A&A 625, A84 (2019)
https://doi.org/10.1051/0004-6361/201833670

Astronomy & Astrophysics

The Universe at extreme magnification

J. M. Diego

Instituto de Física de Cantabria (CSIC-UC), Avda. Los Castros s/n, 39005 Santander, Spain
e-mail: jdiego@ifca.unican.es

Received 18 June 2018 / Accepted 19 March 2019

ABSTRACT

Extreme magnifications of distant objects by factors of several thousand have recently become a reality. Small, very luminous compact objects, such as supernovae (SNe), giant stars at $z = 1-2$, Pop III stars at $z > 7$, and even gravitational waves (GWs) from merging binary black holes near caustics of gravitational lenses can be magnified many thousands or even tens of thousands of times thanks to their small size. We explore the probability of such extreme magnifications in a cosmological context and include the effect of microlenses near critical curves. We show how the presence of microlenses near the critical curve sets a limit on the maximum magnification. We use a combination of state of the art halo mass functions, high-resolution analytical models for the density profiles, and inverse ray tracing to estimate the probability of magnification near caustics. We estimate the rate of highly magnified events in the case of SNe, GWs, and very luminous stars including Pop III stars. Our findings reveal that future observations will increase the number of events at extreme magnifications, and will open the door not only to studying individual sources at cosmic distances, but also to constraining compact dark matter candidates.

Key words. gravitational lensing: micro – gravitational lensing: strong – stars: Population III – dark matter

1. Introduction

When observing distant objects with a given instrument, the limiting magnitude set by the experimental configuration defines a natural limit to the maximum distance, $D_{\rm max}$, at which an object with a given luminosity can be observed. If some of these objects are magnified by a factor μ, it is possible to observe a similar object farther away up to a distance $D_{\rm l}(z) = \sqrt{\mu} D_{\rm max}$ (ignoring k-corrections). For sufficiently large values of μ, relatively faint objects can be observed at cosmological distances. It is interesting to answer the question of what the limit on this relation is and how far we can see relatively faint objects that otherwise (i.e., without magnification) could not be observed. A similar question has been studied extensively in the context of lensed galaxies and to a lesser extent quasars (QSOs), but they are not the focus of this work. Instead, we focus on bright and very small objects that can be amplified by extreme magnification factors ($\mu > 1000$). The probability of having an event magnified by a large factor μ is well known to scale as μ^{-3} (see e.g., Lee & Paczynski 1990; Rauch 1991). In general, the probability of seeing a strongly lensed event depends on several factors: (i) the volumetric density of objects as a function of redshift, (ii) the volume element that depends on the redshift, and (iii) the probability of intersecting a gravitational lens. The maximum magnification at which an object can be observed depends on the redshifts of the lens and background source, the mass and concentration of the lens, and the size of the background sources. Maximum magnification is obtained when a background source is touching a caustic (i.e., when a source or radius R is at a distance R from the caustic). The smaller the source, the closer it can get to the caustic, so smaller sources can be magnified by larger factors. Stars, for instance, could in principle be magnified by factors of several million when touching a galaxy cluster caustic (Miralda-Escude 1991). Diego et al. (2018) discuss how this maximum magnification gets reduced in the presence of microlenses (see also Kayser et al. 1986). The combined effect of macro- and microlensing can result in large magnification factors even relatively far away from the critical curve of the macromodel (Diego et al. 2018). However, even with extreme magnifications of several million, a normal star like the Sun at redshift $z > 1$ would still fall below the detection limit of the most powerful current telescopes. The background source needs to be small, and also very bright.

In Kelly et al. (2018), one such very bright object at $z > 1$ was found to lie very close to a powerful caustic. This star, nicknamed Icarus, holds the record for the most distant star and most extreme magnification ever observed thanks to gravitational lensing. Kelly et al. (2018) describes Icarus as a $z = 1.49$ giant star that is being magnified by the combined effect of a powerful lens (the galaxy cluster MACS J1149) and a microlens. Both the caustic of the microlens and the background star happen to be aligned near the caustic of the cluster resulting in a very high magnification of a few thousand times. This type of observation is the first of its kind and raises the question of how likely this kind of alignment may be. Additional events may have been observed in Rodney et al. (2018) and more recently in Chen et al. (2019). In Diego et al. (2018), the authors discuss the interesting possibility of using observations like this one to constrain the amount of dark matter in the form of compact microlenses (see also Oguri et al. 2018, where an actual constraint is derived based on Icarus). The mass range that can be constrained this way fills the gap of low to intermediate masses for primordial black holes (PBHs). PBHs have been proposed as one of the candidates for dark matter (or at least to account for a fraction of it). PBHs have been proposed also to explain the observation of the relatively abundant low-frequency Laser Interferometer Gravitational-Wave Observatory (LIGO) events. In this work we estimate the probability of observing

luminous stars at $z > 1$, and also discuss other very compact but intrinsically energetic phenomena such as supernovae (SNe) or gravitational waves (GWs).

In order to observe distant gravitationally lensed objects, large magnification factors that compensate for the increase in luminosity distance are needed. The lower probability of lensing is partially compensated by the larger volume, which grows approximately as $(1 + z)^3$ between $z \approx 0.3$ and $z \approx 1.3$ (and more slowly at higher redshifts). More precisely, the volume per redshift interval peaks at $z \approx 2.5$ (with about 55 Gpc^3 in a redshift interval of thickness $\Delta z = 0.1$). Regarding the probability of lensing, background objects at high redshifts have a higher probability of intersecting a gravitational lens along the line of sight. This probability is described by the optical depth. It is well known that the optical depth grows rapidly between $z = 0$ and $z = 1$. Between $z = 1$ and $z = 3$ it continues to grow although at a much slower pace. Beyond $z = 3$–5, the optical depth is still growing, but much more slowly, and beyond $z = 5$ it only grows by values ~1%, especially if we consider large magnification factors where the presence of caustics is required; however, these caustics are expected to be rare for lenses beyond $z \approx 3$. For sources (or events) that trace the star formation history, and at redshifts of the background source between 1 and 3, the low probability of magnification factors can be compensated by the larger volume element and increased volumetric density. In this redshift interval, the probability of seeing extremely magnified events is highest.

An example of extreme magnification is the above-mentioned Icarus event (Kelly et al. 2018) with a magnification factor exceeding 2000. The previous record holder (to the best of our knowledge) was OGLE-2008-BLG-279, for which Yee et al. (2009) infers a magnification of ≈1600. However, this event took place within the boundaries of our Galaxy, not at cosmological distances. Zackrisson et al. (2015) studies high magnifications in the context of compact globular clusters at high redshift as background objects, and uses N-body simulations to estimate the probability of lensing for a given magnification. They find that a survey with a limiting magnitude of 28 (in the AB system) could find approximately one primordial globular cluster per 100 deg^2 at $z > 7$ magnified by a factor $\mu > 300$. No observations of globular clusters at cosmological distances lensed by factors of hundreds have been reported yet.

Earlier work considered the probability of observing distant very luminous and compact objects, like SNe or QSOs, through gravitational lensing (Martel & Premadi 2008; Oguri & Marshall 2010) at high (but moderate in the context of this paper) magnifications (i.e., $\mu < 100$). In Broadhurst et al. (2018), the authors argue that the low-frequency GW events observed by LIGO could be interpreted as gravitationally lensed events originating at $z > 1$ instead of the implied low-redshift events. In order for lensing to work in this case, the intrinsic rate of coalescence events at high redshift needs to be over an order of magnitude higher than previously assumed (however this rate is largely unknown), and the magnifications involved need to be of order 10^2 or 10^3.

At redshifts $z > 3$ the probability of observing strongly lensed objects declines, unless the intrinsic volumetric density of background objects or their luminosity compensates for the reduction in lensing probability of the larger factor μ. In Windhorst et al. (2018), the authors propose using the James Webb Space Telescope (JWST) to observe the first Pop III stars at $z > 8$ through caustic crossing events and with magnifications greater than several thousand times. Pop III stars can be very luminous (millions of times the luminosity of the Sun) and are expected to be relatively abundant at $z > 10$. Windhorst et al. (2018) proposes monitoring massive gravitational lenses at $z \approx 0.4$–0.8 and estimates that, in the most optimistic scenario, one caustic crossing by a Pop III star could be observed every 3 years after targeting very massive galaxy clusters. Observing Pop III stars through caustic crossing would be very exciting since it would allow us to individually study some of the first stars. However, the question of how likely is it to observe this type of event in any surveyed random portion of the sky (not necessarily including a massive cluster) is still without an answer. We would need to take into account not only the most massive clusters, but all gravitational lenses that are capable of producing caustics at high redshift.

On cosmological scales, earlier work has looked at the probability of lensing at moderate magnifications ($\mu \approx 3$–10), either using analytical models (Turner et al. 1984; Fukugita et al. 1992; Kochanek 1996) or N-body simulations (Hilbert et al. 2007; Takahashi et al. 2011). Based on N-body simulations, Hilbert et al. (2007) compute the probability of lensing for magnifications between 3 and 10, and find that for a source at $z = 4$ the probability of being magnified by a factor larger than 10 is $\approx 3 \times 10^{-5}$ (in a subsequent work this probability was revised toward higher values after including the effect of a smooth distribution of baryons but not microlenses; see Hilbert et al. 2008). Takahashi et al. (2011) extends the calculation up to $\mu = 100$ and finds that for a source at $z = 5$ the probability of being magnified by a factor larger than 10 is $\approx 5 \times 10^{-5}$, in agreement with the result of Hilbert et al. (2007). At $\mu = 100$ and for a source at $z = 5$, the probability (per logarithmic interval) drops by an additional two orders of magnitude. These early studies ignore, however, the role of microlenses near critical curves that can be important, as discussed in Diego et al. (2018; but see also Venumadhav et al. 2017; Oguri et al. 2018).

In this work we compute the probability of magnification for bright and compact objects at $z > 1$, taking into account the role of microlenses. We rely on an accurate mass function to compute the number density of lenses and an analytical model to map the distribution of mass in individual halos. The analytical model allows us to reach spatial resolutions not available with N-body simulations, and explore the regime of high magnification near the caustics. We focus on very luminous stars (including Pop III stars) and SNe; all of them are powerful sources with relatively small radii and they are abundant enough. Their abundance results in non-negligible probabilities for one of them being very close to a caustic.

We also discuss briefly gravitational waves, although this particular case was explored in Broadhurst et al. (2018). Here we present in greater detail the lensing probability calculation of Broadhurst et al. (2018), and extend the discussion for the situation where microlenses are present.

We do not discuss the case of QSO lensing since magnifications in this case are not as extreme as in the case of stars, SNe, or GW. The magnifications involved in QSO lensing is normally modest reaching values of a few hundred μ at most (Walsh et al. 1979; Weymann et al. 1980; Wang et al. 2017). This limitation in the magnification follows from the much larger size of accretion disks, which are orders of magnitude larger than stars or SNe (except when QSOs are observed at X-ray wavelegths, in which case the size of the X-ray emitting region can be considerably smaller, and comparable to the size of a SNe; Mosquera et al. 2013). However, microlensing in QSO plays an important role and the abundant literature on this subject makes fundamental contributions to understanding the role of microlenses on the flux

variations of multiply imaged QSOs (see e.g., Wambsganss et al. 1990). The smaller size of stars or SNe compared with the size of QSO accretion disks makes microlensing even more relevant as the flux fluctuations can be considerably larger than in the case of QSO microlensing.

The paper is organized as follows. In Sect. 2 we give a brief description of the well-known lensing formalism that is relevant for the discussion in this work. In Sect. 3 we lay the groundwork for the computation of the lensing probability of a single halo. Section 4 extends the calculation to a cosmological context where all halos and redshifts are taken into account to compute the lensing optical depth with an emphasis on the optical depth at extreme magnification. In Sect. 5 we discuss the impact of microlenses on the lensing probability. In Sect. 6 we estimate the rates of events for super-luminous stars, Pop III stars, SNe, and GW. In Sect. 7 we discuss the results, and we conclude in Sect. 8. We adopt a flat cosmological model with $\Omega_m = 0.3$, $\Lambda = 0.7$, and $h = 0.7$. Using a model with slightly different cosmological values (like the cosmological model inferred by the *Planck* mission) has no impact on our conclusions since we present order of magnitude estimations of the expected rate of extremely magnified events.

2. Formalism

The basic equation in lensing is the lens equation that relates the real position in the sky of a background source β, the apparent position of the observed image θ, and the deflection angle produced by the lens at that position $\alpha(\theta, \Sigma)$,

$$\beta = \theta - \alpha(\theta, \Sigma), \tag{1}$$

where $\Sigma(\theta)$ is the surface mass density of the lens at the position θ. The vector deflection angle is obtained from the derivatives of the scalar lensing potential,

$$\psi(\theta) = \frac{4GD_{\rm l}D_{\rm ls}}{c^2 D_{\rm s}} \int {\rm d}^2\theta' \Sigma(\theta') \ln(|\theta - \theta'|), \tag{2}$$

where $D_{\rm l}$, $D_{\rm s}$, and $D_{\rm ls}$ are the angular diameter distances to the lens, to the source, and from the lens to the source, respectively. Given the surface mass density Σ, we can compute the lensing potential and its derivatives to obtain the deflection field. Once the deflection field is known, the magnification in the image plane can be derived as a combination of derivatives of the deflection field, and through the lens equation. The magnification in the source plane can be computed using the inverse ray shooting technique. In the image plane, regions where the magnification diverge are known as critical curves. The corresponding mapping of these curves into the source plane (through the lens equation) produce a different set of curves known as caustics. Near a critical curve we can perform a Taylor expansion on the lens equation, and after retaining only the first orders find that

$$|\beta - \beta_o| = \frac{|\theta - \theta_0|^2}{\Theta}, \tag{3}$$

where Θ is a constant (usually expressed in arcseconds) that depends on the lens model and redshifts of the lens and source, β_o is the position of the caustic, and θ_o is the corresponding position of the critical curve. As a consequence, near a critical curve the magnification falls as the inverse of the distance to the critical curve,

$$\mu = \frac{\mu_o}{|\theta - \theta_o|}, \tag{4}$$

where μ_o (usually expressed in arcseconds) is a constant that depends on the slope of the lens potential at the position of the critical curve (shallower potentials result in larger values of μ_o). Later in the paper we refer to the tangential magnification $\mu_{\rm t}$ and radial magnification $\mu_{\rm r}$ with the total magnification being the product of the two: $\mu = \mu_{\rm t}\mu_{\rm r}$. In the source plane, the magnification then decreases as

$$\mu = \frac{\mu_o/\sqrt{\Theta}}{\sqrt{|\beta - |\beta_o|}}. \tag{5}$$

Finally, the probability of having magnification higher than a certain value μ is equal to the probability of being at a distance to the caustic less than the corresponding $\Delta\beta = |\beta - \beta_o|$, i.e., $P(>\mu) \propto \Delta\beta \propto \mu^{-2}$. The differential probability is then ${\rm d}P(\mu)/{\rm d}\mu \propto \mu^{-3}$. From expression (5), the normalization factor $\mu_o/\sqrt{\Theta}$ can range between ~1 for small halos to 20–40 for the most massive clusters. For intermediate lenses with normalization $\mu_o/\sqrt{\Theta} \approx 10$, a source at a distance of 1 milliarcsec (typically a few parsecs at $z > 1$) would have a total magnification of ≈300.

3. Single lens

The above discussion is for regions that are close to a caustic. Here we explore the magnification properties of a single lens, but for all possible distances and magnifications. Even though most of the results in this section are widely known, it will be instructive for the remainder of the paper, in particular to understand the shape of the probability of lensing at intermediate magnifications and the difference between computing the magnification in the source plane with the inverse ray shooting method or the deprojection method.

We assume a cored elliptical Navarro, Frenk, and White (NFW) profile (Navarro et al. 1997) with ellipticity $e = 0.2$, where a NFW profile symmetric in $r = \sqrt{x^2 + y^2}$ is modified after substituting r with $r' = \sqrt{x^2/(1-e) + y^2(1+e)}$. In the central part of the NFW halo, we assume the profile to be $(r + r_{\rm c})^{-1}$ (and not the usual r^{-1}), where the core radius scales with mass as $r_{\rm c} = 20(M_{\rm vir}/10^{15})$ kpc. The core radius avoids unphysical divergences at $r = 0$, but also reproduces the profile of many massive lenses where the BCG usually exhibits a core, or flattened Sersic profile (i.e., with Sersic index $n \approx 1$). Adopting a standard NFW profile with no core should have minimal impact on our calculations, especially at high magnifications where most of the contribution to the lensing probability comes from regions near the tangential critical curves, that is, far from the core region. We should note that for small halos, the presence of the core can reduce their contribution to the optical depth since the central region of some of these halos may be subcritical when a core is present. The contribution of the small halos to the optical depth is discussed in more detail in the following section. For the virial radius, we adopt the scaling $R_{\rm vir} = 1.5 * (M_{\rm vir}/10^{15})^{1/3}$ Mpc, and concentration C given by the model in Prada et al. (2012):

$$C = 7.28 \left(\frac{M_{\rm vir}}{10^{12}}\right)^{-0.074}. \tag{6}$$

We note that in the previous equation, we ignore the evolution with redshift. Since most of the lenses concentrate around redshift 0.5, we expect this to be a small effect, especially for massive halos for which the dependency of the concentration with redshift is weaker (see e.g., Fig. 12 in Prada et al. 2012). The ellipticity $e = 0.2$ is typical of halos. Different ellipticities have

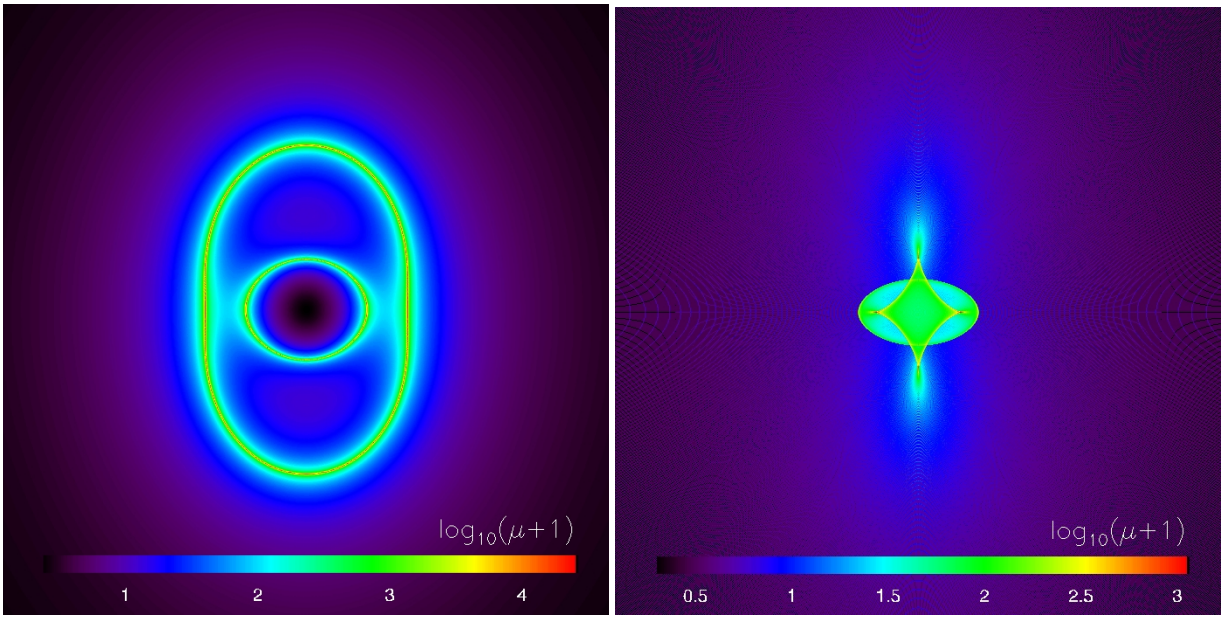

Fig. 1. Magnification pattern for a halo at $z_l = 0.4$ with mass $2 \times 10^{15}\,M_\odot$ and for a source at redshift $z_s = 4$. *Left panel*: logarithm of the magnification and critical curves in the image plane (field of view is 2.71 arcmin across). *Right panel*: logarithm of the magnification and caustics in the source plane (field of view is 1.355 arcmin across) obtained by inverse ray shooting.

little impact on the derived probability of lensing (provided the concentration remains constant). The concentration, on the other hand, plays a more significant role. Halos that are more concentrated increase the probability of lensing, especially of high magnifications. Our results should be considered conservative as we are ignoring projection effects (and the effect of baryons) that could increase the concentration of a halo (and consequently the probability of lensing) if it is aligned in the line of sight or if two small halos fall in the same line of sight. We simulate halos within a region extending up to 2.3 times the virial radius (a sufficiently large area is needed in order to minimize edge effects). This is sufficient to contain the entire halo when the ellipticity is $e = 0.2$. The simulated halos have typical spatial resolutions of 0.6 kpc per pixel for a lens at $z \approx 0.6$ and mass of $2 \times 10^{15}\,M_\odot$ and about 50 pc for a halo of $10^{12}\,M_\odot$ at the same redshift. This resolution is significantly higher than can be obtained with cosmological N-body simulations. This is important to reproduce high magnifications with accuracy. For the halo profiles, we consider a NFW model. Earlier work focusing on smaller galaxies have used different profiles; the isothermal model is one of the most popular. However, to explore in detail the regime of extreme magnifications, intermediate-mass halos are the most relevant and for these types of halos, the NFW profile gives a good description of the total mass (dark matter plus baryons), especially for high-mass halos. We note that we explicitly ignore the role of baryons in the central galaxy in the halo (other than by assuming a core radius in the NFW profile, as described above). Baryons are expected to increase the convergence in the central part of halos (although they can also decrease it due to feedback). This has a non-negligible impact on small halos that may be subcritical without the help of the baryonic component (large halos are already supercritical (i.e., their surface mass density is above the critical surface mass density in parts of the halo) and adding the baryons has little effect on the probability of magnification. The effect of ignoring the baryon effect is discussed in more detail in Sect. 4.

The magnification is computed based on the inverse ray shooting method (see e.g., Kayser et al. 1986). This method is appropriate for computing the total magnification including all possible counterimages. The total flux needs to be accounted for when considering extreme magnifications, since counterimages may appear separated by a small fraction of an arcsec, and hence be unresolved or barely resolved (or a macroimage may break into multiple microimages separated by a few milliarcseconds or less). In Fig. 1 we show the magnification for a halo at $z_l = 0.4$ with mass $2 \times 10^{15}\,M_\odot$ and for a source at redshift $z_s = 4$. The left panel shows the central portion of the halo with the magnification computed in the image plane (assuming a point source in the background). The critical curves where the magnification diverges are clearly visible. The color scale indicates the magnification in logarithmic units. The right panel shows the magnification in the source plane as derived from the inverse ray shooting technique. The field of view in this case is half the size of that in the left panel. The caustics are also clearly visible with the diamond shape caustic mapping into the tangential critical curve, and the elliptical caustic mapping into the radial critical curve. In the context of this paper, the most striking aspect that is of interest to this work is the clear discontinuity in the magnification at the caustics. Contrary to what happens with the magnification in the image plane, where the magnification appears to be continuous across the image, the source plane shows a sharp discontinuity at the caustics. This has interesting implications for the probability of high magnification. To compute the probability of magnification, two different approaches are commonly used in the literature. The fast deprojection method is used normally when computing lensing probabilities for large lenses, like galaxy clusters. This is an inexpensive method that simply deprojects the image plane (using the lens equation) into the source plane. Since the magnification, μ, is defined as the ratio of areas between the observed images and the sources, the area of a pixel with magnification μ in the image plane is reassigned to an area in the source plane that is μ times smaller.

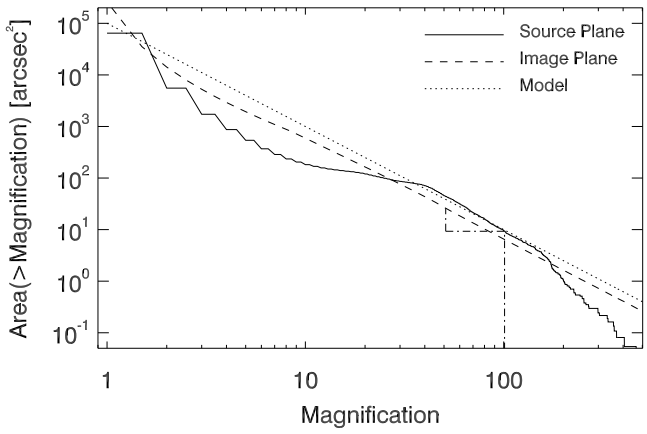

Fig. 2. Area above a given magnification for a single halo at $z_l = 0.4$ with mass $2 \times 10^{15}\, M_\odot$ and for a source at redshift $z_s = 4$. The area computed in the image plane does not follow the μ^{-2} law at low and moderate magnifications ($\mu < 40$). At very high magnifications ($\mu > 150$), resolution effects break the μ^{-2} law, but the area can be safely extrapolated to high values of the magnification above $\mu \approx 50$.

This method is very fast, but it does not give the right probability for the total magnification by ignoring the multiplicity of images. When multiple images are produced, several regions in the image plane (with different magnifications) project back into the same area in the source plane.

Instead, an inverse ray shooting method, where pixels that originate in the image plane and land in the source plane are counted for every pixel in the source plane, is the appropriate method for computing the total magnification. In this work we are interested only in the total magnification since we assume that counterimages of a small background source with extreme magnification form very close to the critical curve and the multiple images (carrying the total magnification) may appear as a single unresolved image. Figure 2 shows the difference between the two methods for the cluster shown in Fig. 1. The dashed line corresponds to the case where the area is computed in the image plane and later divided by μ (deprojection method). This method overestimates the area at low and intermediate magnifications and underestimates it at high magnifications. The inverse ray tracing method is shown as a solid line in Fig. 2. The inverse ray shooting method suffers from resolution limitations at very high magnifications. In this regime, the width of the region around the caustic that has such high magnification is smaller than the pixel size used to do the mapping between the image and source planes. As a consequence, the probability of lensing does not follow the μ^{-2} law at the highest magnification. This can be seen in the solid line at magnifications higher than $\mu \approx 150$.

The remedy for this problem is to simply extrapolate the μ^{-2} law from this point. The plateau observed at intermediate magnifications ($10 < \mu < 40$) corresponds to the gap in magnification in the source plane between the regions inside the caustics and the regions outside the caustics. Generally, the larger the caustic region, the larger this gap and the more pronounced the plateau. Above a certain value of the magnification ($\mu \approx 40$), the area follows the standard law μ^{-2} expected for fold caustics. This value of the magnification ($\mu \approx 40$) corresponds approximately to the local minimum of the magnification at the center of the diamond shaped caustic. Halos that are subcritical (i.e., their surface mass density is lower than the critical surface mass density) do not have a plateau and fall faster than the μ^{-2} law without reaching high magnifications. As a halo approaches the critical point, its probability of lensing starts approaching the μ^{-2} law. To better show the relationship between the two methods, in Fig. 2 we show the connection between the two in the regime of high magnification, where both curves fall like μ^{-2}. If we take a point with magnification $\mu = 50$ on the dashed line and divide the area by two moving downward in the vertical direction, and later move in the horizontal direction up to the point where the magnification is $2\mu = 100$, we end up in the solid curve. In other words, the deprojection method counts the area twice, but it only accounts for the magnification of one of the counterimages. This is easy to understand: at high magnification most of the flux is divided into two counterimages, each one carrying approximately half the total magnification. At lower magnifications, there may be three or more dominant counterimages making the relationship between the two methods less clear.

4. Lensing probability for a cosmological volume

The probability of having an event with magnification higher than certain value is a key ingredient when computing the rates of lensed events. The number of observed events from the redshift interval $z_s \pm \Delta z/2$ that are lensed with magnification higher than μ by a population of lenses between redshift 0 and redshift z_s is given by

$$\frac{\mathrm{d}N(>\mu, z_s)}{\mathrm{d}z_s} = \mathrm{d}V(z_s)\, \Delta z\, R(z_s)\, P(>\mu, z_s), \tag{7}$$

where $\mathrm{d}V(z_s)\Delta z$ is the volume element at z_s contained in the spherical shell with thickness Δz, R is the intrinsic rate of events or number of events per unit volume and unit time at the redshift z_s, and $P(>\mu, z_s)$ is the probability of having a magnified event at z_s that is being amplified with magnification higher than μ. When R represents a rate of events (transients), it needs to be divided by the factor $(1 + z_s)$ in order to account for the time stretch between the observed rate and the intrinsic rate.

The product $V(z_l)\, \Delta z\, R(z_s)$ corresponds to the total number of events per unit time in the volume being lensed ($z_s - \mathrm{d}z_s < z < z_s + \mathrm{d}z_s$). The probability $P(\mu, z_s)$ is defined as

$$P(>\mu, z_s) = \int_0^{z_s} \mathrm{d}z_l \frac{\mathrm{d}p(>\mu, z_l, z_s)}{\mathrm{d}z_l}, \tag{8}$$

where $\mathrm{d}p(>\mu, z_l, z_s)/\mathrm{d}z_l$ is the fraction of the total area (at the distance of the source) that is being magnified by a factor larger than μ (see e.g., Turner et al. 1984),

$$\frac{\mathrm{d}p(>\mu, z_l, z_s)}{\mathrm{d}z} = A_T(z_s)^{-1} \frac{\mathrm{d}V(z_l)}{\mathrm{d}z_l} \int \mathrm{d}M \frac{\mathrm{d}N}{\mathrm{d}M\mathrm{d}z} A_N(\mu, M, z_l, z_s), \tag{9}$$

where $\mathrm{d}V(z_l)/\mathrm{d}z_l$ is the volume element at the redshift of the lens integrated over the entire sky, $\mathrm{d}N/\mathrm{d}M\mathrm{d}Z$ is the halo mass function, $A_N(>\mu, M, z_l, z_s)$ is the area with magnification higher than μ for a lens with mass M at redshift z_l lensing a background source at redshift z_s, and $A_T(z_s)$ is the area in physical units of the spherical shell at redshift z_s. That is,

$$A_T(z_s) = 4\pi D_a(z_s)^2, \tag{10}$$

where $D_a(z_s)$ is the angular diameter distance at the redshift of the source. The definition of $P(>\mu, z_s)$ is essentially the same as the optical depth of lensing by a factor larger than μ, so we can refer to $P(>\mu, z_s)$ as $\tau(>\mu, z_s)$.

For the mass function, we consider the one from Watson et al. (2013); it is appropriate for a wide mass range

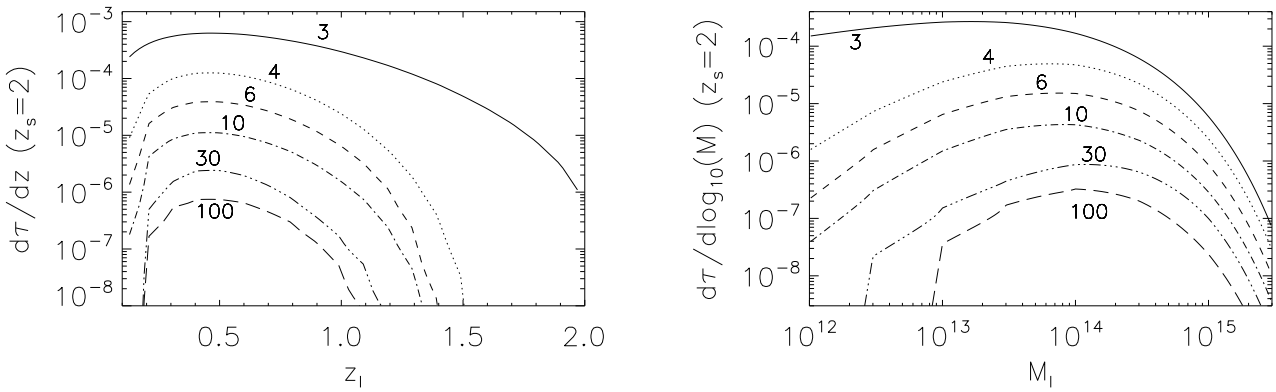

Fig. 3. *Left panel*: contribution to the optical depth for a source at redshift $z_s = 2$ as a function of the redshift of the lens z_l and for all lens masses. *Right panel*: contribution to the optical depth for a source at redshift $z_s = 2$ as a function of the mass of the lens M_l and for all lens redshifts. In both cases, the numbers by each curve indicate the magnification that the curve contributes to. For high magnifications, only the most massive halos contribute to the magnification.

(from small groups to the most massive clusters). Alternative mass functions, for example the Tinker mass function, are expected to produce similar results since the main differences between these mass functions appear at the regime of very massive clusters, which have a small impact for our study. The Watson et al. (2013) mass function is more accurate than the Tinker function in this regime, due to the larger volume of the JUBILEE simulation (Watson et al. 2014). We impose a lower cutoff in the mass and compute the integral between the mass limits $M_{\rm min} = 10^{12}\,M_\odot$ and $M_{\rm max} = 3\times10^{15}\,M_\odot$. The selected mass range is appropriate for computing the probability of lensing at high magnification, as we show in Sect. 4. At low magnifications, the mass range needs to be extended toward lower masses, since even subcritical halos can contribute in this regime. However, a lower mass limit of $10^{12}\,M_\odot$ is sufficient for our purposes since we focus on the high-magnification regime. To model the quantity $A_N(>\mu, M, z_l, z_s)$ we simulate caustics for different halo masses, halo redshifts, and background sources, then compute $A_N(>\mu, M, z_l, z_s)$ for each model through inverse ray tracing (see Sect. 3). Since inverse ray tracing has limited resolution in areas of high magnification, we compute the lensing probability $P(>\mu, z_l, z_s)$ only up to values of $\mu = 100$. At values of μ significantly higher than $\mu = 100$, we observe resolution effects. However, $\mu = 100$ is high enough that above this number, $P(>\mu, z_l, z_s)$ scales as the usual μ^{-2} law, typical of fold caustics, and can be safely extrapolated for each halo toward higher magnifications. Obviously, this extrapolation is performed only for those halos that exhibit supercritical behavior (i.e., have caustics). For subcritical halos no extrapolation is needed as they simply do not reach the value $\mu = 100$. This computation is repeated for different halo masses, halo redshift, and source redshift. We adopt an adaptive resolution scheme, so the magnification around small halos are computed at higher resolution than larger halos in order to properly resolve the caustic regions. The global probability of lensing of a source at redshift z_s is computed after integrating the lensing probabilities of all individual halos in the mass range considered, and up to the redshift of the source. In this work we assume a maximum redshift for the lenses of $z_{\rm max} = 2.0$, since most lenses that contribute to the high magnification are below $z = 2$. For the sources, we assume $z_{\rm max} = 5.0$. The increase in probability between redshift $z_s = 5$ and $z_s = 10$ is modest, at the level of 20% (Zackrisson et al. 2015).

As mentioned earlier, the optical depth $\tau(>\mu, z_s)$ is given by Eq. (8). The optical depth defined this way gives the probability of magnification with a factor larger than μ of a given background source at z_s with infinitesimally small radius.

Figure 3 shows the differential contribution to the optical depth for a source at $z_s = 2$ as a function of the redshift of the lens (left panel) and as a function of the mass of the lens (right panel). The numbers next to each curve show the magnification at which the contribution is computed. In the left panel, we integrated over the entire mass range. At low magnifications, contributions from halos come from the entire redshift range with a peak around $z \approx 0.5$. However, as the magnification increases, the redshift range of the halos that contribute to that magnification becomes narrower, although it maintains the peak at $z \approx 0.5$. The reason for this decrease is made more evident in the right panel where now we have integrated over the total redshift range. High magnifications are produced only by large lenses. The sharp decline in optical depth for $\mu = 100$ at 10^{13} solar masses marks the transition between subcritical and supercritical halos. At lower magnifications ($\mu < 10$) small halos play an important role and cannot be neglected.

The right panel of Fig. 3 predicts that the halos that contribute the most to the optical depth at high magnification are in the range of $10^{14}\,M_\odot$ (at magnification factors of approximately 3, halos ten times smaller are the main contributors). The calculation points out that larger lenses, although less numerous, can magnify more than one background object, as evidenced by galaxy clusters that can lens tens of sources each, while smaller halos typically lens one background object at most.

This appears to be in tension with current observations, and also some earlier theoretical work that suggest that less massive halos are the most typical lenses (see however Fig. 5 in Hilbert et al. 2007, where a similar result is also found). This is supported in part by results from optical surveys, such as SDSS, where it is found that most known gravitational lenses are less massive halos (Inada et al. 2012), not the more massive halos expected from Fig. 3. Earlier works also predicted the lensing effect from these early-type galaxies. There are different reasons for this tension that deserve a discussion in this section. From the theoretical side, most of the previous calculations found in the literature focus on low magnifications ($\mu \sim 3$). In the regime of low magnifications, the role of low-mass halos is more relevant, as is shown in the right panel of Fig. 3. At higher

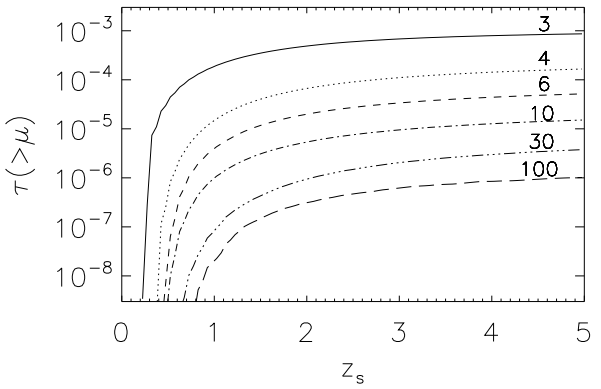

Fig. 4. Optical depth for different magnifications as a function of redshift. The numbers by each curve indicate the corresponding magnification.

magnifications more massive halos take a leading role, as shown by the same figure (see also Keeton et al. 2005, where a similar trend is observed also for single highly magnified images). In earlier predictions, it is also customary that the optical depth is computed in the context of galaxy-galaxy lensing, which explicitly ignores the contribution to the optical depth from groups and clusters. These calculations draw the population of lenses from the velocity dispersion of early-type galaxies (Fukugita et al. 1992; Collett 2015). The velocity dispersion of galaxies derived from observations sets a natural limit on the maximum velocity dispersion, at the maximum radius of the stellar component. Also, much of the literature explicitly focuses on Spherical Isothermal Sphere (SIS) profiles, which are known to increase the optical depth of low-mass halos compared with the NFW halos used in our work (see e.g., Figs. 9–11 in Lapi et al. 2012, Fig. 2 in Porciani & Madau 2000, or Fig. 9 in Perrotta et al. 2002 where the dependency with M is explicitly shown). Regarding observations, the discrepancy between our calculation and observations (where most strong lenses are expected to be smaller halos) occurs because we ignore the cooling role played by baryons since they can increase the relative contribution of small halos to the optical depth of lensing. Baryons can promote the subcritical central region of a small halo to supercritical values. This is in agreement with observations, where most QSOs and lensed galaxies have been found so far around early-type galaxies in optical surveys, such as SDSS (Inada et al. 2012), and not in more massive halos as predicted by our calculation. However, the observed magnifications for these lensed QSO are moderate (a few tens at most). For extreme magnifications, larger halos are more relevant and for these massive halos, the role of the baryonic component or the central galaxy diminishes since these halos are already supercritical (Meneghetti et al. 2003).

Large halos are better modeled by NFW profiles, while smaller halos are better represented by SIS profiles. Kochanek & White (2001) argues that around the galaxies embedded in halos, the change from NFW to SIS "explains why many lenses found in groups of galaxies were associated with the galaxies in the group rather than the group halo, even though the group halo had to be more massive than its component galaxies." It is unclear to what extent many of the elliptical galaxies identified as lenses in optical surveys, such as SDSS, are not in reality the central galaxy of more massive halos with virial masses several times $10^{13}\,M_\odot$ (some are clearly identifiable as such). Early-type galaxies with large velocity dispersion are known to be at the centers of halos with several times $10^{13}\,M_\odot$.

A more accurate model would take into account this transition between the low-mass halos (SIS-like) to the more massive halos (NFW-like) (see e.g., Keeton 1998; Porciani & Madau 2000). This transition is mostly due to the more important role that baryons (cooling) play in small halos (Kochanek & White 2001). Porciani & Madau (2000) show that below a halo mass scale of $\approx 3.5\times10^{13}$ a SIS model is more appropriate, while for masses above this value, a NFW profile is a better description. A hybrid model where small substructures (including the central galaxy) are modeled as SIS while the main halo follows a NFW-like profile may be the most suitable models. Such detailed modeling is, however, beyond the scope of this paper. Since this work focuses on the highest magnifications, for which more massive halos are more relevant, we consider for simplicity only NFW profiles, but it should be kept in mind that the contribution from small halos or substructures (especially at low magnifications) is underestimated by adoptiong a NFW profile. Our results on the optical depth should then be considered conservative.

Figure 3 also clearly shows that the integrated optical depth decreases rapidly with increasing magnification. This is shown in more detail in Fig. 4, where again the number next to the curves indicates the magnification. As found in earlier work, the probability of lensing declines sharply below $z_s = 1$ and changes much more gently above $z_s = 2$. The relative change between $\mu = 3$ and $\mu = 10$ is much more abrupt than between $\mu = 10$ and $\mu = 30$ or between $\mu = 30$ and $\mu = 100$. This is a direct consequence of the transition between the subcritical and supercritical regime shown in Fig. 2, where the probability of lensing drops faster than μ^{-2} at low magnification factors. The same pattern is observed when integrating over cosmological volumes as shown in Fig. 5. We note that in this plot the μ^{-2} is satisfied at $\mu < 100$ for all redshifts.

The optical depth is, in general, in agreement with previous results based on N-body simulations (see Fig. 5 in Hilbert et al. 2007); however, it should be noted that the optical depth in Hilbert et al. (2008) may still be affected by the limited resolution of the N-body simulation, which would impact smaller halos to a greater extent. However, since we neglect the effect of baryons, our result should be considered conservative, as mentioned above. As noted in Fig. 4 of Hilbert et al. (2008), when baryons are added the mass range that contributes the most to the optical depth shifts from $\approx 8\times10^{13}\,M_\odot$ to $\approx 3\times10^{13}\,M_\odot$. Hilbert et al. (2008) concludes that baryons can increase the optical depth by up to a factor of 2. As mentioned earlier, at low magnifications ($\mu < 10$), the role of small halos is important so our estimation of the optical depth may be biased low by a factor of ≈ 2 in this regime. However, for the high magnifications we are interested in, the error introduced in the optical depth by ignoring the effect of baryons is expected to be less than that factor of 2.

Finally, the baryonic component distorts the magnification through microlensing. As discussed in more detail below, microlensing reduces the maximum magnification of a small background source that could be attained by a macromdoel. The reduction in magnifying power is proportional, to the first order, to the surface mass density of microlenses, i.e., to the baryonic component. Since around critical curves the stellar component is, in general, larger for small halos than for massive halos, extreme magnification of small background objects is possible only when the stellar component near the critical curves is relatively small, for example in groups and clusters.

5. Role of microlenses

The above results are derived under the assumption that the caustic is not disrupted by substructure and that the $\mu^{-1} \propto \sqrt{|\beta-\beta_0|}$

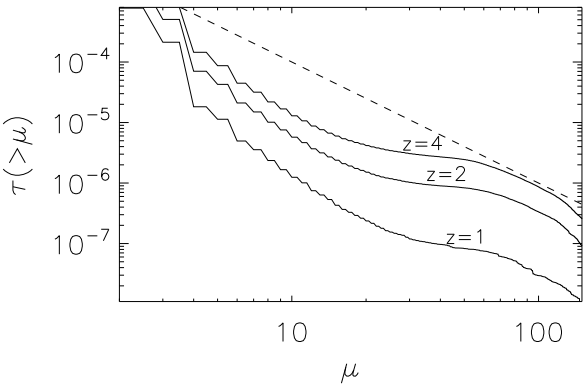

Fig. 5. Optical depth for different redshifts as a function of magnification. The dashed line is the power law μ^{-2}.

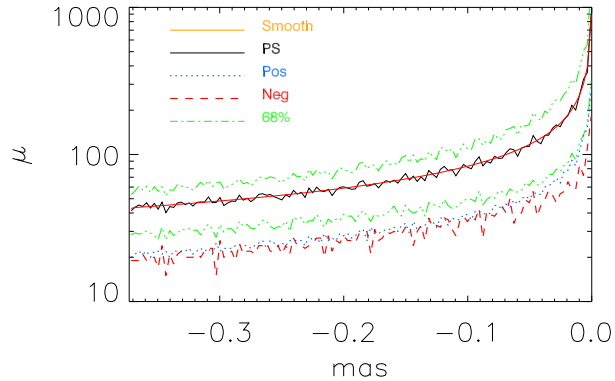

Fig. 6. Mean magnification in the source plane and its variability as a function of distance to the macrocaustic.

law is maintained everywhere in the vicinity of the caustic in the source plane, where β_0 is the position of the caustic (in the image plane the corresponding law is $\mu^{-1} \propto |\theta - \theta_0|$, where θ_0 is the position of the critical curve). The probability density function (or pdf) of the magnification is simply $dN/d\mu \propto \mu^{-3}$. However, as discussed in Diego et al. (2018), in the presence of substructure, including microlenses with stellar or even substellar masses, the caustic can suffer severe distortion and the $\mu^{-1} \propto \sqrt{|\beta - \beta_o|}$ is no longer a valid approximation. The maximum magnification – which in theory can reach values as high as $\sim 10^7$ for a background object with a size similar to that of the Sun and for a powerful lens such as a massive galaxy cluster – can be significantly lower when microlenses are present. The maximum magnification is still proportional to the inverse of $\sqrt{R}$, where R is the radius of the background object, but the proportionality constant is now related to the Einstein radius of the microlenses rather than the much larger Einstein radius of the macrolens. To be more precise, at low effective optical depth of microlensing[1], the proportionality constant is related to the Einstein radius of the microlenses, but magnified by $\sqrt{\mu_{\rm macro}}$, where $\mu_{\rm macro}$ is the magnification of the macrolens in the absence of microlenses (see Diego et al. 2018, for a discussion of this effect). When microlenses are involved, we distinguish between the total magnification μ (from the combined potential of the macromodel plus microlens) and the magnification of the macromodel $\mu_{\rm marco}$.

At high effective optical depth of microlenses even this scaling with $\sqrt{\mu_{\rm macro}}$ breaks down due to overlapping microcaustics, and the maximum magnification becomes a function of the macromodel surface mass density of microlenses and the distance to the caustic of the macromodel (Diego et al. 2018). In this section, we explore the departure of the pdf from the standard law $dP/d\mu \propto \mu^{-3}$, when the caustic of a macromodel is in the presence of microlenses. We make use of simulations similar to those described in Diego et al. (2018), and the reader is directed to that reference for details on the simulations. Here we simply summarize the most practical aspects of the simulation. For the macromodel we chose a synthetic model that mimics the caustic region of a small galaxy. Adopting a relatively small galaxy is convenient since the magnification drops more quickly as we move away from the caustic, which it would do if we were working with a caustic of a massive cluster. A rapid drop in the magnification allows us to study a wide range of magnifications in a relatively small area, and include regions in the source plane (relatively far away from the caustic) where the impact of microlenses is much less significant. The results derived in this section can easily be translated to larger lenses. For simplicity, we adopt a model that produces a vertical critical curve (and vertical caustic) so the magnification from the macromodel remains constant when moving in the vertical direction and falls as $1/|x - x_0|$, where x_0 is the position of the critical curve. The values of total κ and γ are 0.6667 and 0.3333, respectively, at the position of the critical curve. For simplicity, we fix γ and only modify κ on either side of the critical curve with a slope of 10^{-7}, which is appropriate for galaxy halos (the slope of κ goes as the inverse of the Einstein radius of the lens in the vicinity of the critical curve for popular models like the isothermal model). The simulation covers a region of $0.178'' \times 0.00093''$ in the lens plane with a resolution of 31 nanoarcsec per pixel, resulting in approximately 1.7 billion pixels. The dimension in the horizontal direction is much larger than in the vertical direction. This is a consequence of the tangential magnification $(1/(1-\kappa-\gamma))$ being much higher than the radial magnification $(1/(1-\kappa+\gamma))$. The much wider extension in the horizontal direction is required in order to capture the change in the properties of the magnification as one moves away from the caustic, but more importantly because the caustics from microlenses that are relatively far away from each other can still overlap in the source plane near the caustic. The deflection field from the microlenses is computed from the Spera et al. (2015) model, normalized to a surface mass density of $\Sigma_o = 19\, M_\odot\, {\rm pc}^{-2}$. This is similar to the value used in Diego et al. (2018) and a typical value in the outskirts of a galaxy. This surface mass density results in values of $\kappa_* \sim 0.008$ (or $\kappa_*/\kappa = 0.012$), where κ_* is the convergence due to the microlenses.

To compute the magnification in the source plane we use an inverse ray tracing algorithm. In Fig. 6 we show the average total magnification (black solid line) in the source plane after binning the source plane in regions of width $3\,\mu$as and as a function of the distance to the macromodel caustic. The red solid line shows the corresponding magnification of the macromodel, when microlenses are not included in the simulation. The average magnification is basically the same, reflecting the conservation of flux. The dot-dashed green line marks the 68% region of the magnification variation. These curves lie very close to the dispersion values suggesting that, to first order, the pdf of the magnification when microlenses are present and at a fixed distance to the caustic can be approximated by a Gaussian. The width of this Gaussian scales with the effective surface mass density (i.e., the surface mass density corrected by the factor $\mu_{\rm macro}$). The dotted blue and dashed red lines show the average

[1] This definition was introduced in Diego et al. (2018) to account for the fact that a microlens in a region with magnification $\mu_{\rm macro}$ and with mass M behaves as a microlens with an effective mass $M\mu_{\rm macro}$.

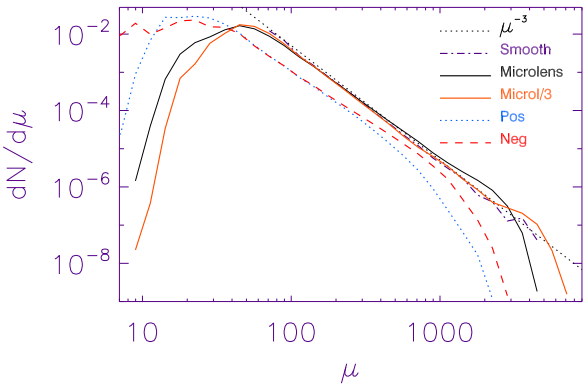

Fig. 7. Probability of magnification in the source plane. The black solid line shows the probability of the total magnification when microlenses are present. The black dot-dashed line shows the probability computed in the same area, but when microlenses are not included (i.e., the smooth model). The black dotted line is the μ^{-3} law expected for smooth caustics. The blue dotted and red dashed curves are the probability of magnification from the regions with positive and negative parity, respectively.

magnification as a function of the distance to the macrocaustic computed on the sides with positive and negative parity respectively. Each side carries roughly half the total magnification and behaves very similarly in terms of average magnification at a fixed distance. However, this similar behavior is not maintained for the higher moments of the pdf. As discussed in earlier works (Chang & Refsdal 1979, 1984; Kayser et al. 1986; Schechter & Wambsganss 2002; Diego et al. 2018), the pdf on the side with negative parity presents larger fluctuations than the corresponding pdf on the side with positive parity. The same result is found in our simulation, but covering a range of magnifications not studied in earlier work. Figure 7 shows the pdf of the total magnification (solid black) and also for the magnification on the sides with positive (dotted blue) and negative (dashed red) parities computed over a region of $370 \times 56\,\mu\mathrm{as}^2$ and including the macrocaustic.

The pdf of the total magnification when microlenses are included resembles the magnification of the smooth model except at the highest magnifications where, as discussed in Diego et al. (2018), the maximum magnification is lower than the corresponding maximum magnification of the smooth model. However, at more modest magnifications ($\mu \approx 1000-3000$), the model with microlenses has an increased probability compared with the smooth model. The contrary is found above $\mu > 2000$, where the probability of the magnification no longer follows μ^{-3}, but falls sharply as a log-normal. Also, as discussed in Diego et al. (2018), the side with negative parity exhibits a higher probability of high magnification than the side with positive parity. This is compensated by a higher probability of low magnification on the side with negative parity, confirming earlier results that show how fluctuations on the side with negative parity can be larger than on the side with positive parity (Schechter & Wambsganss 2002). The departure from the μ^{-3} law starts to be evident at $\mu \approx 500$ (below this value a pdf evaluated over a wider region would show agreement between the smooth model and the model with microlenses at lower magnifications). The value of $\mu \approx 500$ above which the pdf deviates from the μ^{-3} law can be approximately derived from theoretical arguments. Following Diego et al. (2018), the optical depth of microlensing τ can be approximated by

$$\tau = \frac{\Sigma}{M}\pi\frac{4GM}{c^2}\frac{D_{\rm ds}}{D_{\rm d}D_{\rm s}}\mu_{\rm t} = 4.2\times10^{-4}\Sigma(M_\odot/{\rm pc}^2)\frac{a_2\mu_o}{\theta}, \qquad (11)$$

where $\Sigma(M_\odot/\mathrm{pc}^2) = 19$, $a_2 = 1.5$, and $\mu_o = 0.46''$ for the model considered in this work; $a_2 = mu_r^{-1}$ represents the inverse of the magnification in the direction parallel to the macrocaustic; while μ_o defines the strength of the lens and enters in the expression $\mu = \mu_o/\theta$, where μ is the total magnification and θ is the distance to the critical curve of the macromodel. As discussed in Diego et al. (2018), the above expression for τ is derived under the assumption that all microlenses have the same mass. When compared with actual simulations based on a realistic mass function for the microlenses, Diego et al. (2018) find that Eq. (11) overestimates the true optical depth by a factor of ≈3. Accounting for this factor of 3 to correct the optical depth we find that τ saturates (i.e., $\tau = 1$) when $\theta = 1.83$ mas. At this distance the total magnification from the macromodel is $\mu = \mu_o/\theta = 250$ a factor ≈2 from the estimated value of $\mu \approx 500$, where the pdf of the total magnification for the model with microlenses start to deviate from the μ^{-3} law. The value of the magnification where this departure from μ^{-3} is observed is inversely proportional to $\Sigma(M_\odot/\mathrm{pc}^2)$. Figure 7 shows one more example corresponding to a situation where the $\Sigma(M_\odot/\mathrm{pc}^2) = 6.3$, that is, 3 times lower than in the previous case. The pdf of the total magnification for this case is shown as a red solid line and the deviation from μ^{-3} takes place clearly at a higher magnification ($\mu \approx 2000$). Interestingly, the departure from μ^{-3} is more abrupt, but the log-normal cutoff seems to be maintained. Lower values of $\Sigma(M_\odot/\mathrm{pc}^2) = 6.3$ are hence more favorable to observe more extreme events of many thousands, but they have a lower probability for intermediate magnifications of $\mu \approx 1000$. On the contrary, the higher the value of $\Sigma(M_\odot)/\mathrm{pc}^2$, the sooner the pdf of the total magnification will deviate from μ^{-3} (although compensating the pdf of μ with an increase toward magnification ∼1000).

In order to explore in more detail the departure from the μ^{-3} law we compute the magnification at a resolution ten times higher (i.e., 3.1 nanoarcsec per pixel or ≈600 times the diameter of the Sun at $z = 1.5$) in a smaller region that zooms in around the macrocaustic. Since the maximum magnification scales as the inverse of the square root of the source radius, the results presented here are valid for sources with a radius ≈1000 times the radius of the Sun. A source with a radius R that is 100 times smaller than this could be amplified by a factor 10 times larger when it is at a distance R from the caustic. Although the simulated region is smaller, we still account for the contributions from the microlenses that lie far away from the macro critical curve. Figure 8 shows the pdf in this smaller region for the smooth model (black dot-dashed curve) and the model with microlenses (blue dotted and blue dashed curves). The difference between the blue-dotted and the blue-dashed is due to the different fraction of the caustic region that is used to compute the pdf. The blue-dotted curve is derived from a narrow region (width ≈5μas) very close to the macrocaustic, while the blue-dashed curve is computed from a wider region (width ≈ 25μas). The widths of these regions, as well as the caustic zone, are shown in Fig. 9, where for convenience we have rotated the simulated region so the macrocaustic is oriented in the horizontal direction.

The pdf in the narrower region follows the log-normal distribution originally discussed in Diego et al. (2018), which is typical of situations where many microcaustics overlap (this is the saturation regime discussed in Diego et al. 2018). When computing the pdf over a wider region, the pdf starts to resemble the μ^{-3} law at lower magnifications ($\mu \sim 2000$). The departure from this law at lower magnifications is largely due to the fact that the zoomed region does not include areas farther away from the macrocaustic that still contribute to these magnification bins. The black dot-dashed line shows the pdf of the smooth

model (i.e., the same zoomed region but without microlenses). The black solid line shows the μ^{-3} law expected for smooth caustics. We note how the excess of probability (a factor of ≈2) when microlenses are present is more evident in this plot around $\mu \approx 2000$. On the contrary, at $\mu \approx 10\,000$ the probability of magnification is at least two orders of magnitude lower in the presence of microlenses, setting a natural limit to the maximum magnification around this value. If the contribution from microlenses is smaller, it is possible to increase the maximum magnification as shown by the red curves in Fig. 8. These curves are derived from a model where the surface mass density of microlenses is three times lower than the model in the blue curves. As shown in Fig. 7, below $\mu \approx 2000$ the pdf of the magnification closely follows the μ^{-3} law, but beyond $\mu \approx 2000$ the pdf deviates from this relation, and at magnifications $\mu \approx 10\,000$ the pdf is already falling log-normally. Since microlenses are expected to be ubiquitous up to the boundaries of most macrolenses, it is expected that observed extreme magnifications higher than a few tens of thousands will be extremely rare.

If microcaustics are disturbed by microlenses and the maximum magnification is of order 10^4 rather than 10^7 as predicted for smooth caustics in clusters and small background sources like luminous stars, this may explain the apparent lack of observations of caustic crossing events from bright stars at $z = 1$–2 (with the only exception of the Icarus/Iapyx events discussed in Kelly et al. 2018). To boost the flux of a distant luminous star by 10 magnitudes (so it can be detected) we need high magnifications (of order $\sim 10^4$) that, as discussed above, can be two orders of magnitude less likely when microlenses are present. Since lower surface mass density of microlenses distorts the caustic less, the most extreme magnification factors are possible only with the smallest number of microlenses. This implies increasing the redshift of the background source so the critical curves move farther away from the central region and the surface mass density of microlenses gets lower. Alternatively, we can identify portions of the critical curves at lower redshift, which have relatively small contributions from microlenses. For example, the critical curves connecting two clusters in the process of merging produce near pristine critical curves between the two clusters like the ones observed in the G165 cluster (Frye et al. 2019; Griffiths et al. 2018).

6. Expected rates for compact bright sources

In this section we present rough estimates of the expected rates for different types of compact background sources. In particular we focus on the expected rate of extreme magnification events (with $\mu > 100$–1000) for SNe, luminous stars at redshifts 1–2, Pop III stars at $z > 6$, and gravitational waves from binary black holes (BBH). Rates of QSOs are not be considered in this work as they have been studied in detail elsewhere (including the impact of microlenses). Also, since the accretion disk of a QSO is orders of magnitude larger, about 10 light-days (with 1 light-day ≈ 170 AU), the maximum magnification for a QSO is ∼ 100 times lower (see also Blackburne et al. 2011; Guerras et al. 2013, who show that the size of the light emitting region can be significantly larger than this).

Supernovae can also reach relatively large sizes (although not as large as QSOs). More interestingly, the expanding photosphere from a SN explosion that is intersecting a caustic will exhibit very distinctive changes in the light curve as a consequence of the varying magnification (see e.g., the recent work of Dobler & Keeton 2006; Goldstein et al. 2018; Foxley-Marrable et al. 2018, and references therein).

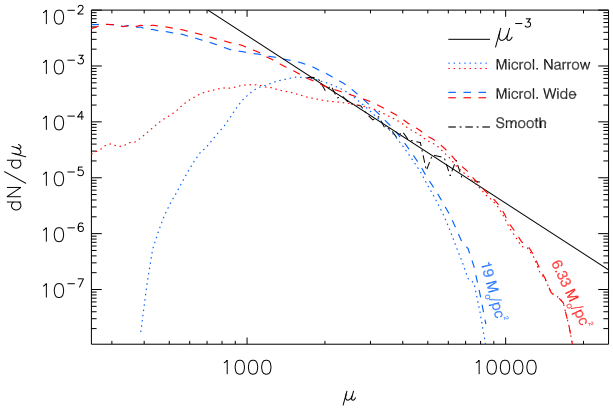

Fig. 8. Pdf of the magnification computed in a smaller region around the macrocaustic at ten times the resolution. The black solid line shows the expected μ^{-3} for smooth caustics. The black dot-dashed line shows the pdf of the smooth model computed through the inverse ray shooting method. The dotted lines are the pdf of the model with microlenses computed over the narrow region in Fig. 9, while the dashed lines are for the pdf computed over the wide region. Blue lines are for the model with $\Sigma = 19\,M_\odot\,\mathrm{pc}^{-2}$ and red lines are for a model with $\Sigma = 6.33\,M_\odot\,\mathrm{pc}^{-2}$.

During the early phases of the SNe (including the progenitor or pre-explosion phase), if the progenitor is at a fraction of 1 AU away from a caustic, the fainter but much more compact photosphere can be magnified by extreme factors owing to the relatively small size of the growing photosphere. After the initial explosion, the SN expands at rates of $\approx 10^4\,\mathrm{km\,s^{-1}}$ for several days (or even weeks) resulting in sizes for the photosphere of ≈6 AU one day after the explosion and ≈30 AU after 5 days (and reaching $\sim 3 \times 10^{-4}$ pc at the time of maximum brightness). If the SN is located at a few tens of AU from a caustic (or microcaustic), its maximum magnification will peak when the photosphere touches the caustic. As the photosphere expands, the total magnification will start to decline leaving a distinctive signature in the light curve.

The probability of observing extremely magnified events at redshift z depends on the product of the volume element at z, the volumetric density of objects (or events) at the same redshift and the probability of observing an event above a magnification μ. Figure 10 shows the combined effect of the volume element, the evolution of the rates and the lensing probability. The dashed line shows the number of objects (or event rates) per redshift interval that trace the star formation rate (SFR) as a function of redshift. For this particular example, we use the intrinsic rate of SNe explosions (of all types). The solid line shows the number of these events that would be magnified by a factor $\mu > 100$. The dotted line shows the more realistic case where the observations are flux limited. We define the flux limit as the maximum distance at which an object would be observed without lensing, D_o. In particular, for this curve we adopt D_o as the luminosity distance at $z = 0.3$. Since the received flux scales as $D_l(z)^{-2}$, any event beyond $z = 0.3$ brighter than the flux limit must have a magnification greater than $\mu = D_l(z)^2/D_o^2$ (the same law applies for gravitational waves, where the signal-to-noise ratios of the lensed objects scale as $\sqrt{\mu}/D_l(z)$).

The location where the dotted and solid lines cross correspond to $\mu = 100$, magnifications are higher than 100 to the right of this point and lower than 100 to the left of this point. At low magnifications ($\mu < 50$) the images may be resolved, and the total magnification may not be the best approach to describe the

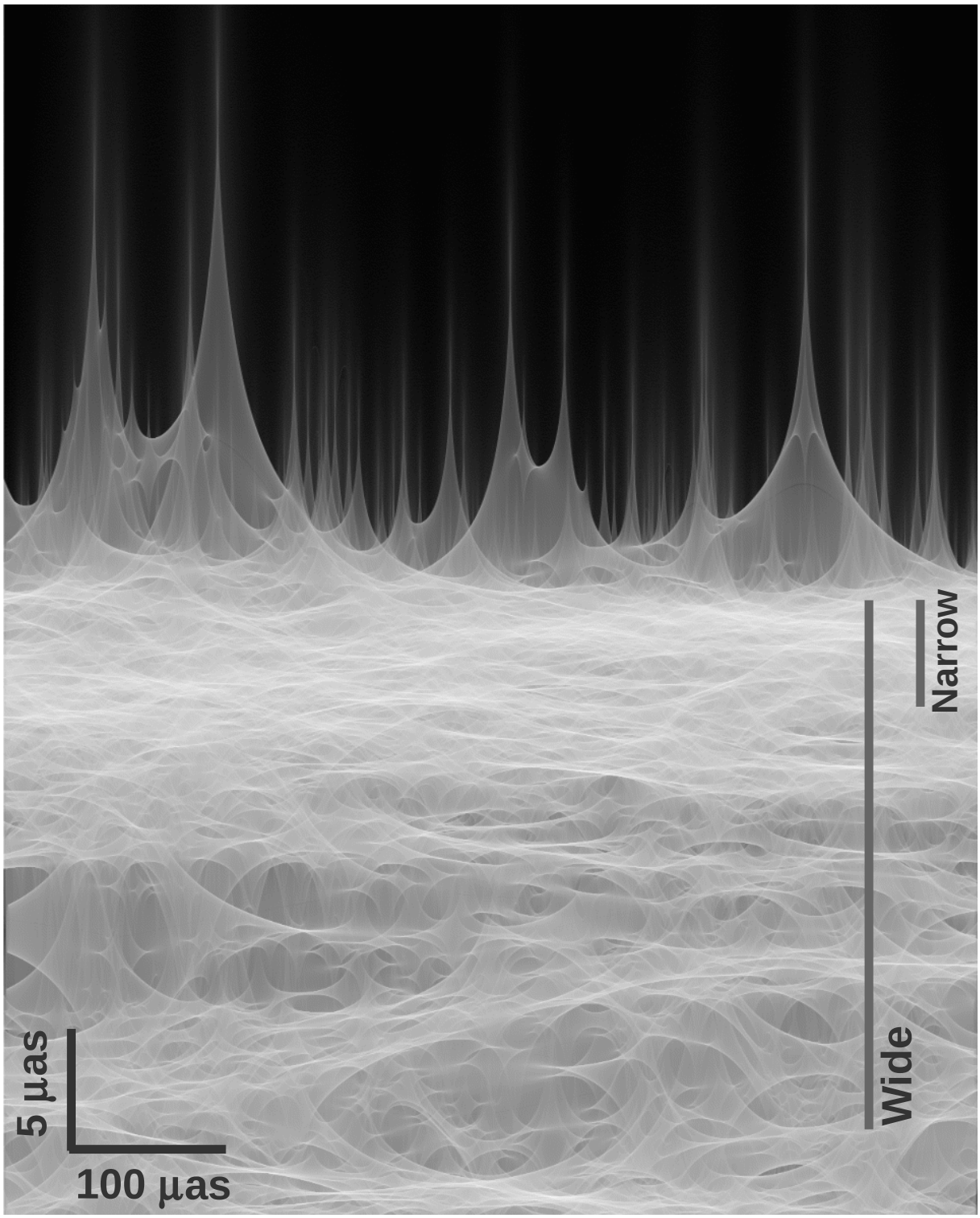

Fig. 9. Magnification in the caustic region in the presence of microlenses with $\Sigma = 19\,M_\odot\,\mathrm{pc}^{-2}$. The original simulated region has been compressed by a factor ∼20 in the horizontal direction in order to show a larger region. The two vertical lines indicate the extension over which the pdf in Fig. 8 is computed for the narrow and wide regions. If microlenses were not present, the caustic from the macromodel would be a single straight horizontal line with a region of zero magnification above it, and magnification decreasing as $1/\sqrt{d}$ below it, where d is the distance to the caustic.

observations. Instead, the magnification carried by each image should be considered, which is ≈1/2 the total magnification for μ-factors larger than a few tens. Consequently, the dotted line at $z \approx 1$ (where the total magnification is ≈20) would be overestimated by a factor ≈4.

6.1. SNe at z=1–3

The volume in a shell of fixed thickness δz is maximum at $z \approx 2.5$. The star formation rate density (SFRD) reaches its maximum value at $z \approx 2$ (Madau & Dickinson 2014). Hence, the

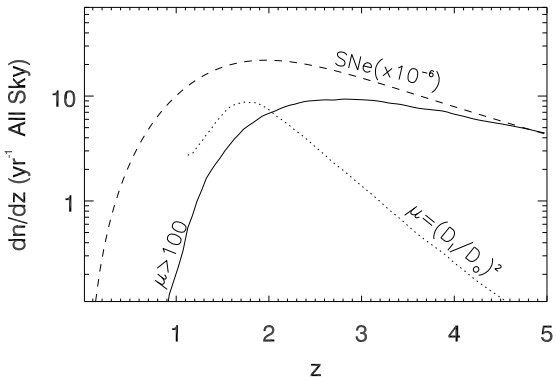

Fig. 10. Rate of events (all sky and in the source frame) for a model that traces the star formation rate and normalized to 8×10^{-5} events per year per Mpc^3 (dashed line). The SNe rate has been re-scaled by a factor 10^6 for presentation purposes. Also shown are the product of the rate times the optical depth for $\mu > 100$ (solid line) and the rate of lensed events with the same apparent magnitude as a similar unlensed event at $z = 0.3$ (dotted line).

volumetric rate of any event that traces the SFR will be maximum at around $z \approx 2$, as shown in Fig. 10. The combined effect of maximum volume and maximum rate at $z \approx 2$ makes this an ideal redshift interval to search for events at extreme magnification. Despite the numerous detections of SNe, so far only three SNe have been gravitationally lensed with relatively large factors (in the range of a few tens at most; Quimby et al. 2014; Kelly et al. 2015; Goobar et al. 2017). Here, we consider the case of SNe (of all types) at high redshifts ($z \approx 1$–3) that are being lensed by extreme factors. As detailed below, during the much fainter phase of the first moments of the explosion, the SN is still small enough that magnifications of $\mu \sim 1000$ or higher can take place if the SN occurs close enough to a caustic.

After a SN explosion, there is typically an intense and short burst of flux lasting from a few seconds to a few hours. This phase is known as the shock breakout. The breakout takes place ≈1.5 h after the explosion (Garnavich et al. 2016; Bersten et al. 2018). After ≈1 day the photosphere has grown by ≈1000 solar radii assuming an expansion rate of 10^4 km s^{-1} (Pejcha & Prieto 2015). We focus on this period of 1 day after the explosion (or 2–3 days in the observed frame for SNe at $z = 1$–2), where the size of the SN is smaller than $1000\,R_\odot$ and the maximum magnification can reach factors $\mu > 1000$. During this first day, the spectrum is concentrated mostly in the UV band. At $z = 1$–2, a significant portion of this emission would be redshifted into the visible band making the SN visible in the more sensitive optical bands. For a supernova like SN 2016gkg with absolute magnitude −17 at its maximum, the estimated magnitude before the shock breakout was $M_V \approx -12$ and reached $M_V \approx -15$ at the maximum of the shock breakout (Bersten et al. 2018). If such SNe can be observed without the help of gravitational lensing magnification at $z = 1$ during its maximum, i.e., with apparent magnitude $m \approx 27$ attainable for instance in the deep-drilling fields planned for the Large Synoptic Survey Telescope (LSST), the same SNe, when observed amplified by a factor $\mu \sim 1000$ may reveal not only a much brighter main peak, but also the shock breakout ($m \approx 25$) and even the last phase of the precursor ($m \approx 26.5$ ignoring k-corrections).

To estimate the rate SNe at $z \approx 1$–3 we adopt a model that traces the SFR (Madau & Dickinson 2014). A model following the SFR is clearly motivated given the relatively short delay time of ≈1 Gyr (for type Ia) between the star formation and the SN explosion (Ruiter et al. 2009). We normalize the model to a rate of 0.8×10^{-4} Mpc^{-3} yr^{-1} for all types of SNe at $z = 0$ (Barbary et al. 2012). For this model, the volumetric rate for any type of SNe at $z = 1$ is 4.6×10^{-4} Mpc^{-3} yr^{-1}, and 7×10^{-4} Mpc^{-3} yr^{-1} at $z = 2$ (both rates given in the source frame, i.e., not corrected by time dilation). For type Ia SNe the corresponding rate would be about 3–4 times lower (Barbary et al. 2012). Hence, the number of type Ia SNe explosion per year and redshift interval is estimated to be $\approx 10^7/2$ at $z \approx 1$ rising to $\approx 2 \times 10^7/3$ at $z \approx 2$ (see Fig. 10), where the dividing factors 2 and 3 account for the time dilation factor. For magnification higher than $\mu = 1000$, the number of events per year is very small (≈0.1 event per year in the interval between $z = 1$ and $z = 3$). For more modest magnification factors, at $z = 1$, we expect ≈0.5 SN per year with magnification higher than $\mu = 100$. At $z = 2$ due to the increase in the intrinsic rate, volume element and optical depth, we expect ≈3 SNe per year (where we have divided the intrinsic rate by a factor $1 + z = 3$). The increase in optical depth at higher redshifts partially compensates the decline in intrinsic rates beyond $z = 2$, so the number density of lensed SNe with $\mu > 100$ remains more or less constant between $z = 2$ and $z = 3$. Above $z = 3$, the modest increase in optical depth is not enough to compensate the declining intrinsic rate and volume element per redshift interval. At magnifications $\mu = 100$, the flux is boosted by 5 magnitudes, enough to see a sufficiently bright pre-shock phase with absolute magnitude $M_V = -12$ at $z = 1$ with a telescopes like JWST. The shear number of events at $z \approx 2$ maximizes the probability that one of these events takes place very close to a caustic.

Surveys covering a large region of the sky with good cadence should be able to identify these SNe as exceptionally luminous ones. One such survey is LSST which will monitor approximately 20 000 deg^2 down to magnitude 24.5 (5σ in r band) with a cadence of a few days (Ivezic et al. 2019; Abell 2009). Hence, LSST should see ≈2 strongly lensed SNe ($\mu \sim 100$ or more) per year between $z = 1$ and $z = 3$. The observed time delay between the multiple images can be used to independently determine the Hubble constant (Vega-Ferrero et al. 2018). Over the life time of LSST, there is a 50% chance that at least one SNe has magnification higher than 1000; this means that the SNe is very close to a critical where the rapidly expanding photosphere will probe the structure of the overlapping microcaustics (see e.g., Schneider & Weiss 1986).

6.2. SNe at $z = 6$–12

Supernovae can be extremely luminous, and thus can be observed from large distances. Type Ia SNe have bolometric luminosities on the order of 10^{43} erg s^{-1} (see e.g., Sukhbold 2019), which is about 10^{10} more luminous than the Sun. Their corresponding absolute bolometric magnitudes is then ≈−20. The first type Ia SNe may have emerged after the formation of the first generation of white dwarfs. Since white dwarfs form from stars with a maximum mass of $\approx 10\,M_\odot$, the minimum time to form a white dwarf is the lifetime of a star with $\approx 10\,M_\odot$, that is only about 30 million years (see e.g., Table 1 of Windhorst et al. 2018). The delay time between the formation of the star and the SNe can be also relatively small and less than 1 Gyr as shown, for instance, by Maoz et al. (2010). Hence, SNe Ia can appear very early on in the Universe, although it is not known at what redshift the first SNe Ia may take place (after the formation of the white dwarf, the binary system needs to form).

In this section we consider the case of type Ia SN at $z = 6$–12 that can potentially be observed by JWST during the phase of

maximum brightness and are relevant for cosmology (at least to test for possible changes in the equation of state of dark energy. It should be kept in mind, however, that the rates assumed in this section are somewhat speculative, given the uncertainties in the formation rate of type Ia SNe at high-z and the unknown delay time between the formation of the star and the SNe which may significantly lower the rate at the highest redshift if the delay time is greater than 0.5 Gyr. Nevertheless, detection of such distant events through lensing could be used to infer the intrinsic rate (and delay times). Due to the larger size of the SNe during its maximum ($\sim 10^4\,R_\odot$), for a SNe sufficiently close to a caustic, the maximum magnification for a SNe can be about an order of magnitude lower than the magnification during the early phases of the explosion.

For SNe at redshift $z > 6$, JWST's Near-Infrared Camera (NIRCam) is arguably the instrument best suited to study these events. We consider the two wide filters F322W2 (covering the range $\sim$2.5–4 μm) and F444W ($\sim$3.8–4.8 μm). F322W2 is more sensitive than F444W, and it is ideal for SNe Ia at $z < 10$, while F444W is better suited for SNe Ia exploding at $z > 10$. At $z > 12$, the peak emission in the redshifted spectra of the SNe moves to longer wavelengths than those covered by the F444W. At $z < 8$, a SN explosion can be observed by JWST in the F322W2 filter even without the help of lensing.

We assume an absolute magnitude of $M_V = -19$ during maximum for type Ia SNe (Phillips 1993; Brown et al. 2010). The apparent magnitude is given by $M_V + DM(z) + k(z)$ where DM is the distance modulus and $k(z)$ is the k-correction. The distance modulus ranges from 48.85 at $z = 6$ to 50.58 at $z = 12$.

For the k-correction we estimate $k(z) = -1.88$ for F322W2 at $z = 8$ and $k(z) = -2.04$ for F444W at $z = 12$[2]. Without any lensing, a SN at $z = 6$ would have AB magnitude of 28 in F322W2 that could be detected with moderately deep observations. At $z = 12$ a type Ia SN would have AB magnitude of 29.5, so in order to see it with reasonable integration times in the less sensitive F444W would require lensing factors of 10 or more.

To estimate the rate of SNe at high redshift, we adopt a model that traces the SFR (Madau & Dickinson 2014), but normalized to a rate of $1.25 \times 10^{-4}\,\mathrm{Mpc}^{-3}\,\mathrm{yr}^{-1}$ for type Ia at $z = 1$ (Barbary et al. 2012). For this model, the volumetric rate of SNe Ia at $z = 6$ is $3 \times 10^{-5}\,\mathrm{Mpc}^{-3}\,\mathrm{yr}^{-1}$ and $5 \times 10^{-6}\,\mathrm{Mpc}^{-3}\,\mathrm{yr}^{-1}$ at $z = 12$. At the low end of the redshift interval, between $z = 6$ and $z = 7$ there are $365 \times 10^9\,\mathrm{Mpc}^3$. The total number of SNe explosions in this redshift interval is $\approx$1.7 million per year (after applying the time dilation factor), of which only $\sim$1 per year is expected to be lensed with total magnification $\mu > 100$. At lower magnifications, $\approx$20 SNe are expected to be lensed with a total magnification higher than 10, and with apparent magnitudes brighter than $m = 25.5$ mag. Even though it is extremely unlikely that JWST will find one of these SNe, a wide survey like LSST can potentially detect some of these lensed type Ia SNe (see Rydberg et al. 2018) that could be later followed up by JWST.

At the high end of the redshift interval between $z = 11$ and $z = 12$, there are $\approx 225 \times 10^9\,\mathrm{Mpc}^3$. The total number of SNe explosions in this redshift interval is $\approx$85 000 per year, of which $\approx$1 per 10 year is expected to be lensed by a factor larger than 100 (i.e., apparent magnitude AB $\approx$ 24 mag and hence detectable by LSST). The probabilities of detecting one of these SNe is low, but not negligible. A follow-up with JWST would provide valuable information about these first SNe.

[2] Patrick Kelly (priv. comm.).

As mentioned at the beginning of this subsection, it is not well known what is the abundance of type Ia SNe at high redshift. However, what is known is that other SNe types, such as core collapse, should be more abundant at $z > 6$, and some of them could equally be observed through extreme magnifications.

6.3. Luminous stars at $z \approx 1.5$–2.5

The low probability of observing high-z lensed SNe is partially a consequence of the relatively low number density of objects at high redshift. In order to compensate for the low probability of having a high magnification, there needs to be a large number of background objects in order to have a higher probability of one of them being close enough to a caustic. Bright stars (with absolute magnitude $M_V \approx -9$ or brighter) at $z \approx 1.5$–2.5 (i.e., with distance modulus in the range $\approx$45–47) could be observed with apparent magnitudes $\approx$27–29 if they are magnified by factors of a few thousand, hence within reach of future telescopes like JWST. As shown in Fig. 10, the optimal range for detecting events or objects that trace the SFR is between $z = 1.5$ and $z = 2.5$.

The mass-luminosity relation is well established for main sequence stars,

$$\frac{L}{L_\odot} = A\left(\frac{M}{M_\odot}\right)^\alpha, \tag{12}$$

where $A \approx 1$ and $\alpha = 3$–4. For massive stars, the slope α is shallower ($\approx$2.7 for stars with $10 < M_\odot < 50$ Vitrichenko et al. 2007) and for the most massive and luminous stars the luminosity can reach the Eddington limit, i.e., $L \propto M$ (see Eq. (3) and Sect. 3 in Windhorst et al. 2018). Owing to the high value of the exponent α, massive stars can be extremely luminous, reaching luminosities several million times that of the Sun. However, at cosmic distances ($z > 1$) and without the help of gravitational lensing magnification, their apparent magnitudes would be very faint (AB > 36) and well beyond the reach of current telescopes. With magnification factors of 1000 or more, some of these stars may get enough boost in their flux to be observable in deep observations. Small stars are more common and can be magnified by higher factors. The maximum magnification occurs when the star is at one radius distance from the caustic, $\mu_{\max} \propto 1/\sqrt{R}$, where R is the star radius. Since the luminosity of a star, L, scales as $L \propto R^a$, a lensed star has a maximum luminosity scaling as $L_{\max} \propto R^{a-0.5}$. Since $a > 1$, larger stars will always be more luminous at maximum magnification than smaller stars, and hence they are more likely to be observed during caustic crossing events even though they are more rare.

The recent observation of the Icarus star at $z = 1.49$ by Kelly et al. (2018) represents the first example of a redshift $z > 1$ star magnified by factors of several thousand. Kelly et al. (2018) estimates an absolute magnitude for Icarus of $M_V = -9 \pm 0.75$ mag. At the redshift of Icarus ($z = 1.49$) the distance modulus is 45.21 and the k-correction is estimated to be -1.1 magnitudes (in F125W) resulting in an apparent magnitude of $m = 35.11 \pm 0.75$ mag (without magnification). The magnification for the Icarus event was estimated to be $\mu \approx 2000$–4000, giving a boost of 8.25–9 magnitudes, so that the event could be observed with $m = 26.11$–$26.86(\pm 0.75)$[3].

In this section we address the question of how common these events may be. In order to do this, we perform a series

[3] As this manuscript was being finalized, Icarus experienced a new microlensing event in early June 2018 with similar peak brightness to that of the original event in 2016.

of approximations adopting a conservative approach in those assumptions that are more uncertain. The number of bright stars at a given redshift can be estimated as a constant fraction of the number of stars in a galaxy times the number of galaxies,

$$N_*(L_*, z) = \int \mathrm{d}L N_*(L_*, L) * \Psi_{\rm gal}(L, z) * F_{\rm SF}, \quad (13)$$

where $\Psi_{\rm gal}(L, z)$ is the luminosity function of galaxies and $N_*(L_*, L)$ is the fraction of stars with luminosity L_* in the galaxy with luminosity L. The factor $F_{\rm SF}$ accounts for the fact that massive stars have short lives, so they will be predominantly found in galaxies where star formation is taking place. Elliptical galaxies for instance have low SFRs, so the fraction of very luminous stars is expected to be much smaller. We set this fraction to $F_{\rm SF} = 0.5$ in the redshift interval considered. Observations show that the more numerous small galaxies ($M_* < 10^{10}\,M_\odot$) at $z = 1$–2 are mostly star forming galaxies (Ilbert et al. 2013, 2015; Muzzin et al. 2013).

For the luminosity function we adopt the best-fit Schechter parameters to the UV luminosity function in Reddy & Steidel (2009) at $z \approx 2$. After integrating the Schechter function between absolute magnitudes −23 and −16 we find a number density of ≈0.08 galaxies per Mpc^3 (and a factor ≈2 times less if the integral is carried out down to absolute magnitude −17). Assuming to the first order no evolution in the luminosity function between $z = 1.5$ and $z = 2.5$, the total number of galaxies in this redshift interval is $\approx 4 \times 10^{10}$.

To estimate the number of bright stars that may experience magnifications higher than 1000 (through temporary microlensing episodes), first we compute the number of galaxies that are intersecting a caustic. We assume that a galaxy is intersecting a caustic if its center is at a distance from the caustic such that the magnification of its nucleus is ≈100. This means that the galaxy will produce multiple images (usually three or more), with two of the counterimages carrying most of the magnification (i.e., two images, each with magnification of ≈50 when distance from the caustic is approximately the radius of the galaxy). This is a conservative assumption since in most cases galaxies intersecting caustics are observed with average magnification ≈30–50 (e.g., the galaxy of Icarus). Since the average of the optical depth for $\mu > 100$ between $z = 1.5$ and $z = 2.5$ is $\approx 10^{-7}$, we expect ≈4000 galaxies intersecting caustics between $z = 1.5$ and $z = 2.5$. For each one, a portion of the galaxy at any time will be crossing a caustic. If a star is at a distance of a few parsecs from the caustic (i.e., at separations of 1 mas or less) where the magnification from the macromodel can reach factors of several hundred (see Sect. 2) and microlensing may become important, temporarily boosting the magnification to factors of a few thousand. At these magnifications, a very luminous star at $z \approx 1.5$–2.5 could be elevated to magnitudes brighter than 28. If a galaxy of radius $R_{\rm gal}$ is intersecting a caustic, the probability that a given star in that galaxy is within a distance d (typically a few parsecs) of the caustic is $d/R_{\rm gal}$, which is $\sim 10^{-3}$ for a galaxy of a few kpc in radius. Hence, approximately one in a thousand stars in the galaxy has a chance of intersecting a microcaustic over the lifetime of JWST. If that galaxy has a number of super-luminous stars (with luminosities exceeding $10^6\,L_\odot$) similar to our Galaxy (i.e., ∼100), this means that for every ten galaxies intersecting a caustic, one of them would have a super-luminous star close enough to the caustic to be observable. This puts the number of potentially observable background stars at ≈400. This is a very rough estimate and ignores many subtleties.

An alternative and perhaps more precise estimation can be obtained as follows. We assume that the number of bright stars in a given galaxy correlates with its luminosity. We assume that there are n_* bright stars in a faint galaxy with absolute magnitude −16 (the number n_* will be estimated later). The number of stars in a galaxy with absolute magnitude M will then be

$$N_*(L_*, L) = n_* 10^{-(M+16)/2.5}, \quad (14)$$

where we have used the relation between magnitude and luminosity, $M - M_o = -2.5 \log(L/L_o)$ with $M_o = -16$ and L_o the corresponding luminosity of a galaxy hosting n_* bright stars. The number of very luminous stars in our Galaxy and surroundings (LMC and SMC) with luminosities above $10^6\,L_\odot$ or bolometric absolute magnitudes exceeding −10 mag is of the order of 100 (see e.g., Humphreys & Davidson 1979; Hamann et al. 2006; Crowther et al. 2010, 2016; Hainich et al. 2014). The total number is probably a factor of a few higher, but we adopt the conservative number of 100 for a galaxy like the Milky Way (plus its satellites). By requiring Eq. (14) to reproduce the number of bright stars in our Galaxy ($M = -21$ mag, $z = 0$, $N_* \approx 100$ bright stars), we find that $n_* \approx 2.5$. This is well motivated since the specific SFR of the Milky Way falls within the typical values for galaxies at $z = 0$ (Pérez-González et al. 2008; Licquia & Newman 2015). We should note, however, that this simple recipe fails in cases like the Large and Small Magellanic Clouds (LMC/SMC), where a significantly smaller number of bright stars would be expected, while the number of observed extremely bright stars are as numerous in the LMC/SMC as in our own Galaxy. This may be consequence of interactions between the LMC/SMC and our Galaxy, which may trigger star formation episodes at the perigalactic passages of the LMC/SMC with our Galaxy (Harris & Zaritsky 2004). The expression above ignores such temporary episodes or interactions, and hence should be considered a conservative approximation. After integrating over the galaxies brighter than magnitude $M = -16$, we find ≈1 bright star per Mpc^3, or $\approx 5 \times 10^{11}$ bright stars between $z = 1.5$ and $z = 2.5$, of which ≈50 000 will be in areas where the magnification is higher than 100 and can momentarily undergo microlensing events by massive stars or BHs that boost the total magnification to $\mu > 1000$, hence potentially observable by current telescopes. At macro-magnifications higher than 1000, the number drops by a factor of 100, so we expect ≈500 stars at macromodel magnification $\mu > 1000$ between $z = 1.5$ and $z = 2.5$. In the context of the Icarus star with a macromodel magnification ≈500 there are four times the number of bright stars (brighter than absolute bolometric magnitude −10 mag) moving in areas with macromodel magnification higher than 500; in other words, ≈2000 over the entire sky or ≈1 star per 20 square degrees. These stars would appear very faint, but within reach of deep observations at AB ∼ 28–29 magnitudes. Occasionally, these stars may experience microlensing events where their flux may increase between 1 and 3 magnitudes. The duration of these microlensing events depends on the cluster microlens, the macromodel magnification, the relative velocity between the bright star and the caustic, and the angle between the star motion and the caustic orientation. For typical configurations, this timescale is on the order of days to weeks.

The assumptions made above regarding the number of bright stars is quite uncertain[4], but we should expect a reasonable number of stars that could be followed up. Icarus represents the first entry in a list of background stars at cosmic distances grazing

[4] Many of these extremely luminous stars can be also variable, which adds another element of uncertainty since a star may be super-luminous only during a short period of time.

in a field of microcaustics at high magnification. This list will grow quickly in the near future. Having a census of background stars that are moving near macrocaustics is of great interest since it opens the door to the statistical study of the composition of the microlenses. Whether these microlenses are simply stars and remnants or more exotic microlenses like primordial black holes is something that could be determined with sufficient statistical information about the fluctuations in flux of the background stars (assuming time variability and microlensing features in the light curves can be distinguished).

The planned Deep Drilling Fields of LSST (Ivezic et al. 2019; Abell 2009) will cover 38.4 deg^2 in four separate fields to a much greater depth (although how deep is still undecided) and also better cadence. These fields are expected to reach magnitude 28 mag in r in the coadded data (and single night depths of ≈26 mag). At such depths, these fields should contain approximately two stars similar to Icarus. These stars will exhibit changes in their flux due to microlensing on timescales of days to weeks. They can also be followed later by JWST during one of the microlensing peaks to study their spectrum in detail.

Although not discussed in this work, observations of caustic crossings by Pop III stars will also be possible with future large telescopes like the ELT (in construction and expected to be operative in the mid-2020s) or the proposed LUVOIR (~2035).

6.4. Pop III stars

At even higher redshifts, Pop III stars can be used as powerful and very small background sources (Bromm et al. 2001; Windhorst et al. 2018). These stars are supposed to be extremely luminous owing to their high masses (Bromm et al. 2001; Susa et al. 2014), although recent work suggests that they may be less massive than previously thought (Ishigaki et al. 2018). The higher redshift has the advantage of having maximum optical depth, and the critical curves that form farther away from the centers of the lenses. At these separations, the impact of microlenses is smaller, thus allowing for higher magnification (at the expense of having fewer short episodes of microlensing). Despite luminosities that can reach tens of millions $L_\odot$, the large distance makes their detection extremely challenging. Even deep observations made with powerful telescopes like the JWST may not reach the depth required to see them (Rydberg et al. 2013). However, thanks to gravitational lensing, some of these stars may get a boost of several magnitudes that could make them detectable. As discussed in Windhorst et al. (2018), JWST's NIRCam will be able to look for highly magnified Pop III stars as caustic crossing events. JWST is expected to routinely reach AB ≈ 28.5–29 in medium-deep observations. In this section we present an order of magnitude estimation of the probability of observing these events. The number density of faint galaxies observed in deep observations (e.g., in the GOODS and HUDF fields) is approximately between ≈1 and a few galaxies per arcmin^2 in the redshift interval $z = 8-12$ (Finkelstein et al. 2015). Similar results are derived when observing faint galaxies through powerful gravitational lenses like the Hubble Frontier Fields clusters (Ishigaki et al. 2015). The number of galaxies in the entire sky above $z = 8$ (and within our horizon) can then be several hundred million galaxies and perhaps up to several billion. With such large numbers, even a low probability of intersecting a caustic results in a significantly large number of galaxies.

From the previous section, microlenses negate the possibility of observing these stars at magnifications higher than 10^5 (assuming that the observations are made in regions where the contribution form microlenses is small). For critical curves at high redshift ($z_s > 7$), we expect the role of microlenses to be smaller than for sources at $z_s \approx 1-2$. How high the maximum magnification is depends on how small the role of microlenses is at the position of critical curves for sources at high redshift. To estimate the surface mass density of microlenses at the position of a high-redshift critical curve, we compute the critical surface mass density for a lens at $z = 0.4$ and a source at $z = 10$, $\Sigma_{\rm crit} = 1.8\times10^3\,M_\odot\,{\rm pc}^{-2}$, and assume that at the position of the critical curve the convergence typically adopt values close to $\kappa \approx 0.5$[5], that is $\Sigma = \kappa\times\Sigma_{\rm crit} \approx 900\,M_\odot\,{\rm pc}^{-2}$. Adopting a mass-to-light ratio of 200 (see e.g., Girardi et al. 2002), we obtain that the surface mass density of stars is $\Sigma \approx 4.5\,M_\odot\,{\rm pc}^{-2}$, close to the low-$\Sigma$ model adopted earlier with $\Sigma \approx 6.33\,M_\odot\,{\rm pc}^{-2}$. Hence, if the surface mass density of microlenses at the position of high-redshift critical curves is in the range of a few $M_\odot\,{\rm pc}^{-2}$, we should expect maximum magnifications not exceeding a few tens of thousands. We adopt the value $\mu_{\rm max} = 5\times10^4$ as a reasonable limit (for a star with a radius several tens of times the radius of the Sun; see the typical radii for these stars in Tables 2–4 in Windhorst et al. (2018). If we set the detection limit to $m = 28.5$ mag, (a reasonable limit for medium-deep observations with NIRCam in JWST), this implies that only sources brighter than $m = 40.25$ mag can be observed through caustic crossing events. Following Windhorst et al. (2018) Sect. 4.4, we find the minimum mass that can be observed at this maximum magnification. These are are $70\,M_\odot$, $190\,M_\odot$, and $250\,M_\odot$ at $z_s = 7$, $z_s = 12$, and $z_s = 17$, respectively (these masses become 27, 36, and 42 in the much shorter but brighter phase of the asymptotic giant branch). Any mass below these limits would require a magnification higher than $\mu_{\rm max}$.

Visbal et al. (2018) estimates that at $z < 20$, the SFRD is already $\approx10^{-4}$–$10^{-3}\,M_\odot$ Mpc^{-3} yr^{-1}. A similar value is obtained by Sarmento et al. (2018), but in the redshift range $z = 8-12$. At these high redshifts, the SFR is expected to be dominated by the production of Pop III stars (see e.g., Sarmento et al. 2018). We make the simplifying assumption that most Pop III stars being produced at these redshifts are at a characteristic mass scale of $100\,M_\odot$. This is close to the peak of the mass function found by Susa et al. (2014) for the first stars (see also Hirano et al. 2014). This leads to $\approx10^{-6}$–10^{-5} stars Mpc^{-3} yr^{-1} or 2–20 Mpc^{-3} during the typical lifetime of a massive Pop III star ($t_{\rm mean} \approx 2$ Myr or d$z \approx 0.05$ at $z = 10$). This volumetric density translates into a surface density of ≈0.05–0.5 stars with $100\,M_\odot$ per mag arcsec^2 between $z = 10$ and $z = 10.05$. This estimate is relatively uncertain, but is consistent with the observational limit in the near-IR of $m \gtrsim 31$ mag arcsec^{-2} (see Windhorst et al. 2018). Adopting the upper value (i.e., 0.5 stars per arcsec^2 in d$z = 0.05$ or 10 stars per arcsec^2 and redshift interval d$z = 1$), the surface brightness from such a density and redshift interval would be $m \approx 41$ mag arcsec^{-2}. After integrating over a redshift interval between $z = 7$ and 20 we estimate the surface brightness would be AB ≈ 36 mag arcsec^{-2}, comfortably below the limit $m \approx 31$ mag arcsec^{-2}. The number density of stars can be increased by a factor 100 and still not exceed this limit. Hence, as an upper limit we estimate a maximum number of stars of 1000 arcsec^{-2} and per redshift interval at $z = 10$. This estimate is in good agreement with the value estimated by Windhorst et al. (2018), who estimate the surface density of bright ($L > 10^6\,L_\odot$) Pop III stars at $z > 7$ and brighter than AB ≈ 38.5 mag to be in the range 1–1000 per arcsec^2

[5] Near critical curves, arcs typically have small radial magnification ($|1-\kappa+\gamma| \approx 1$) and high tangential magnification ($|1-\kappa-\gamma| \approx 0$).

Adopting the compromise value of 10 stars per arcsec^2 or $\approx 5 \times 10^{12}$ stars in the observable sky, the number of stars with $\mu > 100$ would then be $\approx 5 \times 10^6$ or ≈ 100 per square degree. At $\mu > 1000$, the number of lensed stars would be 1 per square degree. As a consequence, a surveyed area of 1 square degree should contain at least one Pop III star at $z > 7$ that is lensed by a factor $\mu > 1000$ with $m < 31$ mag, and within reach of the JWST. The number density of extremely lensed Pop III stars can be increased by targeting known lenses at intermediate redshifts. Windhorst et al. (2018) estimates that the rate of extreme lensing events can be as high as 1 yr^{-1} and per 3 massive lenses surveyed. As discussed in Windhorst et al. (2018), Pop III stars are short-lived but their remnants are expected to host bright accretion disk that could be equally lensed. These accretion disks are expected to live much longer than the Pop III stars, so their number density (and rate of lensed events) may be correspondingly higher.

6.5. Gravitational waves

Broadhurst et al. (2018) suggest that the low-frequency events observed by LIGO can be reinterpreted as gravitational waves (GWs) originating at $z > 1$ that are magnified by factors $\mu > 100$. Broadhurst et al. (2018) relies on the same lensing probability presented in this work, so we do not repeat the calculations presented in that work. Here, we briefly discuss their result and the impact of microlenses near critical curves in the context of GWs, which was not explored in Broadhurst et al. (2018). As proposed by Broadhurst et al. (2018), the low-frequency GWs observed by LIGO would be produced by binary black holes (BBHs) with intrinsic chirp masses of about 30 solar masses (which is a factor ≈ 2 times lower than the values inferred by the LIGO collaboration) and that they would be redshifted so the observed chirp mass (derived from the observed GW frequency and its derivative) appears a factor of 2 higher. In order for the strongly lensed GW interpretation to work in the context of the LIGO observations, the intrinsic rate of BBH coalescence needs to be $\approx 3 \times 10^4$ events per year and per Gpc^3 and between $z = 1.5$ and $z = 2.5$. At this rate, the number of events in this redshift interval is $\approx 1.5 \times 10^7$ per year, of which ≈ 4 events per year will be magnified by factors above $\mu = 100$, and hence are within reach of the LIGO detectors. Whether the low-frequency events detected by LIGO are gravitationally lensed, as proposed by Broadhurst et al. (2018), is something that can be tested with future data; at these high magnifications we would expect to see multiple images of the same source with similar magnifications. Broadhurst et al. (2018) suggests that many of these multiple images are not observed since they fall within the detection limit of LIGO (due to the time delay between images and Earth's rotation). At the high magnifications considered in Broadhurst et al. (2018) ($\mu > 100$), the effects of microlensing are expected to be important for small sources. However, given the long wavelength of gravitational waves we need to consider wave optics and microlenses with masses below 30 $M_\odot$ to produce negligible magnification of the GW. The maximum magnification of a GW by a lens of mass M is $\mu_{\max} \approx x/(1 - e^{-x})$, where $x = (4/\pi) \times 10^5 (M/M_\odot)(\nu/Ghz)$ (Schneider et al. 1999). For a typical GW detected by LIGO ($\nu \sim 250$ Hz) and a microlens of mass $<10\, M_\odot$, the maximum magnification is $\mu_{\max} \approx 1$. At masses above 100 $M_\odot$ and frequencies of 250 Hz, the maximum magnification of GW scales as $\mu(M > 100\, M_\odot)_{\max} \approx 3 \times M/100\, M_\odot$. At higher frequencies or much higher masses, the exponential part tends to zero and $\mu_{\max} \approx x$. Even though the magnification can be modest or negligible, there can be interference effects that can leave observable signatures even if the microlens mass is relatively low (see e.g., Lai et al. 2018). As discussed in Diego et al. (2018), a microlens with mass M near a caustic where the macromodel tangential magnification is μ_t behaves as a microlens with effective mass $\mu_t M$. This means that at magnification factors of several hundred, a microlens with a mass of a few solar masses behaves as a microlens with an effective mass of several hundred solar masses. Hence, interference patterns (and some extra magnification) in the GW may occur if the GW is strongly magnified near a caustic, which is disturbed by microlenses. Figure 11 shows examples of time delays for a massive microlens of mass 400 $M_\odot$ (small panels on the left) and for a smaller microlens of mass 10 $M_\odot$ (large panel on the right). In both cases, time delays between 1 and 100 millisecond can be produced. Since the inverse of the frequency is also of order 1 millisecond, interference can take place in both cases. Moreover, if the lens plane is populated by many microlenses (as expected), Diego et al. (2018) shows how there can be a greater number of multiple images, and also the time delay between them resulting in complex interference patterns. From the figures in the left panel, a trend is observed between the time delay and the magnification of the macromodel. For a fixed source position (with respect to the center of the microcaustic), the time delay grows as $\sqrt{\mu_{\rm macro}}$. This means that relatively small microlenses sufficiently close to a critical curve can result in time delays that are large enough for interference to take place. On the other hand, as shown by Oguri et al. (2018), the size of the microcaustics shows a weaker dependency with $\mu_{\rm macro}$, so in order for these effects to be important with small microlenses, the separations between the source and the microcaustic must be small (almost independently of the macromodel magnification). These smaller separations are required to produce situations like the one shown on the right side of Fig. 11, where the probability of intersecting the smaller microlens is lower than in the panels on the left side (with a higher mass and hence larger microcaustics). Hence, even though measurable interference patterns can be produced by stars of a few solar masses, their associated caustics are smaller. On the other hand, their larger number (compared with 400 $M_\odot$ stars) compensates for the smaller size of their caustics, thus increasing the probability of intersecting a small caustic. The superposition of multiple caustics also helps to increase the time delay, so microlenses near macromodel critical curves are expected to produce a rich phenomenology in terms of time delay and interference patterns. This regime is studied in more detail in Diego et al. (2019).

Finally, GW from binary neutron stars (BNS) mergers or binary neutron star and black hole (BNSBH) mergers can be equally magnified. In the case of BNS mergers, their larger number increases the probability of being lensed, but the lower masses limits the maximum distance at which BNS can be observed without lensing. Since the minimum magnification needed to detect a lensed GW effect scales as $(D_l(z)/D_o)^2$ (the strain on a lensed GW scales as $\sqrt{\mu}/D_l(z)$), to see an event similar to GW170817 (which is at a distance of 40 Mpc) but at a distance of 1000 Mpc ($z \approx 0.2$) would require a magnification of $\mu \approx 600$. The optical depth at $z_s = 0.2$ for $\mu \approx 600$ is extremely small ($<10^{-8}$). At higher redshifts, the increase in volume and intrinsic rate of BNS mergers is not enough to compensate the rapid increase of magnification needed to promote the event above the detection threshold, so the prospects of seeing a lensed BNS are very dim (except for modest magnifications between $\mu = 1$ and $\mu = 2$, where there may actually be a chance of seeing such events). In the case of BNSBH mergers, we expect a similar scenario; even though the mass of the BH component may be high, the intrinsic chirp mass remains low making the

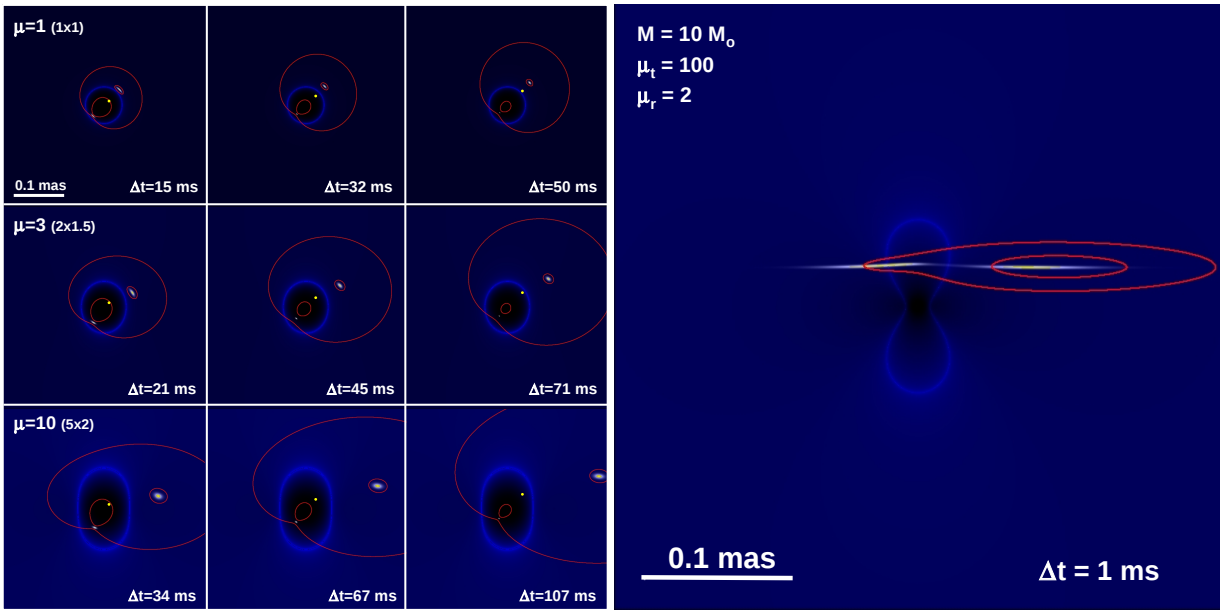

Fig. 11. Time delay between the brightest counterimages (GW) produced by microlensing. The group of nine figures on the left is for a microlens of mass 400 $M_\odot$ at $z = 0.5$ and a source at $z = 1.5$. Magnification in the lens plane is shown in blue/black. The critical curve can be seen as the bright blue contour. The multiply lensed images are shown in white. Each row corresponds to a different magnification. The numbers in parentheses next to the magnification indicate the tangential and radial magnifications respectively (i.e., $\mu_t \times \mu_r$). The microlens is isolated in the *top row* ($\mu = 1$) and embedded in a potential from a macrolens with magnification $\mu = 3$ (*middle row*) and $\mu = 10$ (*bottom row*). Each column represents a different source position (marked with a small yellow dot). The red contours are isochrones. The time difference between isochrones is given in the bottom right corner of each panel. The *right figure* shows one example of a much lighter microlens (with mass 10 $M_\odot$), but much closer to the macrolens critical curve. In particular, the macro-magnification at this position is 200 (100 in the tangential direction and 2 in the radial direction). The redshifts are the same as in the left panels. For this configuration, the time delay between microimages is on the order of 1 millisecond. The source is much closer to the center of the caustic, hence the smaller time delay. Since the microcaustic is smaller, at separations from the center of the microcaustic similar to those on the left side, no multiple images are produced.

detectability of strongly lensed GW from BNSBH very unlikely. However, as noted by Diego et al. (2019), the lower chirp mass produces strains with higher frequencies, and it is at these frequencies where the effects of interference from stellar mass microlenses can be appreciated best.

Similarly, GWs from SNe explosions are expected to be detectable only within our local neighborhood (tens of kpc), so the probability of them being lensed is even more remote.

7. Discussion

Massive stars are usually born after fragmentation of large clouds (Tan et al. 2014). This means that massive stars tend to form groups or small clusters (Susa et al. 2014). If one of these minihalos is close to a lens caustic, the probability that one of the stars in the group will cross a caustic grows compared with the case where luminous stars are isolated. A good example is the R136 group in the heart of the Tarantula Nebula (in the LMC) which contains several of the most luminous stars known. If one of these groups, with typical sizes of a few parsecs, is moving across a field of micro-caustics, each bright star in the group can undergo multiple microlensing episodes over a period of time greatly increasing the rate of events. This may create problems when interpreting the data since the number density of bright stars would degenerate with the number density of microcaustics. In the starburst region of the Tarantula Nebula, Schneider et al. (2018) finds a flattening in the slope of the IMF at the high-mass end suggesting that very massive (and hence luminous) stars may be relatively more abundant than would be inferred from standard IMF. Hence, the number of stars between $z = 1.5$–3 with macromodel magnification higher than 500 could be more than our estimated number of 2000 based on standard IMF functions.

We have also considered a different type of luminous source, SNe explosions. In particular, type Ia SNe are of great interest for cosmology since they can be used as standard candles to constrain the evolution of the equation of state of dark energy. Even though the explosions from SNe can release tremendous amounts of energy (Woosley et al. 2002), detecting them at redshifts higher than $z = 1$–2 is challenging for current survey telescopes (typically with modest apertures) without the aid of lensing. In a recent work, Rydberg et al. (2018) consider strongly lensed distant SNe that could be observed by LSST. In their work, they also discuss the possibility of detecting Pop III SNe at $z < 7$. In particular, they estimate that LSST may detect 1–2 Pop III between $z = 5$ and $z = 7$ in the deep survey (magnitude AB ≈ 28.5 in 10 deg^2) although with a much more modest magnification factor than considered in this work (see also Oguri & Marshall 2010, for an estimate at lower redshifts). Studying in detail the light curves of strongly lensed SNe can be used to infer the underlying population of underlying microlenses (Rauch 1991)

At redshifts $z > 7$, the first stars could be observed through caustic crossing events with telescopes like JWST, as proposed by Windhorst et al. (2018). It is believed that at the end of their life, Pop III stars with masses in the range 15–40 $M_\odot$ explode as core-collapse supernovae (CC SNe) and with masses in the range 140–260 explode as very energetic pair-instability (PI) SNe, with luminosities up to 100 times higher than Type Ia or Type II SNe (Heger & Woosley 2002) and leaving no remnant. At masses between 40 and 140 $M_\odot$ and above $\approx$260 $M_\odot$ the star collapses

into a proto-neutron star or BH without producing a SN. Since SNe explosions from Pop III stars are considerably brighter than regular SNe, they can be detected by JWST (Wise & Abel 2005; Whalen 2013; Whalen et al. 2013). Both JWST and WFIRST should be able to detect SNe explosions from the first stars up to $z \approx 20$ (Whalen et al. 2013). At lower redshifts (5–7), telescopes like LSST could also see lensed SNe from late Pop III (Rydberg et al. 2018).

As pointed out by Diego et al. (2018), microlenses disrupt the caustics from macromodels and reduce the maximum magnification attainable by a small background source. The most favorable scenario for observing faint background sources is in regions around the lens where the contribution from microlenses is small. If the background source is at high redshift, the magnification required to compensate the increase in luminosity distance will be higher, but the contribution from microlenses will also be smaller since the critical curves move outward. Also, microlenses on the side of the critical curve with negative parity can magnify by larger factors at the expense of having longer periods of low magnification, where the background source apparently vanishes from the data. Due to microlenses, radial curves are less likely to produce extreme magnification, since the optical depth of microlensing is larger, although they are more likely to produce fluctuations with more modest magnification factors μ of a few tens or few hundreds. We can in principle estimate the maximum magnification expected at a given observed lensing arc based on an estimate of the surface mass density and a lens model. Similarly, if a given arc known to intersect a caustic is monitored for some time, we can infer the surface mass density of microlenses including nonluminous ones such as remnants or primordial black holes (PBH). The most favorable situation is for arcs intersecting caustics at high redshift. If one of these arcs is known to contain a sufficiently high number of bright sources, for instance Pop III stars, and no extreme events are observed, we can directly use the lack of fluctuations to impose tight constraints on the abundance of PBH (assuming we are observing far enough from the center of the lens so the contribution from the stars and remnants is small, and a population of PBH would dominate the surface mass density of microlenses). Bright background sources at high redshift are more likely to undergo extreme magnification than the similar sources at low redshift since the corresponding critical curves are observed in regions of the lens that are significantly less influenced by microlenses. Bright Pop III stars at high redshift are hence ideal targets for extreme magnification (Windhorst et al. 2018).

Although not explored in detail in this paper, the role of microlenses can be important in the case of strongly lensed GW. If the volumetric rate of BBH mergers is on the order of 10^4 per year and Gpc^3, we have shown how we should expect approximately 1 GW per year magnified by large factors of 100 or more. At these magnifications, the amplification takes place near a critical curve where microlenses play an important role. We discussed how microlenses have a minor impact on the overall magnification of the GW, but for the right mass range, time delays of order 1 millisecond can produce interference patterns that could be observed both in the spectrum of the GW, but also as a different signal in the detectors since they are probing different portions of the caustic region (although this second effect is expected to be much smaller than the first). We showed how in the case of a single microlens of a few solar masses near a critical curve, where the magnification can be more than 100, time delays of order 1 millisecond are possible. In a more realistic scenario, where the lens plane is populated by many microlenses, their corresponding microcaustics can overlap in the source plane resulting in larger time delays and complex interference patterns.

8. Conclusions

Giant arcs lurk around critical curves of massive galaxies, galaxy groups, and clusters. These arcs can contain very luminous stars or SNe, and normally intersect the critical curves. If a super-luminous star, Pop III star, or SNe happens to lie at a few pc from the caustic, it can be magnified by factors of hundreds. At these magnifications, microlenses can distort significantly the pdf of the magnification (compared with the case without microlenses) producing temporary episodes of extreme magnification with $\mu > 1000$. The length of these events depend on the relative velocity between the source and the caustic. The presence of microlenses also disrupts the caustic of clusters, preventing a caustic crossing event from reaching magnifications of millions. Instead, when microlenses are present, the maximum magnification of the caustic crossing event is in the range of tens of thousands. We computed the probability of extreme magnification with a state of the art mass function and analytical models for the individual lenses, which allowed us to reach high spatial resolution. The magnification was computed with inverse ray tracing, the appropriate algorithm to account for the total flux of extremely magnified, unresolved images. Our results are conservative since we ignored projection effects that increase the optical depth of lensing. We also ignored the role of baryons that can increase the optical depth, especially from smaller halos, by up to a factor 2. When microlenses are present, a trade off in the magnification between the macromodel caustic and the microcaustics results in an increase in the probability of magnifications around a few thousand, which can be a factor ~ 2, although the exact number depends on the macromodel and microlensing configurations. We find that the probability of having extreme magnification higher than 1000 for a source at $z = 2$ is $\approx 3 \times 10^{-9}$. A source or event population with a volumetric density of $10^{-3}\,\mathrm{Mpc}^{-}3\,\mathrm{yr}^{-1}$ (comparable to the rate of SNe of all types) would then produce approximately one event per year with $\mu > 1000$ at $z = 2$. When accounting for microlensing near critical curves, this number can be increased by a factor of ≈ 2. Among the most promising targets are Pop III stars for which we predict that as many as 1 star per square degree could have magnification higher than 1000. In combination with sporadic boosts from microlensing, these stars can be found with JWST in deep fields reaching AB $\approx$ 30. Extremely magnified super-luminous stars at $z = 1.5-3$ should be also relatively abundant. We predict that there should be $\approx$2000 stars in the entire sky and in the redshift interval $z = 1.5-3$ with macromodel magnifications similar to that of Icarus. The counterimages of these stars will be normally unresolved and affected by microlensing events. At typical relative velocities and microlensing surface mass densities, these events should take place every few years and with magnifications of a few thousand over periods of days to weeks (Diego et al. 2018), sufficient to be observed by future telescopes.

Surveys like LSST will cover about half the sky with a cadence of a few days, sufficient to see strongly lensed SNe above $z = 1$. We estimate that about 5 strongly lensed SNe ($\mu > 100$) per year above $z = 1$ should be discovered by LSST. On the Deep Drilling Fields, we estimate LSST should detect approximately 2 stars, similar to Icarus. These stars can be later followed up by JWST to get their spectrum and identify small fluctuations in the flux due to microlensing.

If the rate of GWs at $z > 1$ is approximately one order of magnitude higher than the inferred rate from naively

extrapolating the local rate, experiments like LIGO, which are instantaneously sensitive to the entire sky (neglecting the geometric factor), should routinely detect strongly lensed GWs with magnification factors of 100 or more. At these magnifications, small microlenses of a few solar masses can trigger interference patterns in the GW that could be detected in their spectrum.

Acknowledgements. The authors wish to thank Mitch Struble, Ravi Sheth, Gary Bernstein, and Ariel Goobar for the very useful comments; Rogier Windhorst for carefully reviewing the manuscript; and Pat Kelly for important suggestions, and for providing the *k*-corrections for JWST. The author also wishes to thank the referee for raising interesting questions about the manuscript. J.M.D. acknowledges the support of projects AYA2015-64508-P (MINECO/FEDER, UE), funded by the Ministerio de Economia y Competitividad. J.M.D. acknowledges the hospitality of the Physics Department at the University of Pennsylvania for hosting him during the preparation of this work.

References

Barbary, K., Aldering, G., Amanullah, R., et al. 2012, ApJ, 745, 31
Bersten, M. C., Folatelli, G., García, F., et al. 2018, Nature, 554, 497
Blackburne, J. A., Pooley, D., Rappaport, S., & Schechter, P. L. 2011, ApJ, 729, 34
Broadhurst, T., Diego, J. M., & Smoot, III, G. 2018, ArXiv e-prints [arXiv:1802.05273]
Bromm, V., Kudritzki, R. P., & Loeb, A. 2001, ApJ, 552, 464
Brown, P. J., Roming, P. W. A., Milne, P., et al. 2010, ApJ, 721, 1608
Chang, K., & Refsdal, S. 1979, Nature, 282, 561
Chang, K., & Refsdal, S. 1984, A&A, 132, 168
Chen, W., Kelly, P. L., Diego, J. M., et al. 2019, ArXiv e-prints [arXiv:1902.05510]
Collett, T. E. 2015, ApJ, 811, 20
Crowther, P. A., Schnurr, O., Hirschi, R., et al. 2010, MNRAS, 408, 731
Crowther, P. A., Caballero-Nieves, S. M., Bostroem, K. A., et al. 2016, MNRAS, 458, 624
Diego, J. M., Kaiser, N., Broadhurst, T., et al. 2018, ApJ, 857, 25
Diego, J. M., Hannuksela, O. A., Kelly, P.L., et al. 2019, A&A, submitted [arXiv:1903.04513]
Dobler, G., & Keeton, C. R. 2006, ApJ, 653, 1391
Finkelstein, S. L., Ryan, Jr., R. E., Papovich, C., et al. 2015, ApJ, 810, 71
Foxley-Marrable, M., Collett, T. E., Vernardos, G., Goldstein, D. A., & Bacon, D. 2018, MNRAS, 478, 5081
Frye, B. L., Pascale, M., Zitrin, A., et al. 2019, ApJ, 871, 51
Fukugita, M., Futamase, T., Kasai, M., & Turner, E. L. 1992, ApJ, 393, 3
Garnavich, P. M., Tucker, B. E., Rest, A., et al. 2016, ApJ, 820, 23
Girardi, M., Manzato, P., Mezzetti, M., Giuricin, G., & Limboz, F. 2002, ApJ, 569, 720
Goldstein, D. A., Nugent, P. E., Kasen, D. N., & Collett, T. E. 2018, ApJ, 855, 22
Goobar, A., Amanullah, R., Kulkarni, S. R., et al. 2017, Science, 356, 291
Griffiths, A., Conselice, C. J., Alpaslan, M., et al. 2018, MNRAS, 475, 2853
Guerras, E., Mediavilla, E., Jimenez-Vicente, J., et al. 2013, ApJ, 764, 160
Hainich, R., Rühling, U., Todt, H., et al. 2014, A&A, 565, A27
Hamann, W.-R., Gräfener, G., & Liermann, A. 2006, A&A, 457, 1015
Harris, J., & Zaritsky, D. 2004, AJ, 127, 1531
Heger, A., & Woosley, S. E. 2002, ApJ, 567, 532
Hilbert, S., White, S. D. M., Hartlap, J., & Schneider, P. 2007, MNRAS, 382, 121
Hilbert, S., White, S. D. M., Hartlap, J., & Schneider, P. 2008, MNRAS, 386, 1845
Hirano, S., Hosokawa, T., Yoshida, N., et al. 2014, ApJ, 781, 60
Humphreys, R. M., & Davidson, K. 1979, ApJ, 232, 409
Ilbert, O., McCracken, H. J., Le Fèvre, O., et al. 2013, A&A, 556, A55
Ilbert, O., Arnouts, S., Le Floc'h, E., et al. 2015, A&A, 579, A2
Inada, N., Oguri, M., Shin, M.-S., et al. 2012, AJ, 143, 119
Ishigaki, M., Kawamata, R., Ouchi, M., et al. 2015, ApJ, 799, 12
Ishigaki, M. N., Tominaga, N., Kobayashi, C., & Nomoto, K. 2018, ApJ, 857, 46
Ivezic, Z., Tyson, J. A., Abel, B., et al. 2019, ApJ, 873, 111
Kayser, R., Refsdal, S., & Stabell, R. 1986, A&A, 166, 36
Keeton, II, C. R. 1998, PhD Thesis, Harvard University, USA
Keeton, C. R., Kuhlen, M., & Haiman, Z. 2005, ApJ, 621, 559
Kelly, P. L., Rodney, S. A., Treu, T., et al. 2015, Science, 347, 1123
Kelly, P. L., Diego, J. M., Rodney, S., et al. 2018, Nat. Astron., 2, 334
Kochanek, C. S. 1996, ApJ, 466, 638
Kochanek, C. S., & White, M. 2001, ApJ, 559, 531
Lai, K.-H., Hannuksela, O. A., Herrera-Martín, A., et al. 2018, Phys. Rev. D, 98, 083005
Lapi, A., Negrello, M., González-Nuevo, J., et al. 2012, ApJ, 755, 46
Lee, M. H., & Paczynski, B. 1990, ApJ, 357, 32
Licquia, T. C., & Newman, J. A. 2015, ApJ, 806, 96
LSST Science Collaboration (Abell, P. A., et al.) 2009, ArXiv e-prints [arXiv:0912.0201]
Madau, P., & Dickinson, M. 2014, ARA&A, 52, 415
Maoz, D., Sharon, K., & Gal-Yam, A. 2010, ApJ, 722, 1879
Martel, H., & Premadi, P. 2008, ApJ, 673, 657
Meneghetti, M., Bartelmann, M., & Moscardini, L. 2003, MNRAS, 346, 67
Miralda-Escude, J. 1991, ApJ, 379, 94
Mosquera, A. M., Kochanek, C. S., Chen, B., et al. 2013, ApJ, 769, 53
Muzzin, A., Marchesini, D., Stefanon, M., et al. 2013, ApJ, 777, 18
Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, ApJ, 490, 493
Oguri, M., & Marshall, P. J. 2010, MNRAS, 405, 2579
Oguri, M., Diego, J. M., Kaiser, N., Kelly, P. L., & Broadhurst, T. 2018, Phys. Rev. D, 97, 023518
Pejcha, O., & Prieto, J. L. 2015, ApJ, 799, 215
Pérez-González, P. G., Rieke, G. H., Villar, V., et al. 2008, ApJ, 675, 234
Perrotta, F., Baccigalupi, C., Bartelmann, M., De Zotti, G., & Granato, G. L. 2002, MNRAS, 329, 445
Phillips, M. M. 1993, ApJ, 413, L105
Porciani, C., & Madau, P. 2000, ApJ, 532, 679
Prada, F., Klypin, A. A., Cuesta, A. J., Betancort-Rijo, J. E., & Primack, J. 2012, MNRAS, 423, 3018
Quimby, R. M., Oguri, M., More, A., et al. 2014, Science, 344, 396
Rauch, K. P. 1991, ApJ, 374, 83
Reddy, N. A., & Steidel, C. C. 2009, ApJ, 692, 778
Rodney, S. A., Balestra, I., Bradac, M., et al. 2018, Nat. Astron., 2, 324
Ruiter, A. J., Belczynski, K., & Fryer, C. 2009, ApJ, 699, 2026
Rydberg, C.-E., Zackrisson, E., Lundqvist, P., & Scott, P. 2013, MNRAS, 429, 3658
Rydberg, C. E., Whalen, D. J., Maturi, M., et al. 2018, MNRAS, submitted [arXiv:1805.02662]
Sarmento, R., Scannapieco, E., & Cohen, S. 2018, ApJ, 854, 75
Schechter, P. L., & Wambsganss, J. 2002, ApJ, 580, 685
Schneider, P., & Weiss, A. 1986, A&A, 164, 237
Schneider, P., Ehlers, J., & Falco, E. 1999, Gravitational Lenses (Berlin: Springer-Verlag)
Schneider, F. R. N., Sana, H., Evans, C. J., et al. 2018, Science, 359, 69
Spera, M., Mapelli, M., & Bressan, A. 2015, MNRAS, 451, 4086
Sukhbold, T. 2019, ApJ, 874, 1
Susa, H., Hasegawa, K., & Tominaga, N. 2014, ApJ, 792, 32
Takahashi, R., Oguri, M., Sato, M., & Hamana, T. 2011, ApJ, 742, 15
Tan, J. C., Beltrán, M. T., Caselli, P., et al. 2014, Protostars and Planets VI, 149
Turner, E. L., Ostriker, J. P., & Gott, III, J. R. 1984, ApJ, 284, 1
Vega-Ferrero, J., Diego, J. M., Miranda, V., & Bernstein, G. M. 2018, ApJ, 853, L31
Venumadhav, T., Dai, L., & Miralda-Escudé, J. 2017, ApJ, 850, 49
Visbal, E., Haiman, Z., & Bryan, G. L. 2018, MNRAS, 475, 5246
Vitrichenko, E. A., Nadyozhin, D. K., & Razinkova, T. L. 2007, Astron. Lett., 33, 251
Walsh, D., Carswell, R. F., & Weymann, R. J. 1979, Nature, 279, 381
Wambsganss, J., Paczynski, B., & Schneider, P. 1990, ApJ, 358, L33
Wang, L., Shu, Y., Li, R., et al. 2017, MNRAS, 468, 3757
Watson, W. A., Iliev, I. T., D'Aloisio, A., et al. 2013, MNRAS, 433, 1230
Watson, W. A., Diego, J. M., Gottlöber, S., et al. 2014, MNRAS, 438, 412
Weymann, R. J., Latham, D., Angel, J. R. P., et al. 1980, Nature, 285, 641
Whalen, D. J. 2013, Acta Polytech., 53, 573
Whalen, D. J., Even, W., Frey, L. H., et al. 2013, ApJ, 777, 110
Windhorst, R. A., Timmes, F. X., Wyithe, J. S. B., et al. 2018, ApJS, 234, 41
Wise, J. H., & Abel, T. 2005, ApJ, 629, 615
Woosley, S. E., Heger, A., & Weaver, T. A. 2002, Rev. Mod. Phys., 74, 1015
Yee, J. C., Udalski, A., Sumi, T., et al. 2009, ApJ, 703, 2082
Zackrisson, E., González, J., Eriksson, S., et al. 2015, MNRAS, 449, 3057