 \text{ km s}^{-1} \text{ Mpc}^{-1}$ ,  $\Omega_M = 0.3065$  (Ade et al. 2016). All magnitudes are stated in the AB system.

## 2 GRAVITATIONAL LENSING THEORY AND PHENOMENOLOGY

We begin with an overview of the main take away points in Section 2.1. For readers interested in more detail, we outline some key elements of gravitational lensing theory in Section 2.2, before motivating our focus on gravitational magnification in Section 2.3, justifying our model for optical depth in Section 2.4, and explaining how we predict arrival time difference for magnified image pairs in Section 2.5.

### 2.1 Overview

We consider strong gravitational lensing of GW and EM signals from mergers of compact object binaries, where the lenses are galaxies, groups, and clusters of galaxies that are serendipitously located along the line of sight to the merger. We ignore any microlensing by stars or compact objects that may be located close to the critical curves of the ‘macrolens’ – i.e. the galaxy/group/cluster lens. Whilst microlensing can be important, especially in high magnification regions of very dense macrolenses (e.g. Diego 2019; Diego et al. 2019; Meena & Bagla 2020; Mishra et al. 2021), it is beyond our current aim of establishing baseline predictions of, and follow-up strategies for, candidate gravitationally lensed GWs. As such our approach is likely to yield lower limits on the detectable rates of lensed GWs.

The geometrical nature of gravitational lensing and the growth of large-scale structure in the universe result in strong lensing being dominated by lenses at redshifts of  $z_L \simeq 0.2\text{--}0.6$  and sources at redshifts of  $z_S \gtrsim 1$ . The typical comoving distance travelled by a

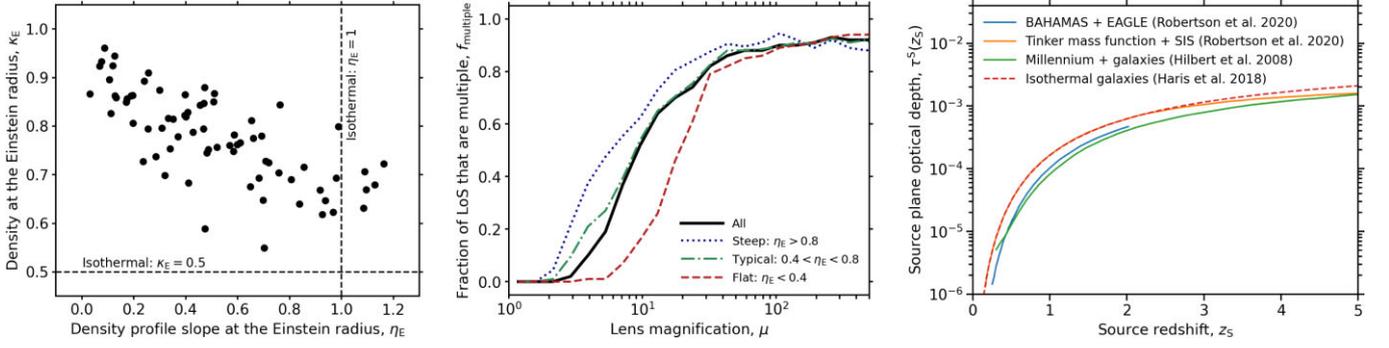
signal from a strongly lensed GW source is therefore expected to be  $\gtrsim 7 \text{ Gpc}$  ( $\gtrsim 10^{26} \text{ m}$ ). This is three orders of magnitude larger than the comoving virial radius of the most massive galaxy clusters in the observable universe (Bianconi et al. 2021). It is therefore valid to adopt the so-called thin lens approximation.

The distance to strongly lensed GW sources also dwarfs their physical size ( $\lesssim 10^3 \text{ m}$ ), allowing us to treat them as point sources of EM and gravitational radiation. This allows us to ignore finite source size effects when considering large lens magnifications. Despite the wavelength of GWs ( $\lambda_{\text{GW}} \lesssim 10^7 \text{ m}$  for LVK) being much longer than EM radiation, the former is still many orders of magnitude smaller than the physical scale of the lenses (Einstein radii corresponding to physical scales of  $\gtrsim 10^{19} \text{ m}$ , i.e.  $\gtrsim 300 \text{ pc}$ , and thus  $\gtrsim 0.1 \text{ arcsec}$  at  $z \simeq 0.2$ ). For the purposes of our calculations, we are therefore safe to treat both GWs and EM radiation in the geometric optics limit and thus ignore wave effects.

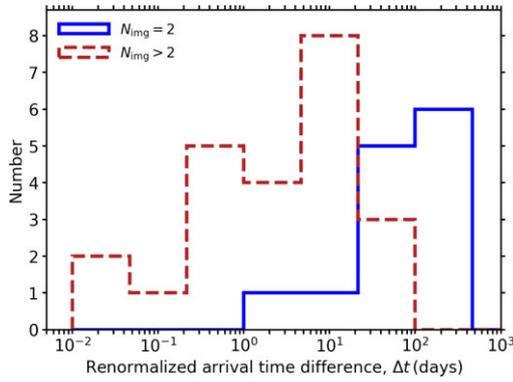
We predict the rates and properties of gravitationally lensed GWs and their EM counterparts as a function of gravitational lens magnification. As discussed in Section 2.3, this is the least model-dependent approach to predicting the ensemble of lensed detections that are feasible with GW detectors and wide-field optical telescopes at the current time. In particular, we side-step the challenge of solving the lens equation (equation 6) and the requirement to assume a form for the deflection potential that this would entail. The main advantage of our approach, beyond computational speed, is that not assuming anything about the structure of the lens population means that our predictions are not biased to any type of lens and are thus relevant across the full range of lens mass ( $10^{12} \lesssim M_{200} \lesssim 10^{15} M_\odot$ ; Robertson et al. 2020). The main disadvantage is that the efficiency of multiple image production at low lens magnification ( $2 < \mu \lesssim 10$ ) depends on lens mass, with individual galaxy-scale lenses (relatively steep density profiles) being more efficient than group- and cluster-scale lenses (relatively shallow density profiles) at low magnification. Care is therefore needed at  $\mu \lesssim 10$ , because some of the predicted magnified detections will be single images.

We adopt a threshold magnification of an individual GW detection of  $\mu > \mu_{\text{th}} = 2$ . This threshold corresponds to a bias by a factor of  $> \sqrt{\mu_{\text{th}}} = 1.4$  in the luminosity distance to a lensed source if the magnification is ignored (equation 12), i.e. comparable with or larger than the typical 90 per cent posterior credible intervals on luminosity distance in LVK’s GW data analysis. This threshold also marks the onset of multiple image formation by a lens – such as an isolated early-type galaxy – that is well-described by an isothermal density profile. It is however important to note that the fainter of two images formed by an isothermal lens is demagnified, i.e.  $0 \leq |\mu_-| \leq 1$ , if the magnification of the brighter image is in the range  $2 \leq \mu_+ \leq 3$  (equation 16). The formation of multiple images therefore does not automatically imply the detectability of multiple images, although detections at  $\mu \gtrsim 10$  will be dominated by multiply-imaged systems (Fig. 1 and Section 2.3), and be accompanied by other similarly magnified images independent of the structure of the lens.

We use analytic expressions to convert the predicted magnifications of gravitationally lensed GWs into predictions of the time since / until another signal from the same source arrived / will arrive at Earth (equations 21 and 22). These expressions are based on two common scenarios for the formation of bright highly magnified image pairs, and are shown to capture the time delay behaviour of the systems that have been discovered to date (Fig. 2). The expressions capture the dependence of time delay on both the mass and the structure of the lens, bearing in mind that group/cluster-scale lenses tend to be denser and flatter than galaxy-scale lenses at their respective Einstein radii (Fig. 1). The details of the time delay and its dependence on lens



**Figure 1.** Left-hand panel: The distribution of density  $\kappa_E$  versus density profile slope  $\eta_E$ , measured from the lens models of 79 group- and cluster-scale lenses. Groups and clusters are denser and shallower at their Einstein radius than the SIS model that is a good description of individual galaxy-scale lenses. Centre: The run of  $f_{\text{multiple}}$  with lens magnification for the lenses shown in the left-hand panel, where  $f_{\text{multiple}}$  is the fraction of lines of sight at a given lens magnification that are multiple, i.e. have image multiplicity of greater than one. The threshold magnification at which lenses are efficient at forming multiple images of distant sources is a function of density profile slope. Right-hand panel: Different predictions of the source plane optical depth as a function of source redshift agree within a factor of  $\lesssim 2$ . The three models shown by solid curves are discussed by Robertson et al., and are based on three different approaches to integrating the optical depth to lens magnification over the full dynamic range of the dark matter halo mass function. The BAHAMAS+EAGLE curve is based on modern high resolution cosmological hydrodynamical simulations, the Tinker curve combines the analytic form of the Tinker et al. (2008) mass function with the cross-section of the singular isothermal density profile model, and the millennium curve is based on the semi-analytic approach of Hilbert et al. (2008) that pastes analytic galaxies into the dark matter-only millennium simulation. The dashed curve, from Haris et al. (2018), is a prediction of the optical depth to multiple imaging that assumes that all gravitational lenses are individual massive galaxies with isothermal density profiles.



**Figure 2.** The distribution of arrival time differences obtained by authors cited in Section 2.5 for known time delay lenses including lensed quasars and Supernova Refsdal, split into those for lenses that produced just two images and those that produced more than two images. Arrival time difference has been renormalized to the same lens geometry of  $\mathcal{D} = 3.3$  Gpc and Einstein radius of  $\theta_E = 1$  arcsec.

properties are described in Section 2.5 and the supporting derivations are presented in Appendix A.

## 2.2 Lensing fundamentals

We consider an infinitely thin mass distribution at redshift,  $z_L$ , and distant sources at redshift  $z_S > z_L$ . The true and unmeasurable positions of the sources on the celestial sphere are  $\beta$  and the observed positions are  $\theta$ . Following precepts laid down by Schneider (1985), Blandford & Narayan (1986), and others, the time for a lensed signal to travel from the source to the observer differs from the traveltime in the absence of lensing by an amount that depends on a dimension-less scalar field called the Fermat potential,  $\tau$ :

$$ct = \mathcal{D} \tau(\theta, \beta), \quad (1)$$

where  $\mathcal{D} = D_L^C D_S^C / D_{LS}^C$ ,  $D_L^C$  and  $D_S^C$  are the comoving distances from the observer to the lens and the source respectively, and  $D_{LS}^C = D_S^C - D_L^C$ . The Fermat potential comprises a geometrical term that

depends on image and source positions, and the deflection potential,  $\psi$ , that describes the so-called Shapiro (1964) delay:

$$\tau(\theta, \beta) = \frac{(\theta - \beta)^2}{2} - \psi(\theta). \quad (2)$$

The deflection potential satisfies the Poisson equation,

$$\nabla^2 \psi = 2\kappa, \quad (3)$$

where  $\kappa$  is the projected density of the lens in units of the critical surface mass density,  $\Sigma_{\text{crit}}$ ,

$$\kappa(\theta) \equiv \frac{\Sigma(\theta)}{\Sigma_{\text{crit}}}, \quad (4)$$

where

$$\Sigma_{\text{crit}} = \frac{c^2}{4\pi G} \frac{D_S}{D_L D_{LS}}, \quad (5)$$

and  $D_L$ ,  $D_{LS}$ , and  $D_S$  are angular diameter distances.

By Fermat's principle, the measurable positions of gravitationally lensed sources (commonly referred to as images) are located at the stationary points of the surface described by equation (2). Taking the gradient and setting  $\nabla \tau = 0$  yields the lens equation:

$$\theta = \beta + \nabla \psi(\theta) = \beta + \alpha(\theta), \quad (6)$$

where  $\alpha = \nabla \psi$  is referred to as the deflection angle. In the strong lensing regime multiple images of a distant source that is observed through a gravitational lens occur if equation (6) has multiple solutions,  $\theta_k$ , for a given source position,  $\beta$ .

The angular offset of multiple images of a distant source from the centre of a lens is often characterized by the Einstein radius of the lens. Formally, this is the radius of a continuous ring-like image of a source that is located precisely on the axis of an axisymmetric lens:

$$\theta_E = \left( \frac{4GM}{c^2} \frac{D_S}{D_L D_{LS}} \right)^{1/2}, \quad (7)$$

where  $M$  is the projected mass interior to  $\theta_E$ . Whilst observed gravitational lenses are only approximately axisymmetric,  $\theta_E$  is a useful approximation for the location of strongly lensed images of

distant sources. It also emphasizes the convenience of  $\kappa$ , because  $\kappa \geq 1$  is sufficient for the formation of multiple images (Subramanian & Cowling 1986), and  $\langle \kappa(\lt \theta_E) \rangle \equiv 1$  for axisymmetric lenses.

The time delay relative to an unperturbed signal, as defined by equations (1) and (2), is not measurable because true source positions are not measurable in the presence of a lens along the line of sight. The measurable quantity is the arrival time difference between two gravitationally lensed images of the same source, which is proportional to the difference between the Fermat potentials traversed by the two signals, here denoted  $k = 1, 2$ :

$$c \Delta t_{1,2} = \mathcal{D}[\tau(\theta_1, \beta) - \tau(\theta_2, \beta)]. \quad (8)$$

Equations (2) and (8) reveal that the arrival time difference depends on the details of the gravitational potential of the lens, as discussed further in Section 2.5 and Appendix A.

The apparent size and brightness of gravitationally lensed images differ from the source in the absence of lensing by a factor  $\mu$ , the lens magnification. This is a geometrical effect described by the determinant of the Jacobian matrix of the coordinate transformation from the ‘source plane’,  $\beta$ , to the ‘image plane’,  $\theta$ :

$$\mu = \frac{1}{\det \mathcal{A}}, \quad (9)$$

where

$$\mathcal{A}(\theta) = \frac{\partial \beta}{\partial \theta} = \delta_{ij} - \frac{\partial^2 \psi(\theta)}{\partial \theta_i \partial \theta_j} = \delta_{ij} - \psi_{.ij} \quad (10)$$

and  $\delta_{ij}$  is the identity matrix. Equation (10) implies that the magnification factor is the ratio of the solid angle subtended by a magnified image,  $\omega_1$ , to the solid angle that the source would subtend if it were not lensed,  $\omega_S$ :

$$\mu = \frac{\omega_1}{\omega_S}. \quad (11)$$

The estimated distance to a magnified source of known luminosity is therefore systematically biased by an amount implied by the inverse square law:

$$\mu = \left( \frac{D}{\tilde{D}} \right)^2 \Rightarrow \tilde{D} = \frac{D}{\sqrt{\mu}}, \quad (12)$$

where  $D$  is the true luminosity distance to the source, and we introduce  $\tilde{D}$  as the luminosity distance that would be inferred if  $\mu = 1$  is assumed for a source for which  $\mu \neq 1$ . Phrasing magnification in terms of distance (equation 12) is useful in the context of gravitationally lensed GWs, because the LVK collaboration assumes  $\mu = 1$  when analysing their data immediately after detection of a GW, and therefore they infer  $\tilde{D}$ .

### 2.3 Image multiplicity versus magnification

In principle, calculating the number of gravitationally lensed images of a given source population involves choosing a deflection field,  $\psi$ , and using equation (6) to compute all possible image positions,  $\theta_k$ , for all possible source positions,  $\beta$ . However, whilst it is straightforward to compute  $\beta$  given  $\psi$  and  $\theta_k$ , the inverse mapping from  $\beta$  to an unknown number of  $\theta_k$  can only be solved analytically for very simple mass distributions, and is time consuming to solve numerically.

Simple models for the extended mass distributions of gravitational lenses describe the projected density of the lens as a monotonic function of projected physical distance from the centre of the lens,  $R = \theta D_L$ :

$$\Sigma(R) \propto R^{-\eta}, \quad (13)$$

with a global power-law slope  $\eta > 0$  at all radii. The singular isothermal sphere (SIS;  $\eta = 1$ ) is the canonical analytically solvable case and is also relevant because models with an isothermal slope (generally singular isothermal ellipses) are appropriate to isolated galaxy-scale lenses (e.g. Koopmans et al. 2006). The projected mass density profile of an isothermal mass distribution is given by:

$$\Sigma(R) = \frac{\sigma^2}{2GR}, \quad (14)$$

where  $\sigma$  is the line-of-sight velocity dispersion of the particles in the lens (Binney & Tremaine 1987), and for which it can be shown that the deflection angle is constant and equal to the Einstein radius:

$$\alpha = \theta_E = \frac{4\pi\sigma^2 D_{LS}}{c^2 D_S}. \quad (15)$$

SIS lenses produce pairs of gravitational images with total magnifications of  $\mu_p > 2$ , where  $\mu_p$  is the sum of the magnifications suffered by the image pair (Turner et al. 1984). The limiting case of  $\mu_p = 2$  corresponds to one image (generally denoted as ‘+’) magnified by  $\mu_+ = 2$  and the other image (generally denoted as ‘-’) demagnified such that  $\mu_- \rightarrow 0$ . In general the magnifications suffered by these images are given by:

$$\mu_+ = 1 + \frac{\theta_E}{\beta} \text{ and } \mu_- = \frac{\theta_E}{\beta} - 1, \quad (16)$$

for  $\beta \leq \theta_E$ . SIS lenses therefore produce two images that are both brighter than the source when  $\beta < \theta_E/2$ , which yields  $\mu_p = \mu_+ + \mu_- > 4$ . Note that the lensing behaviour outlined here relates to the pseudo-caustic catastrophe discussed in Section 2.5. Realistic isothermal lenses (i.e. perturbed from circular symmetry by the ellipticity of the lens and/or external shear) can also produce additional images, for example quadruply imaged quasars (e.g. Millon et al. 2020).

Group- and cluster-scale lenses are more massive than galaxy-scale lenses (hence larger  $\theta_E$ ) and typically have density profile slopes at their Einstein radius that are shallower than isothermal. Strong lensing constraints on the density profile slope of group- and cluster-scale lenses have been discussed extensively in the literature dating back to early theoretical work by Turner et al. (1984) and spanning more recent observational studies of increasingly large samples (e.g. Smith et al. 2001; Gavazzi et al. 2003; Sand et al. 2004; Smith et al. 2009; Richard et al. 2010; Umetsu et al. 2016; Fox et al. 2022). We illustrate this in Fig. 1, based on 79 cluster lens models assembled from Richard et al. (2010), Fischer et al. (2019), and Fox et al. (2022), including some models from the SGAS collaboration (Sharon et al. 2020), and the RELICS collaboration (Coe et al. 2019). It is clear that clusters are both denser (relative to  $\Sigma_{\text{crit}}$ ) and flatter than the isothermal model that is typical of galaxy-scale lenses, where we introduce  $\kappa_E$  and  $\eta_E$  as the density and density profile slope, respectively, at the Einstein radius. For reference, the median cluster in this sample is at a redshift of  $z = 0.375$ , with  $\theta_E = 10$  arcsec,  $\kappa_E = 0.78$ , and  $\eta_E = 0.48$ , where  $\theta_E$  and  $\kappa_E$  are stated for a source redshift of  $z = 1.6$  as the typical source redshift found in Section 4.

We show the implications of the structure of cluster-scale lenses for image multiplicity as a function of magnification in the central panel of Fig. 1. The magnification threshold above which cluster-scale lenses produce multiple images is larger than the threshold relevant to SIS lenses discussed above, and is a function of the profile slope,  $\eta_E$ . The magnification threshold for an individual image (i.e. not the sum over an image pair) at which  $f_{\text{multiple}} \geq 0.5$  for the steepest cluster-scale lenses ( $\eta_E > 0.8$ ) is  $\mu \simeq 6$ , and for the shallowest cluster-scale

lenses ( $\eta_E < 0.4$ ) is  $\mu \simeq 20$ , with a threshold for the median cluster at  $\mu \simeq 10$ .

Detailed predictions of the population of multiple images of distant sources formed by the full population of lenses are not feasible at the current time. This is because a model for the covariance of  $\theta_E$ ,  $\eta_E$ , and  $\kappa_E$  across the relevant portion of dark matter halo mass function ( $10^{12} \lesssim M_{200} \lesssim 10^{15} M_\odot$ ) is available neither from observations nor from simulations. Constructing this model will be an important step in the future development of this work. Nevertheless, the observed lenses and supporting theoretical work are sufficient to identify the relevant region of parameter space that brackets the expected behaviour of the relevant lenses. We therefore base our predictions on lens magnification, and exercise caution in interpreting our predictions in terms of multiple imaging in the low-magnification regime ( $2 < \mu \lesssim 10$ ). We adopt a threshold lens magnification of  $\mu > \mu_{\text{th}} = 2$ , as this is the lowest magnification at which any of the lenses considered here produce multiple images.

## 2.4 Optical depth

In this context ‘optical depth’ refers to the fraction of the celestial sphere that is magnified by a certain amount. It can be expressed either in differential form,  $d\tau = \tau_\mu d\mu$ , as the fraction of the sphere magnified by an amount in the range  $\mu \rightarrow \mu + d\mu$ , or cumulative form,  $\tau(\mu > \mu_{\text{th}})$ , as the fraction magnified by more than  $\mu_{\text{th}}$ . It is also important to distinguish between the optical depth in the image plane,  $\tau^I$ , and the source plane,  $\tau^S$ . The former is the fraction of the observed sky – the sphere on which the images are located – that is magnified, whilst the latter is the fraction of the unobservable sphere on which the sources are located that is magnified. In this article we are interested in the number of gravitationally magnified images of a given population of sources. The source-plane optical depth is therefore the most relevant quantity, defined as follows:

$$\tau^S(\mu > \mu_{\text{th}}) = \int_{\mu_{\text{th}}}^{\infty} d\mu \frac{\tau_\mu^I}{\mu}, \quad (17)$$

following Robertson et al. (2020). In particular, computing  $\tau^S$  from  $\tau_\mu^I/\mu$  ensures that both magnification bias is removed and the optical depth is proportional to the number of images that are potentially detectable. Therefore, we predict the number of gravitationally magnified images and thus implicitly integrate over image multiplicity.

Fortunately, the integral of optical depth over lens mass is relatively insensitive to how the structure of gravitational lenses varies with lens mass. Numerous calibrations of  $\tau^S$  as a function of source redshift agree within a factor of 2 (Fig. 1), and show a consistent picture of non-negligible optical depth for dark matter haloes spanning masses  $10^{12} \lesssim M_{200} \lesssim 10^{15} M_\odot$ , i.e. ranging from individual massive galaxies through to massive galaxy clusters (Robertson et al. 2020). As alluded to in Section 2.3, it is also possible to predict the optical depth to multiple imaging by invoking a model for the projected density profile of lenses. We show a recent example of the predicted optical depth to multiple imaging from Haris et al. (2018) as the dashed curve in Fig. 1, which they describe with a simple analytic function that we write as:

$$\tau_{\text{mul}}(z_S) = \left[ \frac{D_S^C}{62.2 \text{ Gpc}} \right]^3. \quad (18)$$

The close agreement between  $\tau_{\text{mul}}$  and  $\tau^S(\mu > 2)$  in Fig. 1 is striking, however it masks important subtleties. Predictions of  $\tau_{\text{mul}}$  assume that the lenses are all individual galaxies with isothermal

density profiles (e.g. Haris et al. 2018; Ng et al. 2018; Oguri 2018), neither of which reflect the true population of lenses as discussed in Section 2.3. Moreover, the threshold total magnification at which an SIS lens produces image pairs is  $\mu_p = \mu_+ + \mu_- = 2$ . Therefore, formally, Haris et al.’s  $\tau_{\text{mul}} = \tau^S(\mu_p > 2)$  applies just to SIS lenses. Nevertheless, equation (18) is a convenient analytic approximation of  $\tau^S(\mu \geq \mu_{\text{th}}, z_S)$ , and we therefore use it to derive an analytic expression for the differential source-plane optical depth,  $\tau_\mu^S$ , by setting

$$\int_{\mu=\mu_{\text{th}}}^{\infty} d\mu \tau_\mu^S(z_S) = \tau^S(\mu \geq \mu_{\text{th}}, z_S) \simeq \tau_{\text{mul}}(z_S) \quad (19)$$

and using the universal scaling  $\tau_\mu^S \propto \mu^{-3}$  for high magnification (Blandford & Narayan 1986), to obtain:

$$\tau_\mu^S(z_S) = \left[ \frac{D_S^C}{31.1 \text{ Gpc}} \right]^3 \mu^{-3}. \quad (20)$$

We use equation (20) in Section 3 when predicting the rate of detectable gravitationally magnified GWs.

## 2.5 Time delay

The arrival time difference between images is directly proportional to the difference between the Fermat potentials (equation 8). Longer arrival time differences are therefore qualitatively interpreted as being attributable to cluster-scale lenses and shorter differences being attributable to galaxy-scale lenses (e.g. Dai et al. 2020). However, the arrival time difference also depends on the lens magnification and on the structure of the gravitational potential of the lens (Witt, Mao & Keeton 2000; Kochanek 2002). At a deeper level, the arrival time difference for an image pair also depends on the type of lens catastrophe that is responsible for forming the image pair. We discuss this in detail in Appendix A, where we show that for two common catastrophes – the pseudo-caustic of lenses that form just two images (of which the SIS is the archetype) and the fold caustic of lenses that are capable of forming more than two images – the arrival time difference between a pair of magnified images is given by:

$$\frac{\Delta t_{\text{pseu}}}{92 \text{ d}} = \left[ \frac{\theta_E}{1 \text{ arcsec}} \right]^2 \left[ \frac{\mu_p}{4} \right]^{-1} \left[ \frac{1 - \kappa_E}{0.5} \right] \left[ \frac{D}{3.3 \text{ Gpc}} \right], \quad (21)$$

$$\frac{\Delta t_{\text{fold}}}{3.9 \text{ d}} = \left[ \frac{\theta_E}{1 \text{ arcsec}} \right]^2 \left[ \frac{\mu_p}{4} \right]^{-3} \left[ \frac{\eta_E}{1} \right]^{-2} \left[ \frac{\kappa_E}{0.5} \right]^{-2} \left[ \frac{1 - \kappa_E}{0.5} \right]^{-3} \times \left[ \frac{D}{3.3 \text{ Gpc}} \right]. \quad (22)$$

Equations (21) and (22) are compatible with published measurements of arrival time difference. In Fig. 2 we show a compilation of arrival time difference measurements from lensed quasars (Fohlmeister et al. 2008, 2013; Dahle et al. 2015; Millon et al. 2020) and supernova Refsdal (Kelly et al. 2016; Rodney et al. 2016). The arrival time difference distribution for lenses that produce just two images ( $N_{\text{img}} = 2$ ) peaks at  $\Delta t \simeq 100 \text{ d}$ , in good agreement with equation (21) at low magnification. The tail to smaller  $\Delta t$  for  $N_{\text{img}} = 2$  is mainly due to a small number of higher magnification pairs. The arrival time difference distribution for lenses that produce more than two images ( $N_{\text{img}} > 2$ ) is broader and extends to lower  $\Delta t$ , due to the broader range of lensing catastrophes included within this sample, and the stronger dependence of  $\Delta t$  on  $\mu$ ,  $\kappa_E$ , and  $\eta_E$  for fold caustics. For example, these lenses include galaxy-scale quadruply imaged quasars in addition to the wide separation quasar lenses SDSS J1004+4112, SDSS J1029+2623, and SDSS J2222+2745, and the massive galaxy cluster MACS J1149.5+2223 that was responsible for supernova

**Table 1.** Models for gravitational wave source populations.

| Model                | Integration limits            |                               | Mass function..... |            |                                    |                          |                               |          |          | Evolution..... |          |      | $\varrho$ |
|----------------------|-------------------------------|-------------------------------|--------------------|------------|------------------------------------|--------------------------|-------------------------------|----------|----------|----------------|----------|------|-----------|
|                      | $m_{\min}$<br>( $M_{\odot}$ ) | $m_{\max}$<br>( $M_{\odot}$ ) | $\gamma_1$         | $\gamma_2$ | $m_{\text{br}}$<br>( $M_{\odot}$ ) | $m_0$<br>( $M_{\odot}$ ) | $\sigma_m$<br>( $M_{\odot}$ ) | $\phi_p$ | $\alpha$ | $\beta$        | $z_p$    |      |           |
| <b>NS–NS mergers</b> |                               |                               |                    |            |                                    |                          |                               |          |          |                |          |      |           |
| Baseline             | 1                             | 2.5                           | 0                  | 0          | ...                                | ...                      | ...                           | 0        | 2.7      | 2.9            | 1.9      | 1.00 |           |
| Steep                | 1                             | 2.5                           | −2                 | −2         | ...                                | ...                      | ...                           | 0        | 2.7      | 2.9            | 1.9      | 0.73 |           |
| Light                | 1                             | 2                             | 0                  | 0          | ...                                | ...                      | ...                           | 0        | 2.7      | 2.9            | 1.9      | 0.65 |           |
| Heavy                | 1                             | 3                             | 0                  | 0          | ...                                | ...                      | ...                           | 0        | 2.7      | 2.9            | 1.9      | 1.45 |           |
| Delayed              | 1                             | 2.5                           | −2                 | −2         | ...                                | ...                      | ...                           | 0        | 3.3      | 2.7            | 1.5      | 1.03 |           |
| Maximal              | 1                             | 2.5                           | −2                 | −2         | ...                                | ...                      | ...                           | 0        | 4.5      | 2.5            | 1.0      | 1.08 |           |
| Power law            | 1                             | 2.5                           | −2                 | −2         | ...                                | ...                      | ...                           | 0        | 2.7      | 0              | $\infty$ | 1.00 |           |
| Non-evolving         | 1                             | 2.5                           | −2                 | −2         | ...                                | ...                      | ...                           | 0        | 0        | 0              | $\infty$ | 0.89 |           |
| <b>BH–BH mergers</b> |                               |                               |                    |            |                                    |                          |                               |          |          |                |          |      |           |
| Baseline             | 5                             | 500                           | −3.3               | − $\infty$ | 50                                 | ...                      | ...                           | 0        | 2.7      | 2.9            | 1.9      | 1.00 |           |
| Tapered              | 5                             | 500                           | −3.3               | −6         | 50                                 | ...                      | ...                           | 0        | 2.7      | 2.9            | 1.9      | 1.28 |           |
| Steep                | 5                             | 500                           | −3.3               | −16        | 50                                 | ...                      | ...                           | 0        | 2.7      | 2.9            | 1.9      | 1.21 |           |
| Perturbed            | 5                             | 500                           | −3.3               | −16        | 50                                 | 33                       | 2                             | 0.03     | 2.7      | 2.9            | 1.9      | 1.81 |           |
| Delayed              | 5                             | 500                           | −3.3               | − $\infty$ | 50                                 | ...                      | ...                           | 0        | 3.3      | 2.7            | 1.5      | 1.36 |           |
| Maximal              | 5                             | 500                           | −3.3               | − $\infty$ | 50                                 | ...                      | ...                           | 0        | 4.5      | 2.5            | 1.0      | 1.79 |           |
| Power-law            | 5                             | 500                           | −3.3               | −16        | 50                                 | ...                      | ...                           | 0        | 2.7      | 0              | $\infty$ | 1.19 |           |
| Non-evolving         | 5                             | 500                           | −3.3               | −16        | 50                                 | ...                      | ...                           | 0        | 0        | 0              | $\infty$ | 0.64 |           |

RRefsdal. This distribution extends up to  $\Delta t \simeq 100$  d, as expected from equation (21), and extends down  $\Delta t \simeq 1$  h. Such short arrival time differences are compatible with the levels of magnification ( $\mu_p \simeq 40$ ) suffered by image pairs produced by typical galaxy cluster lenses ( $\eta_E \simeq 0.5$ ,  $\kappa_E \simeq 0.8$ ), based on equation (22).

### 3 GRAVITATIONALLY LENSING GRAVITATIONAL WAVES

We begin with an overview of the key take away points in Section 3.1. For readers interested in more detail, we explain how we apply the lensing formalism from Section 2 to GW source populations in Section 3.2, which in turn specifies the information that we need about the GW sources. We then describe the populations of NS–NS and BH–BH mergers that we adopt for our lensed GW predictions and verify that they are compatible with the population of sources that have been detected to date in Section 3.3.

#### 3.1 Overview

The basic ingredients required to predict the rates of gravitationally lensed detections of a distant population of objects are (1) the comoving number density of the objects as a function of redshift and a parameter that summarizes signal strength, (2) the comoving volume element as function of redshift, which in turn relies on the cosmological model, (3) the lensing probability, and (4) the detector sensitivity. The least well understood of these ingredients will therefore dominate the uncertainties in the predictions. It is therefore immediately apparent that the least certain ingredient is the first, namely the comoving merger rate density of binary compact object mergers, its evolution with redshift, and the mass function of compact objects (LVK2021). In contrast, the cosmological parameters are known to  $\simeq 10$  per cent, and the optical

depth to gravitational magnification is known to a factor of  $\simeq 2$  (Section 2.4).

We adopt a range of models for the underlying population of GW sources that are compatible with the population detected to date. This includes mass functions that capture the main features of the BH mass function such as a declining logarithmic slope as a function of mass, possible structure superposed on that broadly declining mass function, and reduced efficiency of BH production from stellar evolution at high mass ( $\gtrsim 50 M_{\odot}$ ) due to the pair instability. We also include a broad range of evolution models that are compatible with the data, that are variants on the double power law that is commonly used to describe the evolution of the cosmic star formation rate density evolution.

In Sections 4 and 5, we concentrate on our *Baseline* models of the BH–BH and NS–NS merger populations, as these represent a broad consensus view of the compact object populations. Importantly, in Section 4 we establish that departures from our *Baseline* models have only a small effect on the predicted rate of lensed detections. The *Baseline* models and indeed all other models are specified in Table 1. In brief the *Baseline* ties redshift evolution to the evolution of the cosmic star formation rate density, and adopts a BH mass function that spans  $5 < m < 50 M_{\odot}$  with a power-law slope of  $-3.3$ , and a flat NS mass function that spans  $1 < m < 2.5 M_{\odot}$ . We also make several simplifying assumptions that are consistent with our aim to make broad analytic predictions of the lensed population, that are consistent with the data, and that impact our predictions at the level of a factor of  $\simeq 2$ . The most important of these is that we assume equal mass mergers.

#### 3.2 Arrival and detection rates

The number of gravitationally magnified GWs arriving at Earth per year can be expressed as a function of their apparent mass and distance,  $\tilde{\mathcal{M}}$  and  $\tilde{D}$ , i.e. before correction for the effects of lens

magnification, as follows:

$$\frac{d^2 N_{\text{arr}}}{d\tilde{\mathcal{M}} d\tilde{D}} \Big|_{\mu \geq \mu_{\text{th}}} = \int_{z_0}^{z_{\text{max}}} dz \frac{\mathcal{R}(\mathcal{M}, z)}{1+z} \frac{dV}{dz} \tau_{\mu}^S(\mu), \quad (23)$$

where  $\mathcal{R}$  is the merger rate per unit comoving volume per unit chirp mass,  $dV/dz$  is the comoving volume element,  $\mu(z) = [D(z)/\tilde{D}]^2$ ,  $z_0$  is the redshift to the distance  $D_0$  that satisfies  $\mu = \mu_{\text{th}} = (D_0/\tilde{D})^2$ , and as discussed in Section 2.3 we adopt  $\mu_{\text{th}} = 2$ . Henceforth, we denote  $z_S$  by  $z$ . The chirp mass,  $\mathcal{M}$ , is defined in the usual way:

$$\mathcal{M} = \frac{(m_1 m_2)^{3/5}}{(m_1 + m_2)^{1/5}}, \quad (24)$$

with all masses scaling with  $(1+z)^{-1}$ , consequently the mass ratio,  $q = m_2/m_1$ , is invariant to lens magnification. When needed, we adopt the convention  $m_1 > m_2$ .

The number of gravitationally magnified GWs that are actually detectable per year at Earth depends on the sensitivity of the GW detectors and whether or not they are operating when the signal arrives. We construct  $p_{\text{det}}(\tilde{\mathcal{M}}, \tilde{D})$  for a given GW run using LIGO (the most sensitive of the GW detectors) sensitivity curves based on Martynov et al. (2016) and take account of the impact of the orientation of the GW detectors with respect to incoming signals from close to the detector's horizon following Chen et al. (2021). Although we do not treat the sky localization of GW signals explicitly in this article, we stress that the Virgo and KAGRA detectors play a crucial role in reducing the sky localization uncertainties of GW detections, as for example demonstrated by the detection of GW170817 (Abbott et al. 2017a). We return to this in Section 5 when we discuss EM follow-up observations.

We write the total number of magnified GWs that are available to be detected per year at Earth as:

$$N_{\text{det}} = \int_0^{D_{\text{max}}} d\tilde{D} \int_{\mathcal{M}_{\text{min}}}^{\mathcal{M}_{\text{max}}} d\tilde{\mathcal{M}} \frac{d^2 N_{\text{arr}}}{d\tilde{\mathcal{M}} d\tilde{D}} p_{\text{det}}, \quad (25)$$

with integration limits chosen to span the full region of parameter space over which  $p_{\text{det}} > 0$ . Note that our approach to describing GW detector sensitivity ignores the presence of a network of detectors and assumes that GW waveforms used to search for GW signals are well matched to the lensed source population. Recent work by Yang et al. (2022) indicates that this contributes a systematic uncertainty of order a factor of 2 to our calculations, which is subdominant.

### 3.3 Gravitational wave source populations

We concentrate on NS–NS and BH–BH mergers of roughly equal mass binaries,  $q = 1$ , as these dominate the GW detections to date (LVK2021). The mass of the compact objects that merge are therefore written as  $m = m_1 = m_2 = 2^{0.2} \mathcal{M}$ , when convenient to do so below. Note that  $\simeq 10$ – $20$  per cent departures from  $q = 1$  have no more than a factor of 2 impact on the volume to which the GW detectors are sensitive, and thus considering only  $q = 1$  is both well-motivated by data and not a significant source of systematic bias in our calculations.

We write  $\mathcal{R}$  as a separable function of mass and redshift:

$$\mathcal{R}(z, m) = \mathcal{R}_0 g(z) f(m) \varrho^{-1}, \quad (26)$$

where  $\mathcal{R}_0$  is the comoving merger rate density of binary compact object mergers in the local universe (Section 3.3.1),  $g(z)$  and  $f(m)$  are functions that describe the distribution of sources as functions of redshift and compact object mass, respectively (Sections 3.3.2 and 3.3.3). Note that  $g(z)$  is normalized to ensure

that  $g(z=0) = 1$  and  $f(m)$  is normalized to integrate to unity over the mass range considered. We choose  $\varrho$  to ensure the predicted number of GW detections that are not lensed is invariant between the population models (Table 1).

#### 3.3.1 Local comoving merger rate density, $\mathcal{R}_0$

We adopt local comoving merger rate densities based on the credible ranges from LVK2021, and consistent with Abbott et al. (2018, 2021b). Formally, these rates assume that the source population does not evolve with redshift, and thus are inconsistent with the broad expectation that the populations do evolve with redshift. However, the difference between local rates based on assuming evolving and non-evolving populations is small compared with other uncertainties (Abbott et al. 2021b). Moreover,  $\varrho \simeq 1$  for all models that we consider (Table 1). Therefore, this has negligible impact on our conclusions, and we adopt  $\varrho = 1$  for all models,  $13 < \mathcal{R}_0 < 1900 \text{ Gpc}^{-3} \text{ yr}^{-1}$  for NS–NS, and  $16 < \mathcal{R}_0 < 130 \text{ Gpc}^{-3} \text{ yr}^{-1}$  for BH–BH.

#### 3.3.2 Redshift evolution, $g(z)$

We parametrize redshift evolution as a double power law for both NS–NS and BH–BH mergers:

$$g(z) = C(1+z)^\alpha \left[ 1 + \left( \frac{1+z}{1+z_p} \right)^{\alpha+\beta} \right]^{-1}, \quad (27)$$

where  $z_p$  is the redshift at which the redshift evolution pivots from  $g \propto (1+z)^\alpha$  to  $g \propto (1+z)^{-\beta}$ , and  $C = 1 + (1+z_p)^{-(\alpha+\beta)}$  ensures that  $g(z=0) = 1$ , following Callister et al. (2020).

Our *Baseline* model adopts the canonical values of  $\alpha = 2.7$ ,  $\beta = 2.9$ , and  $z_p = 1.9$ , based on measurements of the star formation rate density (SFRD) history of the universe (e.g. Madau & Dickinson 2014, and references therein). This is equivalent to assuming that the elapsed time from star formation to merger of binary compact objects is zero. Whilst this is clearly unphysical, connecting  $g(z)$  with the cosmic SFRD is a physically well-motivated ansatz, and broadly consistent with population synthesis models of binary compact object mergers (e.g. Dominik et al. 2013; Santoliquido et al. 2021). This model is consistent with the GW source populations detected to date (LVK2021). Moreover, all but the most extreme delay times between star formation and binary compact object mergers – approaching the age of the universe – have a negligible effect on the predicted rate of lensed GWs (Mukherjee et al. 2021). We explore and confirm this by experimenting with alternative models below.

Whilst LVK2021's detection of redshift evolution is strong, their constraints on  $\alpha$  (their  $\kappa$ ) are quite uncertain ( $\alpha = 2.7^{+1.8}_{-1.9}$ ), and only extend to  $z \simeq 1$ , i.e. 2.4 Gyr later than the pivot redshift  $z_p = 1.9$ . We therefore experiment with a *Delayed* model ( $\alpha = 3.3$ ,  $\beta = 2.7$ ,  $z_p = 1.5$ ) and a *Maximal* model ( $\alpha = 4.5$ ,  $\beta = 2.5$ ,  $z_p = 1.0$ ). The parameters of these models were chosen by fixing the pivot redshift at  $z_p = 1.5$  and  $z_p = 1$ , respectively, so as to bracket the range of  $z_p$  compatible with the cosmic star formation rate history and GW constraints on evolution to date. This forces  $\alpha$  for these two models to be steeper than for the *Baseline* model. We chose  $\alpha = 4.5$  for the *Maximal* model, i.e. at the steep end of LVK2021's constraints, then obtained  $\beta = 2.5$  for this model by requiring the integral of  $g(z)$  to be invariant to choice of model parameters – i.e. invariance of the total number of stellar remnant compact object binaries in the universe. We followed a similar procedure to choose  $\alpha$  and  $\beta$  for the *Delayed* model. For reference, we also use *Non-evolving* ( $\alpha = 0$ ,  $\beta = 0$ ,  $z_p \rightarrow \infty$ ) and *power-law* ( $\alpha = 2.7$ ,  $\beta = 0$ ,  $z_p \rightarrow \infty$ ) models.

3.3.3 Mass function,  $f(m)$ 

In general, we parametrize the mass function as a double power law:

$$f(m) \propto \begin{cases} 0 & : m < m_{\min} \\ (m/m_{\text{br}})^{\gamma_1} & : m_{\min} < m < m_{\text{br}} \\ (m/m_{\text{br}})^{\gamma_2} & : m_{\text{br}} < m < m_{\text{max}} \\ 0 & : m > m_{\text{max}}, \end{cases} \quad (28)$$

where  $m_{\text{br}}$  is the ‘break’ at which the power-law slope changes, and  $m_{\min}$  and  $m_{\text{max}}$  are effectively the integration limits of the mass function.

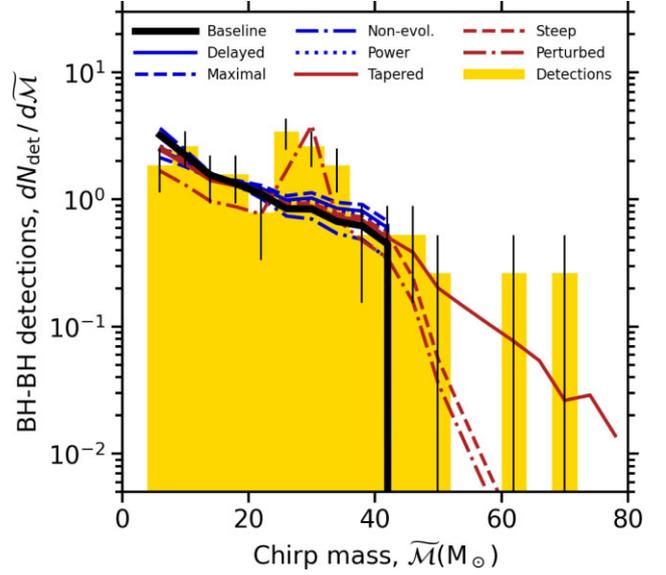
For NS–NS mergers, we adopt a *Baseline* model with  $\gamma_1 = \gamma_2 = 0$ ,  $m_{\min} = 1 M_{\odot}$ , and  $m_{\text{max}} = 2.5 M_{\odot}$ , in line with LVK2021. To test the sensitivity of the predicted rate of lensed NS–NS mergers to these choices, we also use a *Heavy* model that extends to  $m_{\text{max}} = 3 M_{\odot}$ , a *Light* model that is restricted to  $m_{\text{max}} = 2 M_{\odot}$ , and a *Steep* model with  $\gamma_1 = \gamma_2 = -2$ .

For BH–BH mergers, we explore a range of mass functions that are consistent with the theoretical expectation that the efficiency of BH production by stellar evolution declines at  $m \gtrsim 50 M_{\odot}$  due to the pair instability, and the broad features of the GW population detected to date, namely a steeply declining BH mass function and possible structure (i.e. peaks) in the mass function (LVK2021). Our *Baseline* BH mass function is an unbroken power law, implemented as  $\gamma_1 = -3.3$ ,  $\gamma_2 \rightarrow -\infty$ ,  $m_{\text{br}} = 50 M_{\odot}$ , and with  $5 < m < 500 M_{\odot}$ . This model effectively mandates that detections of BH–BH mergers consistent with component masses  $m > m_{\text{br}} = 50 M_{\odot}$  are gravitationally lensed – a physically plausible model that is based on well-understood physics. However,  $\gamma_2 \rightarrow -\infty$  may be rather extreme, and so our *Tapered* model differs from *Baseline* by a less aggressive reduction in efficiency of BH production via  $\gamma_2 = -6$ , and is qualitatively able to match the detection of a few very massive BH–BH mergers without mandating that they are lensed. Our *Steep* model is intermediate between *Baseline* and *Tapered*, with  $\gamma_2 = -16$ , which has been tuned to give a comparable number of apparently very massive BH–BH mergers from both lensing and massive BH formation channels. Finally, our *Perturbed* model differs from our *Steep* model by adding a normally distributed perturbation that is centred at  $m_0 = 33 M_{\odot}$ , with width  $\sigma_m = 2 M_{\odot}$ , and a fraction of the BHs described by the mass function that inhabit this perturbation of  $\phi_p = 0.03$ .

As a consistency check, we show in Fig. 3 that all of the BH models are able to reproduce the broad features of the source-frame chirp mass distribution of BH–BH mergers from LIGO and Virgo’s third run. We use the population of detected BH–BH mergers for this purpose, and not NS–NS merger detections, because of the small numbers of the latter detected to date.

#### 4 RATES AND PROPERTIES OF GRAVITATIONALLY LENSED GRAVITATIONAL WAVES

We present our predictions for the population of gravitationally magnified GWs from BH–BH and NS–NS mergers, concentrating on the *Baseline* model described in Section 3.3 and summarized in Table 1, discussing the stability of our predictions to model choice where relevant. The predictions are organized as follows: detection rates and discovery time-scales (Section 4.1), distance and magnification distributions (Section 4.2), mass distributions (Section 4.3), and time delay distributions (Section 4.4). We compare our predictions with previous work in Section 4.5 and summarize the implications of our predictions for discovery strategies in Section 4.6.



**Figure 3.** The source-plane chirp mass distribution of objects with  $\tilde{M} > 5/2^{0.2} \simeq 4 M_{\odot}$  detected by LIGO and Virgo in their third run, compared with that computed from our BH mass function models (Section 3.3.3), using equation (25) with lensing turned off. The model curves adopt  $\mathcal{R}_0 = 18.3 \text{ Gpc}^{-3} \text{ yr}^{-1} = 30/1.2^{2.7}$  based on the ‘best constrained’ BH–BH merger rate at  $z = 0.2$  given by LVK2021, and reflect a run duration of 11 months with detectors in operation 70 per cent of the time.

**Table 2.** Predicted relative rates of lensed GW detections as a function of GW detector sensitivity and compact object mass function, expressed as the ratio of the number of detectable events magnified by  $\mu > 2$  to the number of detectable events that are not lensed.

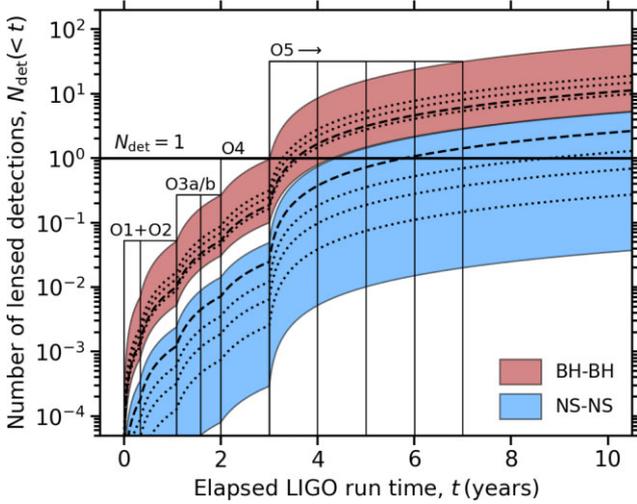
| Model                 | O1       | O2       | O3       | O4<br>'Design' | O5<br>'A+' |
|-----------------------|----------|----------|----------|----------------|------------|
| <b>NS–NS mergers:</b> |          |          |          |                |            |
| Baseline              | 1: 4400  | 1: 3500  | 1: 2500  | 1: 1900        | 1: 970     |
| Heavy                 | 1: 3900  | 1: 3100  | 1: 2200  | 1: 1700        | 1: 860     |
| Light                 | 1: 5200  | 1: 4100  | 1: 2900  | 1: 2300        | 1: 1100    |
| Steep                 | 1: 4800  | 1: 3800  | 1: 2700  | 1: 2100        | 1: 1000    |
| Delayed               | 1: 3900  | 1: 3000  | 1: 2200  | 1: 1700        | 1: 890     |
| Maximal               | 1: 3500  | 1: 2800  | 1: 2000  | 1: 1600        | 1: 890     |
| Non-evolving          | 1: 21000 | 1: 16000 | 1: 11000 | 1: 8600        | 1: 3800    |
| Power law             | 1: 660   | 1: 510   | 1: 360   | 1: 270         | 1: 130     |
| <b>BH–BH mergers:</b> |          |          |          |                |            |
| Baseline              | 1: 1100  | 1: 870   | 1: 750   | 1: 570         | 1: 520     |
| Tapered               | 1: 1100  | 1: 880   | 1: 770   | 1: 590         | 1: 550     |
| Steep                 | 1: 1100  | 1: 850   | 1: 750   | 1: 570         | 1: 540     |
| Perturbed             | 1: 1000  | 1: 800   | 1: 740   | 1: 540         | 1: 600     |
| Delayed               | 1: 980   | 1: 790   | 1: 710   | 1: 580         | 1: 600     |
| Maximal               | 1: 980   | 1: 840   | 1: 810   | 1: 730         | 1: 820     |
| Non-evolving          | 1: 4100  | 1: 3000  | 1: 2400  | 1: 1600        | 1: 1100    |
| Power law             | 1: 320   | 1: 230   | 1: 210   | 1: 120         | 1: 100     |

#### 4.1 Rates of detectable lensed BH–BH and NS–NS mergers

We predict that in the fifth GW run  $\simeq 0.1$  per cent of NS–NS merger detections and  $\simeq 0.2$  per cent of BH–BH merger detections will be magnified by  $\mu > 2$  (Table 2). These relative rates are  $\simeq 5\times$  and  $\simeq 2\times$  larger than in the first run for NS–NS and BH–BH, respectively. Interestingly, the relative rates for lensed NS–NS mergers in the fifth GW run are comparable with the relative rates for lensed BH–BH mergers in the first run. Our predictions are stable to varying the

**Table 3.** Predicted absolute number of detectable lensed (magnified by  $\mu \geq 2$ ) GWs per year as a function of GW detector sensitivity and compact object mass function.

|                       | O1                                    | O2                                    | O3                                    | O4 (Design)                           | O5 (A+)                  |
|-----------------------|---------------------------------------|---------------------------------------|---------------------------------------|---------------------------------------|--------------------------|
| <b>NS–NS mergers:</b> |                                       |                                       |                                       |                                       |                          |
| Baseline              | $1 \times 10^{-5} - 2 \times 10^{-3}$ | $3 \times 10^{-5} - 5 \times 10^{-3}$ | $1 \times 10^{-4} - 2 \times 10^{-2}$ | $3 \times 10^{-4} - 5 \times 10^{-2}$ | $7 \times 10^{-3} - 1.0$ |
| Heavy                 | $2 \times 10^{-5} - 3 \times 10^{-3}$ | $5 \times 10^{-5} - 8 \times 10^{-3}$ | $2 \times 10^{-4} - 3 \times 10^{-2}$ | $6 \times 10^{-4} - 8 \times 10^{-2}$ | $1 \times 10^{-2} - 1.7$ |
| Light                 | $7 \times 10^{-6} - 1 \times 10^{-3}$ | $2 \times 10^{-5} - 3 \times 10^{-3}$ | $7 \times 10^{-5} - 1 \times 10^{-2}$ | $2 \times 10^{-4} - 3 \times 10^{-2}$ | $4 \times 10^{-3} - 0.6$ |
| Steep                 | $8 \times 10^{-6} - 1 \times 10^{-3}$ | $2 \times 10^{-5} - 3 \times 10^{-3}$ | $9 \times 10^{-5} - 1 \times 10^{-2}$ | $2 \times 10^{-4} - 3 \times 10^{-2}$ | $5 \times 10^{-3} - 0.7$ |
| Delayed               | $1 \times 10^{-5} - 2 \times 10^{-3}$ | $4 \times 10^{-5} - 5 \times 10^{-3}$ | $2 \times 10^{-4} - 2 \times 10^{-2}$ | $4 \times 10^{-4} - 6 \times 10^{-2}$ | $8 \times 10^{-3} - 1.2$ |
| Maximal               | $2 \times 10^{-5} - 2 \times 10^{-3}$ | $4 \times 10^{-5} - 6 \times 10^{-3}$ | $2 \times 10^{-4} - 2 \times 10^{-2}$ | $5 \times 10^{-4} - 7 \times 10^{-2}$ | $9 \times 10^{-3} - 1.3$ |
| Non-evolving          | $2 \times 10^{-6} - 3 \times 10^{-4}$ | $6 \times 10^{-6} - 9 \times 10^{-4}$ | $3 \times 10^{-5} - 4 \times 10^{-3}$ | $7 \times 10^{-5} - 1 \times 10^{-2}$ | $1 \times 10^{-3} - 0.2$ |
| Power-law             | $8 \times 10^{-5} - 1 \times 10^{-3}$ | $2 \times 10^{-4} - 3 \times 10^{-2}$ | $9 \times 10^{-4} - 0.1$              | $2 \times 10^{-3} - 0.4$              | $5 \times 10^{-2} - 7.4$ |
| <b>BH–BH mergers:</b> |                                       |                                       |                                       |                                       |                          |
| Baseline              | $4 \times 10^{-3} - 4 \times 10^{-2}$ | $1 \times 10^{-2} - 0.1$              | $4 \times 10^{-2} - 0.3$              | 0.1–1.0                               | 1.4–11                   |
| Tapered               | $6 \times 10^{-3} - 5 \times 10^{-2}$ | $2 \times 10^{-2} - 0.1$              | $5 \times 10^{-2} - 0.4$              | 0.2–1.3                               | 1.6–13                   |
| Steep                 | $5 \times 10^{-3} - 4 \times 10^{-2}$ | $2 \times 10^{-2} - 0.1$              | $5 \times 10^{-2} - 0.4$              | 0.2–1.3                               | 1.6–13                   |
| Perturbed             | $9 \times 10^{-3} - 7 \times 10^{-2}$ | $2 \times 10^{-2} - 0.2$              | $8 \times 10^{-2} - 0.6$              | 0.3–2.0                               | 1.9–16                   |
| Delayed               | $6 \times 10^{-3} - 5 \times 10^{-2}$ | $2 \times 10^{-2} - 0.1$              | $6 \times 10^{-2} - 0.5$              | 0.2–1.4                               | 1.8–14                   |
| Maximal               | $8 \times 10^{-3} - 6 \times 10^{-2}$ | $2 \times 10^{-2} - 0.2$              | $7 \times 10^{-2} - 0.6$              | 0.2–1.6                               | 1.8–15                   |
| Non-evolving          | $1 \times 10^{-3} - 8 \times 10^{-3}$ | $3 \times 10^{-3} - 2 \times 10^{-2}$ | $8 \times 10^{-3} - 7 \times 10^{-2}$ | $3 \times 10^{-2} - 0.2$              | 0.2–2                    |
| Power law             | $2 \times 10^{-2} - 0.2$              | $6 \times 10^{-2} - 0.4$              | 0.2–1.5                               | 0.7–6.0                               | 9.0–73                   |



**Figure 4.** Predicted cumulative number of detected gravitationally magnified BH–BH (red upper shaded band) and NS–NS (blue lower shaded band) mergers for the *Baseline* models described in Section 3.3 and Table 1. Detection of magnified BH–BH mergers are predicted to be a regular occurrence in the fifth GW run and beyond, and detection of magnified NS–NS mergers are predicted to become more likely the longer that LIGO in particular operates at its A+ sensitivity or better in the fifth run. In this figure  $N_{\text{det}}$  is lower than the integral of  $N_{\text{det}}$  from equation (25) by a factor equal to the GW detector duty cycle, which is taken to be 0.6 in the first two runs and 0.7 in later runs. The durations of the first three runs are matched to their actual duration, the fourth run is assumed to last 1 yr (expected to start in 2023), and the fifth and later runs are assumed to span multiple years into the late 2020s. The horizontal solid line marks the detection of one magnified source,  $N_{\text{det}} = 1$ . The width of the shaded bands matches the ranges quoted on the values listed in Table 3, i.e. they reflect the full range of uncertainties discussed by LVK2021. The dotted curves show predictions that correspond to the central values of the three compact object mass functions from LVK2021 that contribute to the ranges on  $\mathcal{R}_0$  that we adopt; the dashed curves correspond to LVK2021’s ‘PDB (pair)’ model.

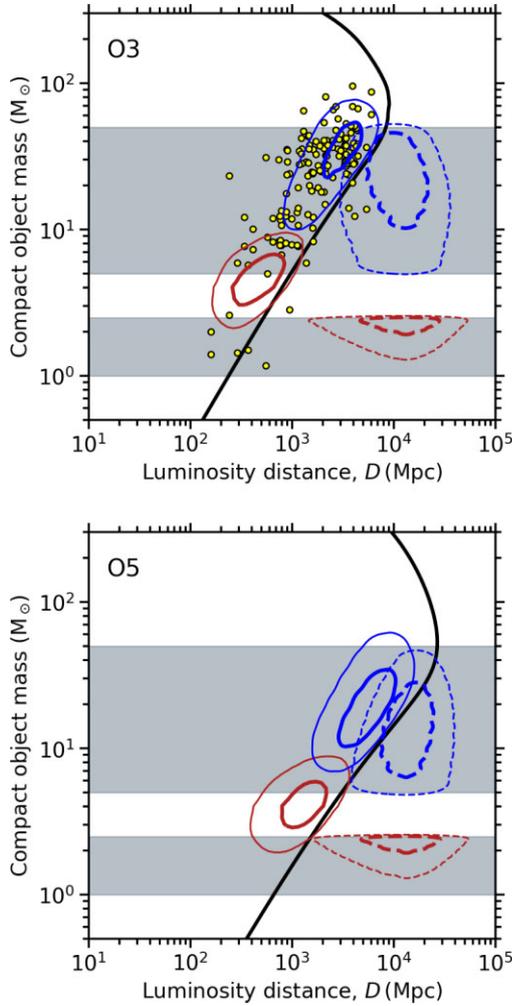
source population models, with the relative rate of lensed detections varying by just  $\simeq 10$ –20 per cent between models (Table 2). The only exceptions are the *non-evolving* and *power-law* models that have been included here for reference.

We predict that the number of detectable gravitationally lensed BH–BH mergers will for the first time approach one per year in the fourth GW run (Table 3), with a handful of detections per year in the fifth run. We also predict that the number of detectable gravitationally lensed NS–NS mergers will lag just one run behind lensed BH–BH mergers, with up to  $\simeq 1$  lensed NS–NS being detectable per year during the fifth run (Table 3). The dominant uncertainty on these rates is the uncertainty on the local rates of compact object mergers (LVK2021). Importantly, the predicted rates vary by less than a factor two across the relevant source population models, with only the *non-evolving* and *power-law* models deviating strongly.

We combine the predicted detection rates from Table 3 with the global fraction of an LVK run for which the GW detectors are obtaining data to show the predicted cumulative number of detectable GWs as a function of run time in Fig. 4. The predicted cumulative number of gravitationally magnified compact object merger detections grows steadily with time, with detection becoming routine in the mid-2020s. Given the strong prospects for first detection in upcoming runs, it is important to consider how to recognize a lensed detection among the growing number of detections that will be made. We address this in the following sections.

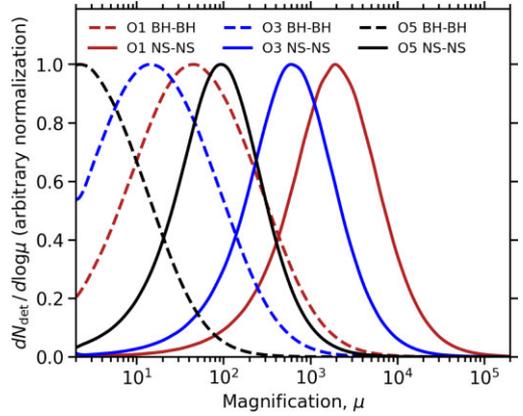
#### 4.2 Distance and magnification distributions

Gravitationally lensed BH–BH and NS–NS mergers are detected as being louder and closer to Earth than they really are, as a consequence of the lens magnification (equation 12). We predict that the centre of the true mass–distance ( $m$ – $D$ ) distribution of both lensed BH–BH and lensed NS–NS mergers is located at  $D \simeq 15$  Gpc for all five GW runs considered in this article (dashed contours in Fig. 5). As the GW detectors become more sensitive from one run to the



**Figure 5.** Predicted distributions of gravitationally magnified NS–NS (red contours) and BH–BH (blue contours) merger detections in the third (top panel) and fifth (bottom panel) GW runs for our *Baseline* models. Solid and dashed contours show the  $\tilde{m}\text{-}\tilde{D}$  (i.e. as inferred in low latency by the LVK collaboration assuming  $\mu = 1$ ) and  $m\text{-}D$  (i.e. true intrinsic values) distributions, respectively. In each case the thicker (inner) and thinner (outer) contours encircle 50 and 90 per cent of the predicted magnified population, respectively. The grey horizontal bands show the mass range encompassed by the mass functions in our *Baseline* models. The points show the masses of the individual compact objects that comprise the population of binary compact object mergers detected through to the end of the third GW run. The thick black curve in each panel shows LIGO’s horizon for equal mass mergers based on Martynov et al. (2016).

next, the distance out to which it is possible to detect gravitationally magnified GWs increases, as seen in the rightward shift of the solid contours between left and central panels in Fig. 5. This signifies the gradual reduction in the lens magnification required to bring faint distant sources within the LIGO horizon. This is summarized in Fig. 6, where we see that the typical lens magnification relevant to BH–BH mergers reduces from  $\mu \simeq 50$  in the first run to  $\mu \simeq 2$  in the fifth run, and relevant to NS–NS mergers from  $\mu \simeq 2000$  in the first run to  $\mu \simeq 100$  in the fifth run. Therefore in the fifth run, the lens population responsible for detectable multiply-imaged BH–BH mergers will be dominated by galaxy-scale lenses, with cluster-scale lenses also contributing detectable singly imaged magnified BH–BH mergers. In contrast, lenses of all mass scales will contribute to the detectable multiply-imaged NS–NS population in the fifth run, with



**Figure 6.** Predicted lens magnification distributions as a function of sensitivity achieved/forecast in the first, third, and fifth GW run, for lensed BH–BH (dashed) and lensed NS–NS (solid) mergers, and based on our *Baseline* models. The black (left-most) dashed and solid curves show that in the fifth run lensed BH–BH will be dominated by low-magnification events  $\mu \simeq 2\text{--}10$ , whilst lensed NS–NS will typically be magnified by  $\mu \simeq 100$ . The latter is comparable with the typical lens magnifications suffered by putative lensed BH–BH in the first run (right-most dashed curve) in the first run. The predicted rate of detectable lensed NS–NS in O5 is higher than for detectable lensed BH–BH in O1 (Table 3) because the intrinsic comoving merger rate density of mergers is higher for NS–NS than for BH–BH (LVK2021). In general, the curves move to the left (lower lens magnification required for detection) as the GW detectors become more sensitive, as seen for example in the LIGO horizon in O5 being rightwards of their horizon in O3 in Fig. 5.

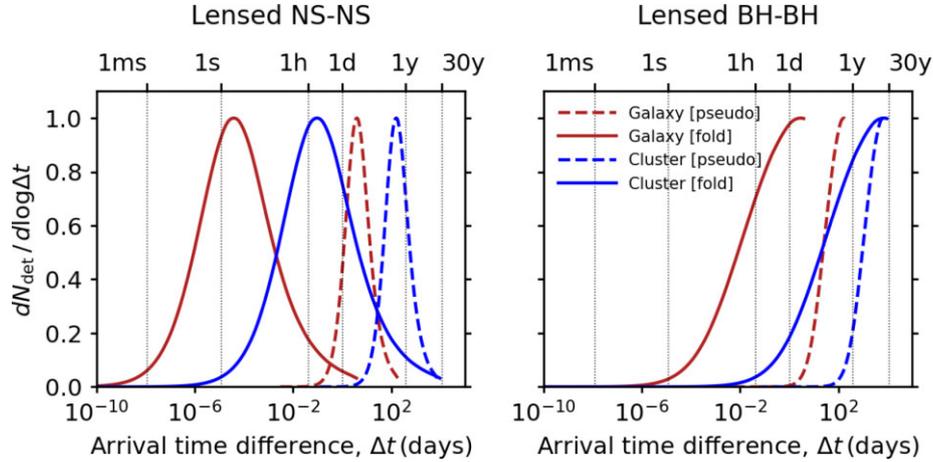
a negligible fraction of detectable singly imaged NS–NS mergers located behind cluster-scale lenses.

### 4.3 Mass distributions

In this section, we use Fig. 5 to consider the  $\tilde{m}$  distribution as a tool for selecting candidate lensed GWs.

The majority of lensed BH–BH mergers do not have anomalously large  $\tilde{m}$ , with just 20 per cent and 5 per cent of them having  $\tilde{m} > 50 M_{\odot}$  in LIGO/Virgo’s third and fifth runs, respectively. The ‘ordinariness’ of the predicted  $\tilde{m}$  values of lensed BH–BH mergers is caused by the declining slope of the BH mass function (LVK2021), as implemented in all of our mass function models (Table 1) – i.e. there are relatively few BH–BH mergers of intrinsically high mass available to be magnified so as to appear to have anomalously large mass. Selecting candidate lensed BH–BH mergers based on  $\tilde{m}$  is therefore inefficient. We stress that we do not interpret Fig. 5 as indicating that most BH–BH detections to date are lensed, and draw attention to the relative detection rates of lensed GWs listed in Table 2.

Selecting candidate lensed NS–NS based on  $\tilde{m}$  appears promising because the predicted  $\tilde{m}$ -distribution of lensed NS–NS mergers peaks within the mass gap between NS and BH (Figs 5). Crucially, the lensed NS–NS  $\tilde{m}$  distribution *peaks* in this region of parameter space, in contrast to the *tail* of lensed BH–BH mergers at  $\tilde{m} > 50 M_{\odot}$  discussed above. The fraction of lensed NS–NS mergers predicted to be inferred by the LVK collaboration to be in this mass gap at  $2.5 < \tilde{m} < 5 M_{\odot}$  is 62 per cent in their fifth run – i.e.  $\simeq 12 \times$  more efficient as an approach for candidate lensed GW selection than the 5 per cent discussed above for lensed BH–BH candidates. This is insensitive to the choice of NS population model (Table 1):  $\gtrsim 50$  per cent of lensed NS–NS detections reside in the lower mass gap for all of the NS mass functions considered.



**Figure 7.** Lensed NS–NS mergers are better matched to the duration of GW detector runs than lensed BH–BH mergers because the arrival time difference between the GW signals from a lensed NS–NS image pair is predicted to be sub-year (left-hand panel), whilst the arrival time difference for lensed BH–BH image pairs is predicted to extend out to numerous years and potentially decades (right-hand panel). The panels show predicted arrival time difference based on our *Baseline* model at the sensitivity forecast for the fifth GW run. Distributions are shown for typical galaxy-scale ( $\theta_E = 1$  arcsec,  $\kappa_E = 0.5$ ,  $\eta_E = 1$ ) and cluster-scale ( $\theta_E = 10$  arcsec,  $\kappa_E = 0.8$ ,  $\eta_E = 0.5$ ) lenses, as discussed in Section 2.5.

#### 4.4 Time delay distributions

We use equations (21) and (22) to transform the predicted magnification distributions derived from our models (Fig. 6) into predicted arrival time difference distributions ( $\Delta t_{\text{pseu}}$  and  $\Delta t_{\text{fold}}$ ) for typical galaxy-scale ( $\theta_E = 1$  arcsec,  $\kappa_E = 0.5$ ,  $\eta_E = 1$ ) and cluster-scale ( $\theta_E = 10$  arcsec,  $\kappa_E = 0.8$ ,  $\eta_E = 0.5$ ) lenses (Fig. 7). We adopt  $\mathcal{D} = 3.3$  Gpc, based on  $z_S = 1.6$  which corresponds to the peak of the predicted true distance ( $D$ ) distribution discussed above, and  $z_L = 0.5$  which corresponds to the peak of the optical depth to strong lensing (Robertson et al. 2020). At fixed lens mass and structure the predicted  $\Delta t_{\text{fold}}$  distributions are broader than the predicted  $\Delta t_{\text{pseu}}$  distributions due to the stronger dependence of the former on lens magnification,  $\mu$ . Also, the predicted  $\Delta t_{\text{fold}}$  distributions peak at smaller arrival time differences than the predicted  $\Delta t_{\text{pseu}}$  distributions, again due to their respective dependence on  $\mu$ .

Gravitationally lensed BH–BH mergers suffer longer arrival time differences than gravitationally lensed NS–NS mergers because the former are typically less strongly magnified ( $\mu \lesssim 10$ ) than the latter ( $\mu \simeq 100$ ; Fig. 6). Importantly, the predicted arrival time differences for lensed NS–NS mergers are dominated by sub-year time-scales, whilst the arrival time differences for lensed BH–BH mergers can extend to numerous years, and in some cases decades. Detection of multiple signals from gravitationally lensed NS–NS mergers are therefore better matched to the length of GW detector runs than detection of multiple signals from lensed BH–BH mergers.

#### 4.5 Comparison with previous work

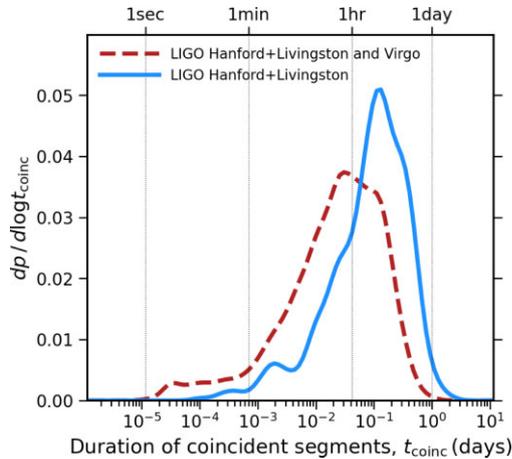
Our predictions for the rates of lensed GWs are broadly consistent with comparable earlier studies (e.g. Li et al. 2018; Ng et al. 2018; Oguri 2018, 2019; Wierda et al. 2021; Xu et al. 2022). This is largely due to these studies adopting a broadly consistent description of the source populations and the lensing optical depth. However, these studies assumed that all lenses are galaxy-scale lenses and thus predicted the rate of multiply-imaged BH–BH (and in some cases NS–NS) mergers under that assumption. In contrast, we predict the rate of gravitationally magnified GWs agnostic to the mass

of the dark matter halo that hosts the lens, and then interpret that prediction in the context of the predicted magnification distribution and the phenomenology of lenses as a function of their mass and structure (Section 2.3). For reasons discussed in Section 2.4, only a subset of the predicted lensed BH–BH population predicted in earlier work will be multiply-imaged due to the inefficiency of group/cluster-scale lenses at forming multiple images at  $\mu < 10$ . This underlines the importance of characterizing the relationship between magnification and image multiplicity as a function of lens mass and structure with upcoming surveys such as the Vera C. Rubin Observatory’s (Rubin’s) Legacy Survey of Space and Time (LSST), ESA’s *Euclid* mission, and spectroscopic follow-up with ESO’s 4MOST instrument.

#### 4.6 Implications for discovery strategies

Looking across the population of GW detections, it is clear from Fig. 5 that essentially all GW detections within  $\simeq 1$  decade in  $\tilde{D}$  of the horizon and with compact object masses in the range  $2 < \tilde{m} < 100 M_\odot$  are candidate lensed detections. It is therefore particularly challenging to identify an individual detection at  $\tilde{m} > 5 M_\odot$  as a candidate lensed event because this overlaps with the BH mass function and the majority of BH–BH detections that are not expected to be lensed. Moreover, selecting candidate lensed BH–BH mergers in the  $\tilde{m}$ – $\tilde{D}$  plane can be sensitive to population model choices (e.g. Broadhurst et al. 2022). In the  $\tilde{m}$ – $\tilde{D}$  plane, selecting candidate lensed NS–NS is  $\simeq 10 \times$  more efficient (and less model dependent) than selecting candidate lensed BH–BH, because the majority of the former resides in the mass gap between NS and BH products,  $2.5 < \tilde{m} < 5 M_\odot$  (Fig. 5). Indeed, objects have already been detected with non-zero probability in this region of the parameter space (LVK2021), some of which have been followed up electromagnetically (Bianconi et al. 2022).

Detection of two GW signals with consistent sky locations can also motivate the identification of candidate lensed objects (e.g. Dai et al. 2020; The LIGO Scientific Collaboration 2021; Bianconi et al. 2022). This strategy has two main disadvantages. First, the significant sky localization uncertainties of a growing catalogue of GW detections (Petrov et al. 2022) are prone to generating false



**Figure 8.** Distribution of coincident segment durations from the third GW run, as discussed in Section 4.6, and based on data obtained from the Gravitational Wave Open Science Centre (Abbott et al. 2021a).

associations between detections. Secondly, each GW detector only collects data of a quality required for detection of binary compact object mergers for  $\simeq 70$  per cent of the duration of a run, and the typical coincident segment length is up to just a few hours. A segment is a period of time that a detector is collecting data and coincidence refers to overlap of segments at the different detectors (Fig. 8). Moreover, the typical run length is a year with inter-run gaps of a similar duration. Putting these practical constraints together, the distribution of coincident segment lengths points to overlapping sky localizations (and other parameter posteriors) of multiple detections within a few hours – i.e. relatively short arrival time differences – as being a good match between the operation of GW detectors and the predicted lensed GW population. This is an excellent match with lensed NS–NS mergers that produce image pairs formed by a fold caustic (Fig. 7). Clearly longer GW runs and within-run operational strategies that offset downtime of the different GW detectors, and thus minimize gaps between coincident segments will be advantageous for the detection of more than one image of a lensed GW.

In summary, multiple arguments point to lensed NS–NS as having significant advantages for securing a robust and unambiguous lensed GW detection. These include short time delays, potential localization to a lensed host galaxy, and selection of candidates in the  $\tilde{m}$ – $\tilde{D}$  plane in low latency that is less model dependent and  $\simeq 10 \times$  more efficient than for candidate lensed BH–BH mergers.

## 5 LENSED KILONOVA COUNTERPARTS: LIGHT CURVES AND OBSERVING STRATEGY

We now discuss the prospects for localizing candidate lensed NS–NS mergers via the EM signal that accompanies them, focusing on prompt optical detection of kilonova counterparts because the detection of kilonovae is less dependent on viewing angle than the detection of SGRBs (Metzger, Thompson & Quataert 2018). In Section 5.1 we give an overview of the kilonova models that we use, in Section 5.2 we present and discuss predicted light curves of lensed kilonova counterparts, in Section 5.3 we show that our reference light curve predictions are relatively insensitive to our assumptions, and in Section 5.4 we discuss observing strategies required to detect the lensed kilonovae.

**Table 4.** Physical parameters of 170817-like and conservative kilonova models used for light-curve predictions.

| Parameter                      | Conservative    | 170817-like       |
|--------------------------------|-----------------|-------------------|
| Chirp mass, $\mathcal{M}$      | $1.2 M_{\odot}$ | $1.188 M_{\odot}$ |
| Mass ratio, $q$                | 0.9             | 0.92              |
| Viewing angle                  | $60^{\circ}$    | $32^{\circ}$      |
| Blue ejecta enhancement factor | 1               | 1.6               |
| Cocoon opening angle           | $0^{\circ}$     | $24^{\circ}$      |

### 5.1 Overview of kilonova models

We use the recent kilonova light-curve models from Nicholl et al. (2021). These models, implemented within the Modular Open Source Fitter for Transients (MOSFIT; Guillochon et al. 2018), include low, high, and intermediate opacity ejecta components, representing the polar dynamical, tidal dynamical, and post-merger (wind) ejecta, respectively. The mass and velocity of each component is determined for a given binary configuration (chirp mass, mass ratio, and tidal deformability) using fits by Dietrich & Ujevic (2017) and Coughlin et al. (2019) to numerical merger simulations.

The luminosity is computed using the  $r$ -process decay heating rate (Korobkin et al. 2012) with time-dependent thermalization efficiency (Barnes et al. 2016) and the Arnett (1982) photon diffusion approximation. Further details of the  $r$ -process decay luminosity in MOSFIT are provided by Villar et al. (2017). The Nicholl et al. models also include a viewing angle dependence based on Darbha & Kasen (2020), and an additional luminosity source from cooling emission of a ‘cocoon’, shock heated by a GRB jet escaping the ejecta, following Piro & Kollmeier (2018). At a given time and wavelength, the apparent magnitude is calculated by summing black-body spectral energy distributions for each component, adjusting for cosmological expansion at the specified redshift, and convolving with the transmission curve of the respective optical/near-infrared filters.

Given the diverse light curves that this model can produce, we elect here to study two representative cases that encapsulate the plausible range in luminosity. The optimistic or ‘170817-like’ model simply takes the best-fitting parameters for GW170817, as determined by Nicholl et al. (using the chirp mass inferred from its GW signal; Abbott et al. 2017a), and evaluate this model at each redshift of interest. In the ‘conservative’ case, we choose a chirp mass ( $\mathcal{M} = 1.2 M_{\odot}$ ) and mass ratio ( $q = 0.9$ ) that are typical of known neutron star binaries that will merge within a Hubble time (Farrow, Zhu & Thrane 2019). These values are similar to the 170817-like model, but three factors conspire to make this conservative case fainter (especially at blue wavelengths): we do not include blue ejecta enhancement by magnetic winds (Metzger et al. 2018) or shock heated GRB cocoon emission (Piro & Kollmeier 2018), and we rotate to a more representative viewing angle of  $60^{\circ}$  from the binary axis. The latter step further suppresses the blue emission (compared to the 170817-like model at  $\approx 30^{\circ}$  off-axis), as less of the low opacity (lanthanide-poor) polar ejecta is within the line of sight for an edge-on observer. Both models are summarized in Table 4.

### 5.2 Reference light curves of lensed kilonovae

The observable light curve of a gravitationally magnified kilonova at a redshift of  $z$  is related to the intrinsic light curve of the kilonova via the lens magnification and can be written in terms of apparent

**Table 5.** Typical distances, redshifts, and magnifications of lensed NS–NS and lensed kilonova counterparts as a function of GW detector sensitivity.

| Parameter         | O1    | O2    | O3   | O4   | O5   |
|-------------------|-------|-------|------|------|------|
| $\tilde{D}$ (Mpc) | 270   | 350   | 500  | 600  | 1250 |
| $D$ (Gpc)         | 12.1  | 12.1  | 12.1 | 12.1 | 12.1 |
| $\mu$             | 2000  | 1200  | 600  | 400  | 100  |
| $2.5\log\mu$      | 8.3   | 7.7   | 6.9  | 6.5  | 5.0  |
| $\tilde{z}$       | 0.059 | 0.075 | 0.10 | 0.13 | 0.24 |
| $z$               | 1.6   | 1.6   | 1.6  | 1.6  | 1.6  |

magnitudes in a given filter as:

$$\begin{aligned} m_{\text{obs}}(z) &= m_{\text{int}}(z) - 2.5\log\mu \\ &= m_{\text{int}}(z) - 5\log(D/\tilde{D}). \end{aligned} \quad (29)$$

A forward prediction of  $m_{\text{obs}}$  for a given kilonova model therefore requires a library of intrinsic light curves that describe  $m_{\text{int}}$  (in our case, based on Nicholl et al.), and a choice of  $\tilde{D}$  and  $D$ . To establish reference light curves for lensed kilonova counterparts to lensed NS–NS mergers, we base our forward predictions on our *Baseline* NS–NS model, the peak of the distributions shown in Fig. 5 for the fourth and fifth GW runs (Table 5), and the transmission curves of the Rubin/LSST *ugrizy*-band filters.

Our reference light curves reveal that lensed kilonova counterparts are faint and fade very quickly (Fig. 9). In the fourth GW run ‘170817-like’ and ‘conservative’ lensed kilonovae will be detectable at  $AB \simeq 22$  and  $AB \simeq 24 - 25$  close to peak, respectively, corresponding to a typical lens magnification for detectable lensed NS–NS mergers of  $\mu \simeq 400$ . Because the sensitivity to faint or less magnified events will improve in the fifth GW run ( $\tau^S \propto \mu^{-2}$ ; equations 19 and 20), the typical lens magnification will be lower, at  $\mu \simeq 100$ . A typical light curve will thus peak 1.5 mag fainter than in the fourth run, at  $AB \simeq 23.5$  and  $AB \simeq 25.5 - 26.5$  for 170817-like and conservative kilonovae, respectively. All the bright events will still be detectable – the overall endeavour is not becoming more challenging – but there will be many more frequent events at lower magnifications that can be explored.

Looking at *gri*-bands as the most sensitive of the optical/near-infrared filters from the ground, a clear signature of lensed kilonova counterparts is that they fade by  $\gtrsim 1$  magnitude per day within the first few days after peak (Fig. 9). As such, they are the fastest fading class of transient sources, with a duration of  $d \lesssim 1$  d (time taken to fade by a factor of two in flux from peak brightness in the observer’s frame). They fade even faster than kilonovae that are not lensed, fast blue optical transients (FBOTs), and rapidly evolving transients (RETs), that all have  $d \gtrsim 2$  d in the observer’s frame (e.g. Perley et al. 2020). This rapid evolution, and in particular its pace relative to kilonovae that are not lensed, is due to the filters probing the cosmologically time-dilated rest-frame UV emission from the distant lensed kilonovae ( $\lambda_{\text{rest}} \simeq 150 - 300$  nm), as opposed to the rest-frame optical emission for nearby kilonovae that are not lensed ( $\lambda_{\text{rest}} \simeq 400 - 700$  nm). Importantly, the pace of evolution of lensed kilonova counterparts is independent of the kilonova model, as can be seen by comparing solid and dashed curves in Fig. 9.

In summary, lensed kilonova counterparts are distinctive because they fade faster than other optical transients studied to date. Also, in the mid/late-2020s they will typically be even fainter than the optical transients that Rubin/LSST is expecting to discover. Fast (spanning a few hours to a few days) and deep ( $AB \simeq 25 - 26$ ) target of opportunity observations will therefore be required to detect them.

### 5.3 Sensitivity of light curves to parameter choices

The features of our reference light curves are robust to changing the chirp mass and mass ratio of the lensed NS–NS merger from which the lensed kilonova emission emanates. In Fig. 10 we show that varying chirp mass in the range  $0.9 < \mathcal{M} < 1.8 M_{\odot}$  affects the apparent magnitude at peak brightness by  $\lesssim 0.5$  magnitudes, and that the slope of the light curves within the first 2 d post NS–NS merger is not significantly altered by variations in chirp mass. We also show that more extreme mass ratio mergers produce brighter kilonovae with light curves that fade more slowly than those with  $q \simeq 0.9 - 1$ , under the assumption that the fraction of ejecta shock heated by the GRB jet is independent of  $q$ . However, in mergers with smaller mass ratios the bulk of the ejecta may be in the tidal plane, so this assumption may not hold if the GRB jet does not interact with such ejecta. Nevertheless, the most extreme mass ratio considered here ( $q = 0.5$ ) still has an observer frame duration of  $d < 2$  d, and thus our conclusion that lensed kilonovae are faster than all other transients is robust to variations in  $q$ . Basing our reference light curves on  $q = 0.9$  is therefore conservative in the context of considering optical follow-up observing strategies in Section 5.4.

Basing our reference light curves on the most probable distances and redshifts derived from our lens models is convenient, however the distributions shown in Fig. 5 are broad. We therefore examine the sensitivity of our reference light curves to these distributions. Sensitivity to  $\tilde{D}$  for a given  $D$  is relatively straightforward to deal with, because the faintness of lensed kilonovae implies that it is only sensible to follow-up the candidate lensed NS–NS mergers with relatively small sky localization uncertainties. Better localized GW detections are located at smaller distances than worse localized detections (Petrov et al. 2022), which implies smaller  $\tilde{D}$  for a given  $D$ , and thus larger lens magnification and brighter kilonovae. Therefore, when observing strategy considerations are folded in, basing the reference light curves on the peak distances and redshifts is conservative.

Nevertheless,  $D$  is clearly unknown when identifying candidate lensed NS–NS mergers in low latency, when the LVK collaboration announces estimates of  $\tilde{D}$ , sky localization posteriors, and the probability that one of the compact objects occupies the lower mass gap at  $3 < \tilde{m} < 5 M_{\odot}$ ,  $p_{\text{gap}}$ . We therefore introduce a  $k$ -correction,  $k(\tilde{z}, z, \lambda)$ , as a convenient way to capture the uncertainty arising from unknown  $D$ . We write  $m_{\text{int}}(z)$  in terms of the apparent magnitudes the kilonova would have if it was located at  $\tilde{z}$ ,  $m_{\text{int}}(\tilde{z})$ , a term to account for the inverse square law that recognizes the impact of the difference between  $\tilde{D}$  and  $D$ , and the  $k$ -correction that accounts for the different amounts of redshifting of the rest-frame spectrum:

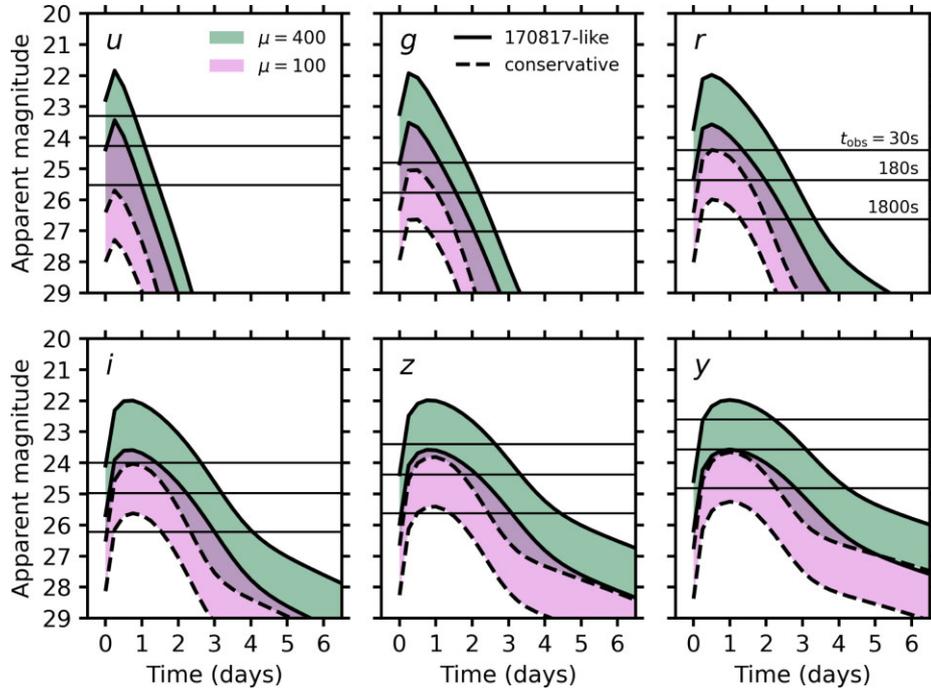
$$m_{\text{int}}(z) = m_{\text{int}}(\tilde{z}) + 5\log(D/\tilde{D}) + k(\tilde{z}, z, \lambda), \quad (30)$$

and then substitute equation (30) into equation (29) to obtain the following simple equation:

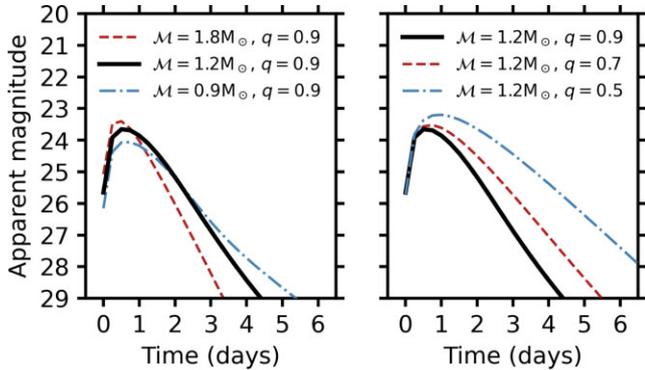
$$m_{\text{obs}}(z) = m_{\text{int}}(\tilde{z}) + k(\tilde{z}, z, \lambda). \quad (31)$$

Equation (31) can be paraphrased as the light curve of the lensed kilonova counterpart to a lensed NS–NS merger is set by the light curve the object would have if it was really located at  $\tilde{D}$  and the  $k$ -correction between  $\tilde{D}$  and  $D$ . In other words, the inverse square law cancels with the lens magnification, and the unknown  $D$  is encoded in the  $k$ -correction.

The  $k$ -corrections at redshifts of interest are not well constrained by existing data. At  $z \simeq 1 - 2$ , most of the optical filters probe rest-frame UV emission, and even AT2017gfo was too faint for UV spectroscopy by the time observations were attempted ( $t = 5.5$  d;



**Figure 9.** Predicted light curves in the observer frame for lensed kilonova counterparts to lensed NS–NS mergers, based on the Vera C. Rubin Observatory’s *ugrizy*-band transmission curves, with each panel labelled according to the respective filter. The solid and dashed curves show the predicted light curve for a kilonova counterpart that is similar to AT2017gfo counterpart to GW170817, and a redder more conservative model, respectively. The upper and lower solid curves show the predicted 170817-like light curve for lens magnifications representative of lensed NS–NS detections in the fourth and fifth GW runs, respectively (Table 5). Similarly, the upper and lower dashed curves are representative of the fourth and fifth runs, respectively. The green (upper) and pink (lower) shaded regions therefore indicate the predicted range of light curves for the fourth and fifth runs, respectively. Thin horizontal lines show the depth expected to be reached by the Vera C. Rubin Observatory with LSSTCam with integration times of 30, 180, and 1800 s, as indicated in the top right-hand panel. The time axis is relative to the time of the NS–NS merger in the observer frame. Note that these light curves correspond to a lensed kilonova counterpart at a redshift of  $z = 1.6$ , therefore the most sensitive filter (*r*-band) corresponds to a rest-frame wavelength of  $\lambda_{\text{rest}} \simeq 230$  nm.



**Figure 10.** Predicted *r*-band light curves for kilonova counterparts to lensed NS–NS mergers, concentrating on the fifth GW run, and demonstrating the sensitivity to the chirp mass (left-hand panel) and mass ratio (right-hand panel) of lensed NS–NS mergers with (as examples) cocoon opening angle of  $30^\circ$  and a viewing angle of  $60^\circ$ . Kilonova counterparts to heavier lensed NS–NS mergers are brighter and fade more quickly than the counterparts to lighter lensed NS–NS mergers. Kilonova counterparts to less equal mass lensed NS–NS mergers are brighter and fade more slowly than the counterparts to more equal mass mergers. However, the duration of the most extreme mass ratio systems ( $q = 0.5$ ) is  $d < 2$  d, and thus shorter than the general population of fast transients discussed in Section 5.2. These characteristics are insensitive to the choice of opening and viewing angles. All other details of the figure are as stated in the caption to Fig. 9. The black curves in both panels are identical.

Nicholl et al. 2017). An estimate of the  $k$ -correction therefore requires models and assumptions about the kilonova physics. The Nicholl et al. (2021) models assume a multicomponent blackbody spectral energy distribution, which is inevitably overly simplistic given the complicated atomic structures of the  $r$ -process elements that comprise the kilonova ejecta. However, even sophisticated radiative transfer models are likely to be unreliable at line-rich UV wavelengths, due to the difficulty of computing the atomic structure of all relevant species and the approximations used to get around this – for example, using statistical properties of atoms rather than individual lines (Kasen et al. 2017), or specifying opacities rather than composition (Bulla 2019; Tanaka et al. 2020).

Importantly, the Nicholl et al. (2021) model fits well the observed luminosity of AT2017gfo (effectively at  $z = 0$ ) in the *Swift* UV bands (2000–3500 Å) at the most relevant time of  $t \simeq 1$  d. The observer-frame light curves of the 170817-like model should therefore be relatively trustworthy in all bands redder than *u* at  $z \simeq 1–2$ , but by  $z \sim 3$  the  $k$ -correction is not well-constrained bluewards of the *i*-band. Since  $z \simeq 3$  is the redshift out to which lensed NS merger candidates can be identified through their detection in the lower mass gap, we can at least be confident that our predictions for a 170817-like kilonova are robust in the *i*, *z*, *y*-bands over the entire redshift range of interest (and in *g* and *r* over much of that range). For the conservative model, the  $k$ -correction is less secure at all wavelengths/redshifts, since this model has not been calibrated directly against an observed kilonova. As a guide, we therefore use the Nicholl et al. (2021) models to estimate  $k$ -corrections 1 d post-merger in the *i*-band across the true redshift range  $\tilde{z} < z < 3$  for a source placed at  $\tilde{z} = 0.1$  in low latency,

obtaining  $-1 \lesssim k \lesssim 0$ . This suggests that the  $k$ -corrections work in our favour and cause lensed kilonova to be a little brighter than if  $k$ -corrections are ignored. Therefore, despite the large uncertainties, the unknown true distance to lensed NS–NS appears unlikely to be a severe impediment to detecting their lensed kilonova counterparts.

#### 5.4 Outline observing strategy

It is unlikely that a wide-field survey telescope will be observing the relevant sky region to the required depth when a candidate lensed NS–NS merger is detected by the GW detectors and announced as a public alert by the LVK collaboration. Target of Opportunity (ToO) optical observations will therefore need to be triggered in response to such a detection, with the upcoming Vera C. Rubin Observatory being the most powerful facility in the mid-2020s when detection rates approach one per year (Section 4.1).

A sensitivity of  $\simeq 26$ – $27$ th magnitude is required to probe a broad range of kilonova models (Sections 5.2 and 5.3). As previously sketched by Smith et al. (2019a), ToO observations that aim to detect lensed kilonova counterparts to candidate lensed NS–NS mergers (magnified into the mass gap between NS and BH) need to be much deeper than those envisaged to follow-up NS–NS mergers that are not lensed (Andreoni et al. 2022). The light curves in Fig. 9 indicate that up to  $\simeq 30$  min per pointing per epoch per filter will be required, depending on the expected magnification of the source. Such deep observations therefore motivate selecting well localized candidate lensed NS–NS mergers. Moreover, the combination of faint flux levels and distinctive short duration of lensed kilonovae motivates prioritizing sensitivity in a single filter over attempting to measure colours from observations through multiple filters.

Adopting nominal parameters of five Rubin/LSSTCam pointings ( $9.6 \text{ deg}^2$  per pointing), three epochs, one filter (preferably  $r$ -band), and  $\lesssim 30$  min integration per pointing per epoch gives a ballpark estimate of  $\lesssim 7.5$  h of observing time per candidate lensed NS–NS – i.e.  $\lesssim 0.25$  per cent of a year’s observing time (assuming  $3000 \text{ h yr}^{-1}$ ). This observing strategy will be sensitive to a broad range of blue and red kilonova models across the first two observer frame days post-GW detection. Two epochs on these time-scales will aim to detect the rapid fading of the rest-frame UV ( $\lambda_{\text{rest}} \simeq 230 \text{ nm}$  for the observer’s  $r$ -band) lensed kilonova emission. Crucially, the cosmological time dilation and redshifting renders lensed kilonovae the fastest fading known transients at  $\gtrsim 1 \text{ mag d}^{-1}$ . The third epoch (say)  $\gtrsim 4$  d post-GW detection would provide a post-hoc template image.

Even a modest factor 2 reduction in the amount of observing time per candidate will help us to increase the number of candidates that can be followed up and/or reduce the observing time requirement. In particular, release of additional mass information in low latency by the LVK collaboration would help us to constrain our lensed kilonova light-curve predictions. For example, the light curves of objects with lower mass ratios are brighter and fade slower than for higher mass ratios (Fig. 10; note the mass ratio  $q$  is lensing invariant and thus low latency posteriors on this parameter are straightforward to interpret). If available, low latency posteriors on  $\mathcal{M}$  would also help considerably, as highlighted by Bianconi et al. (2022) – two of the three mass gap detections ( $p_{\text{gap}} > 0.94$ ) followed up in the third GW run were later identified as having chirp masses of  $\mathcal{M} \simeq 8 M_{\odot}$ . Low latency chirp mass posteriors would therefore enable a cleaner selection of candidates for expensive follow-up observations. Such information will be similarly helpful for ToO follow-up observations with 4-m class telescopes including, CFHT/MegaCam and Blanco/DECam and dedicated GW follow-

up facilities such as BlackGEM and GOTO. The latter facilities also typically have a very broad  $VR$ -band (or similar) filter, that effectively enhances ‘light collecting power’ by a factor two over (say) the  $r$ -band. This can help us to offset the smaller aperture size of these facilities when searching for faint fast lensed transients.

In summary, rapid response deep ToO observations with the Vera C. Rubin Observatory are capable of detecting a lensed kilonova counterpart to a lensed NS–NS merger in the mid-2020s. Designing these observations to be useful for other science interests and gaining access to mass-related information in low latency will both go a long way to optimizing the observations for mutual benefit across the community. GW detections in the mass gap are also scientifically compelling independent of the lensing interpretation. The observations outlined here therefore provide a physically well-motivated approach to exploring the mass gap electromagnetically, and will inform non-lensing interpretations of objects discovered in this region of parameter space.

## 6 SUMMARY

We have investigated gravitational lensing of GWs in the context of the mass functions of compact object remnants of stellar evolution, and of strong gravitational lenses. We developed a new analytic lens magnification-based approach to predicting the detectable rates of lensed GWs with the current generation of detectors, and transformed these predictions into estimates of the arrival time difference between lensed images of the distant GW sources. This enabled us to compare different methods for selecting candidate lensed GWs, and to identify selecting candidate lensed NS–NS mergers that have been magnified into the mass gap between NS and BHs as the most efficient method based on data that are publicly available immediately following the GW detection. Moreover, rapid selection of candidates lends itself to electromagnetic follow-up observations and localization of candidate lensed NS–NS merger to a lensed host galaxy. This is the most direct way to test the interpretation that ‘mass gap’ events have been gravitationally lensed, to shrink the sky localization uncertainties of candidate lensed GWs to the angular scale of gravitational lensing, and thus to potentially achieve a secure first discovery of gravitational lensing of gravitational waves. We therefore combined our lensing predictions with recent models of kilonova light curves from Nicholl et al. (2021) to predict the optical light curves of lensed kilonova counterparts to lensed NS–NS mergers in the filters that will be used by the upcoming Vera C. Rubin Observatory. Our main conclusions are as follows:

#### Lensed GW detection becomes likely from 2023 onwards.

The predicted rates of lensed BH–BH and lensed NS–NS mergers available to be detected will encroach on one per year during the fourth and fifth GW runs, respectively. The predicted rates are relatively insensitive to the details of the mass functions we employed for the source populations, however significant uncertainty is due to the underlying local comoving merger rate density, with lensed rates of  $\simeq 1$ – $10$  and  $\simeq 0.01$ – $1$  per year for lensed BH–BH and lensed NS–NS mergers respectively from the mid-2020s.

**Detection of lensed BH–BH mergers in low latency is challenging.** Although more numerous than lensed NS–NS mergers, lensed BH–BH mergers are challenging to identify in low latency because their masses inferred assuming no magnification is at play ( $\mu = 1$ ) overlap with the population that is not lensed. In other words, in the mass–distance plane essentially every BH–BH detection is a *candidate* lensed BH–BH. This is due to the steep slope of the BH mass function that has emerged from the catalogue of GW detections to date. This does not imply interpreting all BH–BH detections as

being lensed – the relative rate of lensed detections is  $\simeq 1:1000$  (Table 2).

#### Multiple detections of lensed BH–BH mergers are challenging.

Detection strategies based on identifying more than one lensed image of a BH–BH are also challenging. This is because lensed BH–BH mergers are predicted to be dominated by relatively low lens magnifications of  $\mu \simeq 2\text{--}10$ . Consequently, the arrival time difference between lensed images of distant BH–BH mergers is typically long compared with the hour-long time-scales on which GW detectors operate continuously without interruption. Moreover at these magnifications, lenses with relatively flat density profiles such as groups and clusters of galaxies are prone to only produce one lensed image. These issues reduce the efficiency of this selection method and motivate longer GW detector duty cycles that maximize overlap between different GW detectors and detailed work on the relationship between magnification and multiplicity across the lens mass function.

#### Lensed NS–NS mergers are magnified into the mass gap.

Selection of candidate lensed NS–NS mergers in the mass–distance plane in low latency is relatively efficient because  $>50$  per cent of them are predicted to have been magnified into the mass gap between NS and BH. Selection of candidate lensed NS–NS mergers on the  $p_{\text{gap}}$  measure that the LVK collaboration began releasing in low latency in their third run is therefore a useful tool. However, cleaner selection of candidates would be achieved if mass and/or mass ratio posteriors could be made available in low latency (Bianconi et al. 2022).

**Lensed kilonova counterparts are detectable.** Combining our lensing models with Nicholl et al. (2021) kilonova models, we show that lensed kilonova counterparts peak at  $\simeq 22\text{--}26$ th magnitude, depending on the level of lens magnification and whether a 170817-like or more conservative (redder) kilonova model is adopted. Whilst they are faint and fade quickly (the fastest fading transients, with a duration of  $d < 2d$ ), the true redshift of the lensed kilonovae ( $z \simeq 1.6$ ) helps us to stretch the light curves and make detection of the rest-frame UV emission feasible. We concentrate on deep target of opportunity observations with the upcoming Vera C. Rubin Observatory, and show that three epochs of  $r$ -band observations, spanning a sky localization region of  $\Omega_{90} \simeq 50 \text{ deg}^2$ , would be sensitive to a broad range of kilonova physics, and thus enable detection of a lensed kilonova counterpart to a candidate lensed NS–NS. This equates to  $\lesssim 0.25$  per cent of a year’s observing time per candidate lensed NS–NS. The most strongly magnified lensed kilonovae will also be detectable with smaller aperture telescopes, which will also benefit from using a broad  $VR$ -band (or similar) filter.

#### Arrival time differences for lensed NS–NS/kilonovae are short.

Multiple detections of lensed NS–NS mergers are less challenging than for lensed BH–BH mergers because the former are more strongly magnified than the latter. In particular, the arrival time difference between lensed images of distant NS–NS mergers (sub-day) is typically comparable with the duty cycle GW detectors (see above). It is therefore plausible that the second lensed image of a distant NS–NS merger is detected and the lensed kilonova counterpart starts to brighten during the optical follow-up observations that were triggered in response to the first image. This opens up the possibility of measuring the rest-frame UV light curve of a kilonova at a true redshift of  $z \simeq 1\text{--}2$  simultaneously with the NS–NS merger.

#### ACKNOWLEDGEMENTS

GPS dedicates this work to the memory of Derek Fry, and thanks Simon Birrer, Riccardo Buscicchio, Cressida Cleland, Tom Collett, Jeff Cooke, Ben Gompertz, Denis Martynov, Chris Moore, Sam Oates,

Armin Rest, and Jielai Zhang for help and stimulating discussions. GPS acknowledges support from The Royal Society and the Leverhulme Trust. GPS and MB acknowledge support from the Science and Technology Facilities Council (grant numbers ST/N021702/1 and ST/S006141/1). GM acknowledges funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No MARACHAS - DLV-896778. MN is supported by the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement No. 948381) and by a Fellowship from the Alan Turing Institute. DR acknowledges a PhD studentship from the Science and Technology Facilities Council. MJ is supported by the United Kingdom Research and Innovation (UKRI) Future Leaders Fellowship (FLF), ‘Using Cosmic Beasts to uncover the Nature of Dark Matter’ (grant number MR/S017216/1).

This research has made use of data from the Gravitational Wave Open Science Center (gw-openscience.org), a service of LIGO Laboratory, the LIGO Scientific Collaboration, the Virgo Collaboration, and KAGRA. LIGO Laboratory and Advanced LIGO are funded by the United States National Science Foundation (NSF) as well as the Science and Technology Facilities Council (STFC) of the United Kingdom, the Max-Planck-Society (MPS), and the State of Niedersachsen/Germany for support of the construction of Advanced LIGO and construction and operation of the GEO600 detector. Additional support for Advanced LIGO was provided by the Australian Research Council. Virgo is funded, through the European Gravitational Observatory (EGO), by the French Centre National de Recherche Scientifique (CNRS), the Italian Istituto Nazionale di Fisica Nucleare (INFN) and the Dutch Nikhef, with contributions by institutions from Belgium, Germany, Greece, Hungary, Ireland, Japan, Monaco, Poland, Portugal, Spain. The construction and operation of KAGRA are funded by Ministry of Education, Culture, Sports, Science and Technology (MEXT), and Japan Society for the Promotion of Science (JSPS), National Research Foundation (NRF) and Ministry of Science and ICT (MSIT) in Korea, Academia Sinica (AS) and the Ministry of Science and Technology (MoST) in Taiwan.

#### DATA AVAILABILITY

The data presented in this article are available upon reasonable requests to the lead author.

#### REFERENCES

- Abbott R. et al., 2020, *ApJ*, 900, L13  
 Abbott R. et al., 2021a, *SoftwareX*, 13, 100658  
 Abbott R. et al., 2021b, *ApJ*, 913, L7  
 Abbott B. P. et al., 2017a, *Phys. Rev. Lett.*, 119, 161101  
 Abbott B. P. et al., 2017b, *ApJ*, 848, L12  
 Abbott B. P., Abbott R., Abbott T. D., Abraham S., Acernese F., Ackley K., Adams C., Adhikari R. X., 2019, *Phys. Rev. X*, 9, 031040  
 Ade P. A. R. et al., 2016, *A&A*, 594, A13  
 Andreoni I. et al., 2022, *ApJS*, 260, 18  
 Arcavi I., 2018, *ApJ*, 855, L23  
 Arnett W. D., 1982, *ApJ*, 253, 785  
 Atek H. et al., 2015, *ApJ*, 800, 18  
 Bacon D. J., Goldberg D. M., Rowe B. T. P., Taylor A. N., 2006, *MNRAS*, 365, 414  
 Baker T., Trodden M., 2017, *Phys. Rev. D*, 95, 063512  
 Barnes J., Kasen D., Wu M.-R., Martínez-Pinedo G., 2016, *ApJ*, 829, 110  
 Bhatawdekar R., Conzelmann C. J., Margalef-Bentabol B., Duncan K., 2019, *MNRAS*, 486, 3805  
 Bianconi M. et al., 2022, preprint (arXiv:2204.12978)

- Bianconi M., Buscicchio R., Smith G. P., McGee S. L., Haines C. P., Finoguenov A., Babul A., 2021, *ApJ*, 911, 136
- Binney J., Tremaine S., 1987, *Galactic Dynamics*, Princeton Series. Princeton Univ. Press, Princeton
- Blandford R., Narayan R., 1986, *ApJ*, 310, 568
- Broadhurst T., Diego J. M., Smoot George I., 2018, preprint (arXiv:1802.05273)
- Broadhurst T., Diego J. M., Smoot George F. I., 2019, preprint (arXiv:1901.03190)
- Broadhurst T., Diego J. M., Smoot G. F., 2022, preprint (arXiv:2202.05861)
- Bulla M., 2019, *MNRAS*, 489, 5037
- Buscicchio R., Moore C. J., Pratten G., Schmidt P., Vecchio A., 2020a, *Phys. Rev. D*, 102, 081501
- Buscicchio R., Moore C. J., Pratten G., Schmidt P., Bianconi M., Vecchio A., 2020b, *Phys. Rev. Lett.*, 125, 141102
- Callister T., Fishbach M., Holz D., Farr W., 2020, *ApJL*, 896, 32
- Cao S., Qi J., Cao Z., Biesiada M., Li J., Pan Y., Zhu Z.-H., 2019, *Sci. Rep.*, 9, 11608
- Chen H.-Y., Holz D. E., Miller J., Evans M., Vitale S., Creighton J., 2021, *Class. Quantum Gravity*, 38, 055010
- Coe D. et al., 2019, *ApJ*, 884, 85
- Collett T. E., Bacon D., 2017, *Phys. Rev. Lett.*, 118, 091101
- Congdon A. B., Keeton C. R., Nordgren C. E., 2008, *MNRAS*, 389, 398
- Coughlin M. W., Dietrich T., Margalit B., Metzger B. D., 2019, *MNRAS*, 489, L91
- Dahle H., Gladders M. D., Sharon K., Bayliss M. B., Rigby J. R., 2015, *ApJ*, 813, 67
- Dai L., Zackay B., Venumadhav T., Roulet J., Zaldarriaga M., 2020, preprint (arXiv:2007.12709)
- Darbha S., Kasen D., 2020, *ApJ*, 897, 150
- Diego J. M., 2019, *A&A*, 625, A84
- Diego J. M., 2020, *Phys. Rev. D*, 101, 123512
- Diego J. M., Hannuksela O. A., Kelly P. L., Pagano G., Broadhurst T., Kim K., Li T. G. F., Smoot G. F., 2019, *A&A*, 627, A130
- Diego J. M., Broadhurst T., Smoot G. F., 2021, *Phys. Rev. D*, 104, 103529
- Dietrich T., Ujevic M., 2017, *Class. Quantum Gravity*, 34, 105014
- Dominik M., Belczynski K., Fryer C., Holz D. E., Berti E., Bulik T., Mandel I., O'Shaughnessy R., 2013, *ApJ*, 779, 72
- Ellis R., Santos M. R., Kneib J., Kuijken K., 2001, *ApJ*, 560, L119
- Fan X.-L., Liao K., Biesiada M., Piórkowska-Kurpas A., Zhu Z.-H., 2017, *Phys. Rev. Lett.*, 118, 091102
- Farrow N., Zhu X.-J., Thrane E., 2019, *ApJ*, 876, 18
- Finke A., Foffa S., Iacovelli F., Maggiore M., Mancarella M., 2021, *Phys. Rev. D*, 104, 084057
- Fischer T. C. et al., 2019, *ApJ*, 875, 102
- Fohlmeister J., Kochanek C. S., Falco E. E., Morgan C. W., Wambsganss J., 2008, *ApJ*, 676, 761
- Fohlmeister J., Kochanek C. S., Falco E. E., Wambsganss J., Oguri M., Dai X., 2013, *ApJ*, 764, 186
- Fox C., Mahler G., Sharon K., Remolina González J. D., 2022, *ApJ*, 928, 87
- Franx M., Illingworth G. D., Kelson D. D., van Dokkum P. G., Tran K., 1997, *ApJ*, 486, L75
- Gavazzi R., Fort B., Mellier Y., Pelló R., Dantel-Fort M., 2003, *A&A*, 403, 11
- Goldberg D. M., Bacon D. J., 2005, *ApJ*, 619, 741
- Goobar A. et al., 2017, *Science*, 356, 291
- Goobar A. et al., 2022, preprint (arXiv:2211.00656)
- Goyal S., Haris K., Mehta A. K., Ajith P., 2021, *Phys. Rev. D*, 103, 024038
- Graham M. J. et al., 2020, *Phys. Rev. Lett.*, 124, 251102
- Guillochon J., Nicholl M., Villar V. A., Mockler B., Narayan G., Mandel K. S., Berger E., Williams P. K. G., 2018, *ApJS*, 236, 6
- Hannuksela O. A., Haris K., Ng K. K. Y., Kumar S., Mehta A. K., Keitel D., Li T. G. F., Ajith P., 2019, *ApJ*, 874, L2
- Hannuksela O. A., Collett T. E., Çalişkan M., Li T. G. F., 2020, *MNRAS*, 498, 3395
- Haris K., Mehta A. K., Kumar S., Venumadhav T., Ajith P., 2018, preprint (arXiv:1807.07062)
- Hilbert S., White S. D. M., Hartlap J., Schneider P., 2008, *MNRAS*, 386, 1845
- Hou S., Fan X.-L., Liao K., Zhu Z.-H., 2020, *Phys. Rev. D*, 101, 064011
- Ishigaki M., Kawamata R., Ouchi M., Oguri M., Shimasaku K., Ono Y., 2018, *ApJ*, 854, 73
- Jauzac M. et al., 2016, *MNRAS*, 457, 2029
- Kasen D., Metzger B., Barnes J., Quataert E., Ramirez-Ruiz E., 2017, *Nature*, 551, 80
- Kelly P. L. et al., 2015, *Science*, 347, 1123
- Kelly P. L. et al., 2016, *ApJ*, 819, L8
- Kneib J., Ellis R. S., Santos M. R., Richard J., 2004, *ApJ*, 607, 697
- Kochanek C. S., 2002, *ApJ*, 578, 25
- Koopmans L. V. E., Treu T., Bolton A. S., Burles S., Moustakas L. A., 2006, *ApJ*, 649, 599
- Korobkin O., Rosswog S., Arcones A., Winteler C., 2012, *MNRAS*, 426, 1940
- Lai K.-H., Hannuksela O. A., Herrera-Martín A., Diego J. M., Broadhurst T., Li T. G. F., 2018, *Phys. Rev. D*, 98, 083005
- Lasky P. D., Fluke C. J., 2009, *MNRAS*, 396, 2257
- Li Y., Fan X., Gou L., 2019, *ApJ*, 873, 37
- Liao K., Fan X.-L., Ding X.-H., Biesiada M., Zhu Z.-H., 2017, *Nat. Commun.*, 8, 1148
- Li S.-S., Mao S., Zhao Y., Lu Y., 2018, *MNRAS*, 476, 2220
- Liu X., Magaña Hernandez I., Creighton J., 2021, *ApJ*, 908, 97
- Lynds R., Petrosian V., 1989, *ApJ*, 336, 1
- Madau P., Dickinson M., 2014, *ARA&A*, 52, 415
- Margutti R., Chornock R., 2021, *ARA&A*, 59, 155
- Martynov D. V. et al., 2016, *Phys. Rev. D*, 93, 112004
- McIsaac C., Keitel D., Collett T., Harry I., Mozzon S., Edy O., Bacon D., 2020, *Phys. Rev. D*, 102, 084031
- McLeod D. J., McLure R. J., Dunlop J. S., 2016, *MNRAS*, 459, 3812
- Meena A. K., Bagla J. S., 2020, *MNRAS*, 492, 1127
- Mellier Y., Fort B., Soucaill G., Mathez G., Cailloux M., 1991, *ApJ*, 380, 334
- Metzger B. D., Thompson T. A., Quataert E., 2018, *ApJ*, 856, 101
- Millon M. et al., 2020, *A&A*, 640, A105
- Mishra A., Meena A. K., More A., Bose S., Bagla J. S., 2021, *MNRAS*, 508, 4869
- More A., Suyu S. H., Oguri M., More S., Lee C.-H., 2017, *ApJ*, 835, L25
- Mukherjee S., Silk J., 2021, *MNRAS*, 506, 3977
- Mukherjee S., Wandelt B. D., Silk J., 2020a, *Phys. Rev. D*, 101, 103509
- Mukherjee S., Wandelt B. D., Silk J., 2020b, *MNRAS*, 494, 1956
- Mukherjee S., Broadhurst T., Diego J. M., Silk J., Smoot G. F., 2021, *MNRAS*, 506, 3751
- Nakamura T. T., 1998, *Phys. Rev. Lett.*, 80, 1138
- Ng K. K. Y., Wong K. W. K., Broadhurst T., Li T. G. F., 2018, *Phys. Rev. D*, 97, 023012
- Nicholl M. et al., 2017, *ApJ*, 848, L18
- Nicholl M., Margalit B., Schmidt P., Smith G. P., Ridley E. J., Nuttall J., 2021, *MNRAS*, 505, 3016
- Oguri M., 2018, *MNRAS*, 480, 3842
- Oguri M., 2019, *Rep. Progr. Phys.*, 82, 126901
- Oguri M., Takahashi R., 2020, *ApJ*, 901, 58
- Ohanian H. C., 1974, *Int. J. Theoret. Phys.*, 9, 425
- Perley D. A. et al., 2020, *ApJ*, 904, 35
- Petrov P. et al., 2022, *ApJ*, 924, 54
- Piro A. L., Kollmeier J. A., 2018, *ApJ*, 855, 103
- Richard J. et al., 2010, *MNRAS*, 404, 325
- Robertson A., Smith G. P., Massey R., Eke V., Jauzac M., Bianconi M., Ryczanowski D., 2020, *MNRAS*, 495, 3727
- Rodney S. A. et al., 2016, *ApJ*, 820, 50
- Ryczanowski D., Smith G. P., Bianconi M., Massey R., Robertson A., Jauzac M., 2020, *MNRAS*, 495, 1666
- Ryczanowski D., Smith G. P., Bianconi M., McGee S., Robertson A., Massey R., Jauzac M., 2022, *MNRAS*, preprint (arXiv:2204.12984)
- Sand D. J., Treu T., Smith G. P., Ellis R. S., 2004, *ApJ*, 604, 88
- Santoliquido F., Mapelli M., Giacobbo N., Bouffanais Y., Artale M. C., 2021, *MNRAS*, 502, 4877
- Schneider P., 1985, *A&A*, 143, 413

- Schneider P., Ehlers J., Falco E. E., 1992, *Gravitational Lenses*, XIV. Springer-Verlag, Berlin, Heidelberg, New York. 560pp, 112 figs. Also Astronomy and Astrophysics Library
- Sereno M., Jetzer P., Sesana A., Volonteri M., 2011, *MNRAS*, 415, 2773
- Shapiro I. I., 1964, *Phys. Rev. Lett.*, 13, 789
- Sharon K. et al., 2020, *ApJS*, 247, 12
- Singer L. P., Goldstein D. A., Bloom J. S., 2019, preprint (arXiv:1910.03601)
- Smith G. P. et al., 2009, *ApJ*, 707, L163
- Smith G. P. et al., 2019b, *MNRAS*, 485, 5180
- Smith G. P., Kneib J., Ebeling H., Czoske O., Smail I., 2001, *ApJ*, 552, 493
- Smith G. P., Jauzac M., Veitch J., Farr W. M., Massey R., Richard J., 2018, *MNRAS*, 475, 3823
- Smith G. P., Robertson A., Bianconi M., Jauzac M., 2019a, preprint (arXiv:1902.05140)
- Soucail G., Fort B., Mellier Y., Picat J. P., 1987, *A&A*, 172, L14
- Subramanian K., Cowling S. A., 1986, *MNRAS*, 219, 333
- Tanaka M., Kato D., Gaigalas G., Kawaguchi K., 2020, *MNRAS*, 496, 1369
- The LIGO Scientific Collaboration, 2021, *ApJ*, 923, 14
- The LIGO Scientific Collaboration, The Virgo Collaboration The KAGRA Scientific Collaboration, 2021, preprint (arXiv:2111.03634)
- Tinker J., Kravtsov A. V., Klypin A., Abazajian K., Warren M., Yepes G., Gottlöber S., Holz D. E., 2008, *ApJ*, 688, 709
- Treu T. et al., 2016, *ApJ*, 817, 60
- Turner E. L., Ostriker J. P., Gott J. R. I., 1984, *ApJ*, 284, 1
- Umetsu K., Zitrin A., Gruen D., Merten J., Donahue M., Postman M., 2016, *ApJ*, 821, 116
- Urrutia J., Vaskonen V., 2022, *MNRAS*, 509, 1358
- Villar V. A. et al., 2017, *ApJ*, 851, L21
- Walsh D., Carswell R. F., Weymann R. J., 1979, *Nature*, 279, 381
- Wang Y., Stebbins A., Turner E. L., 1996, *Phys. Rev. Lett.*, 77, 2875
- Wang J.-S., Herrera-Martín A., Hu Y.-M., 2021, *Phys. Rev. D*, 104, 083515
- Wempe E., Koopmans L. V. E., Wierda A. R. A. C., Akseli Hannuksela O., van den Broeck C., 2022, preprint (arXiv:2204.08732)
- Wierda A. R. A. C., Wempe E., Hannuksela O. A., Koopmans L. V. E., Van Den Broeck C., 2021, *ApJ*, 921, 154
- Witt H. J., Mao S., Keeton C. R., 2000, *ApJ*, 544, 98
- Xu F., Ezquiaga J. M., Holz D. E., 2022, *ApJ*, 929, 9
- Yang L., Wu S., Liao K., Ding X., You Z., Cao Z., Biesiada M., Zhu Z.-H., 2022, *MNRAS*, 509, 3772

## APPENDIX A: GRAVITATIONAL TIME DELAY

We consider lenses that are approximately circular in projection on the sky, with a density profile local to a gravitationally lensed pair of images of the same source of:

$$\kappa(x) = \langle \kappa \rangle x^{-\eta}, \quad (\text{A1})$$

where  $\kappa$  is defined in equation (4),  $x = \theta/\theta_E$  is the angle from the centre of the lens in units of the Einstein radius,  $\langle \kappa \rangle$  is the density of the lens in the annulus that is bound by the location of the image pair, which for an axisymmetric lens is the convergence at the Einstein radius ( $\kappa_E$ ), and  $0 \leq \eta \leq 1$ . The arrival time difference between the images is conventionally phrased in terms of the image positions via  $\Delta\theta$ , the width of the annulus referred to above, with the dependence on  $\Delta\theta$  differing between different types of image pairs. For singular lenses that form just two images (by virtue of their pseudo-caustic),  $\Delta t_{\text{pseu}} \propto \Delta\theta$  (e.g. Kochanek 2002), whilst for image pairs formed at fold and cusp catastrophes of non-singular lenses,  $\Delta t_{\text{fold}} \propto \Delta\theta^3$  and  $\Delta t_{\text{cusp}} \propto \Delta\theta^4$ , respectively (e.g. Congdon, Keeton & Nordgren 2008). Here we consider  $\Delta t_{\text{pseu}}$  and  $\Delta t_{\text{fold}}$ , as they bracket the observed behaviour of lenses with measured time delays (Section 2.5). In the following sections, we derive expressions for  $\Delta t_{\text{pseu}}$  and  $\Delta t_{\text{fold}}$  in terms of lens magnification,  $\mu$ , and lens structure, i.e.  $\eta$  and  $\langle \kappa \rangle$ .

### A1 Arrival time difference for the pseudo-caustic

Kochanek (2002) showed that the arrival time difference for an image pair formed by a pseudo-caustic is given by

$$\Delta t_{\text{pseu}} = 2\Delta t_{\text{SIS}} \left[ 1 - \langle \kappa \rangle - \frac{1 + (1 - \eta)\langle \kappa \rangle}{12} \left( \frac{\Delta\theta}{\langle \theta \rangle} \right)^2 + \mathcal{O} \left( \left( \frac{\Delta\theta}{\langle \theta \rangle} \right)^4 \right) \right], \quad (\text{A2})$$

where  $\Delta t_{\text{SIS}}$  is the arrival time difference between two images formed by a singular isothermal sphere (SIS) lens. Since  $1 \geq \eta \geq 0$  and  $0.5 \leq \langle \kappa \rangle \leq 1$  (Fig. 1), it is clear that the  $(\Delta\theta/\langle \theta \rangle)^2$  term contributes no more than  $\simeq 10$  per cent to  $\Delta t_{\text{pseu}}$ , even for  $\Delta\theta/\langle \theta \rangle \simeq 1$ . Therefore,  $\Delta t_{\text{pseu}}$  is mainly influenced by the density of the lens,  $\langle \kappa \rangle$ :

$$\Delta t_{\text{pseu}} \simeq 2\Delta t_{\text{SIS}} [1 - \langle \kappa \rangle]. \quad (\text{A3})$$

Therefore,  $\Delta t_{\text{pseu}} \rightarrow 0$  for a flat density profile ( $\eta \rightarrow 0$ ,  $\langle \kappa \rangle \rightarrow 1$ ), and we recover  $\Delta t_{\text{pseu}} = \Delta t_{\text{SIS}}$  for an isothermal lens ( $\eta = 1$ ,  $\langle \kappa \rangle = 0.5$ ).

The arrival time difference for a SIS lens is typically written in terms of the positions of the two images,  $\theta_+$  and  $\theta_-$ :

$$c\Delta t_{\text{SIS}} = \frac{\mathcal{D}}{2} [\theta_+^2 - \theta_-^2] = \mathcal{D} \langle \theta \rangle \Delta\theta, \quad (\text{A4})$$

where  $\mathcal{D}$  is defined adjacent to equation (1). We now use equation (16) to re-write equation (A4) in terms of magnification. Specifically, the sum of the lens magnifications suffered by the image pair is  $\mu_p = 2\theta_E/\beta = 2\langle \theta \rangle/\beta$ , where  $\beta$  is the source position, and  $\theta_{\pm} = \beta \pm \theta_E$ , therefore  $\Delta\theta = 2\beta$ . Putting this together, we obtain

$$c\Delta t_{\text{SIS}} = \frac{4\mathcal{D} \langle \theta \rangle^2}{\mu_p}, \quad (\text{A5})$$

and thus

$$\Delta t_{\text{pseu}} \simeq \frac{8\mathcal{D} [1 - \langle \kappa \rangle] \langle \theta \rangle^2}{c\mu_p}, \quad (\text{A6})$$

We also write equation (A6) as a convenient scaling relation:

$$\frac{\Delta t_{\text{pseu}}}{92 \text{ d}} = \left[ \frac{\theta_E}{1''} \right]^2 \left[ \frac{\mu_p}{4} \right]^{-1} \left[ \frac{1 - \kappa_E}{0.5} \right] \left[ \frac{\mathcal{D}}{3.3 \text{ Gpc}} \right], \quad (\text{A7})$$

where  $\mathcal{D} = 3.3 \text{ Gpc}$  corresponds to  $z_L = 0.5$  and  $z_S = 1.6$ , and we have approximated  $\kappa_E = \langle \kappa \rangle$  and  $\theta_E = \langle \theta \rangle$ .

### A2 Image pairs formed by fold catastrophes

We start from Schneider, Ehlers & Falco (1992) expression for the arrival time difference (their equation 6.21a) between two fold images:

$$c\Delta t_{\text{fold}} = \frac{\mathcal{D} \Delta\theta^2}{6\mu_p \left| \tau_{11}^{(0)} \right|}, \quad (\text{A8})$$

where  $\tau_{11}^{(0)}$  is the partial second-order derivative of the Fermat potential evaluated at the critical point of the lens mapping, i.e. where  $\det(\mathcal{A}) = 0$  at the mid-point between the images. We then use Schneider et al. (1992)'s equation (6.20) to eliminate  $\Delta\theta$ , to obtain

$$c\Delta t_{\text{fold}} = \frac{2\mathcal{D}}{3\mu_p^3 \left| \tau_{11}^{(0)} \right|^3 \left| \tau_{222}^{(0)} \right|^2}. \quad (\text{A9})$$

Next, we address the dependence of  $\Delta t_{\text{fold}}$  on the structure of the lens. Taking  $\tau_{11}$  first, we obtain from equations (2) and (10):

$$\tau_{11} = 1 - \psi_{11} = \mathcal{A}_{11}, \quad (\text{A10})$$

where the subscripts on  $\mathcal{A}$  denote the elements of the Jacobian matrix. Following Schneider et al. (1992), we can write this as

$$\tau_{11} = 2(1 - \kappa), \quad (\text{A11})$$

and then evaluate  $\tau_{11}$  at the mid-point of the image pair to obtain:

$$\left| \tau_{11}^{(0)} \right| = 2 |1 - \langle \kappa \rangle|. \quad (\text{A12})$$

Turning to  $\tau_{222}$ , we differentiate equation (2) to obtain:

$$\tau_{222} = -\psi_{222}, \quad (\text{A13})$$

which shows, in line with many previous works (e.g. Goldberg & Bacon 2005; Bacon et al. 2006; Lasky & Fluke 2009), that second order lensing (flexion) depends on derivatives of the convergence  $\kappa$ .

We therefore adopt  $\tau_{222} = \partial \kappa / \partial \theta$  and write  $\left| \tau_{222}^{(0)} \right|$  as:

$$\left| \tau_{222}^{(0)} \right| = \left| \frac{d\kappa}{d\theta} \right|_{(\theta)}, \quad (\text{A14})$$

where the subscript denotes evaluation of the derivative at  $\theta = \langle \theta \rangle$ . Substituting equation (A1) into equation (A14) then yields

$$\left| \tau_{222}^{(0)} \right| = \frac{\eta \langle \kappa \rangle}{\langle \theta \rangle}. \quad (\text{A15})$$

Substituting equations (A12) and (A15) into equation (A9) we obtain:

$$\Delta t_{\text{fold}} = \frac{\mathcal{D} \langle \theta \rangle^2}{12 c \mu_p^3 \eta^2 \langle \kappa \rangle^2 [1 - \langle \kappa \rangle]^3}, \quad (\text{A16})$$

which can be phrased as a convenient scaling relation:

$$\frac{\Delta t_{\text{fold}}}{3.9 \text{ d}} = \left[ \frac{\theta_E}{1''} \right]^2 \left[ \frac{\mu_p}{4} \right]^{-3} \left[ \frac{\eta_E}{1} \right]^{-2} \left[ \frac{\kappa_E}{0.5} \right]^{-2} \left[ \frac{1 - \kappa_E}{0.5} \right]^{-3} \times \left[ \frac{\mathcal{D}}{3.3 \text{ Gpc}} \right]. \quad (\text{A17})$$

This paper has been typeset from a  $\text{\TeX}/\text{\LaTeX}$  file prepared by the author.