Astronomy & Astrophysics

Imaging dark matter at the smallest scales with $z \approx 1$ lensed stars

Jose M. Diego[1],[*] [iD], Sung Kei Li[2] [iD], Alfred Amruth[2], Ashish K. Meena[3], Tom J. Broadhurst[4,5,6], Patrick L. Kelly[7,8],
Alexei V. Filippenko[9], Liliya L. R. Williams[7,8] [iD], Adi Zitrin[3] [iD], William E. Harris[10], Marta Reina-Campos[10,11] [iD],
Carlo Giocoli[12,13] [iD], Liang Dai[14] [iD], Mitchell F. Struble[15], Tommaso Treu[16], Yoshinobu Fudamoto[17] [iD],
Daniel Gilman[18],[**], Anton M. Koekemoer[19] [iD], Jeremy Lim[2], Jose María Palencia[1] [iD],
Fengwu Sun[20], and Rogier A. Windhorst[21]

[1] Instituto de Física de Cantabria (CSIC-UC), Avda. Los Castros s/n, 39005 Santander, Spain
[2] Department of Physics, The University of Hong Kong, Pokfulam Road, Hong Kong
[3] Physics Department, Ben-Gurion University of the Negev, PO Box 653, Be'er-Sheva 84105, Israel
[4] Department of Physics, University of the Basque Country UPV/EHU, 48080 Bilbao, Spain
[5] DIPC, Basque Country UPV/EHU, 48080 San Sebastian, Spain
[6] Ikerbasque, Basque Foundation for Science, 48011 Bilbao, Spain
[7] Minnesota Institute for Astrophysics, University of Minnesota, 116 Church Street SE, Minneapolis, MN 55455, USA
[8] School of Physics and Astronomy, University of Minnesota, 116 Church Street, Minneapolis, MN 55455, USA
[9] Department of Astronomy, University of California, Berkeley, CA 94720-3411, USA
[10] Department of Physics & Astronomy, McMaster University, 1280 Main Street West, Hamilton L8S 4M1, Canada
[11] Canadian Institute for Theoretical Astrophysics (CITA), University of Toronto, 60 St. George St., Toronto M5S 3H8, Canada
[12] INAF – Astrophysics and Space Science Observatory of Bologna, Via Piero Gobetti 93/3, 40129 Bologna, Italy
[13] INFN–Sezione di Bologna, Viale Berti Pichat 6/2, 40127 Bologna, Italy
[14] Department of Physics, University of California, 366 Physics North MC 7300, Berkeley, CA 94720, USA
[15] Department of Physics and Astronomy, University of Pennsylvania, 209 South 33rd Street, Philadelphia, PA 19104, USA
[16] Physics & Astronomy Department, University of California, Los Angeles, CA 90095, USA
[17] Center for Frontier Science, Chiba University, 1-33 Yayoi-cho, Inage-ku, Chiba 263-8522, Japan
[18] Department of Astronomy & Astrophysics, University of Chicago, Chicago, IL 60637, USA
[19] Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA
[20] Steward Observatory, University of Arizona, 933 N. Cherry Ave., Tucson, AZ 85721, USA
[21] School of Earth and Space Exploration, Arizona State University, Tempe, AZ 85287-1404, USA

ABSTRACT

Recent observations of caustic-crossing galaxies at redshift $0.7 \lesssim z \lesssim 1$ show a wealth of transient events. Most of them are believed to be microlensing events of highly magnified stars. Earlier work predicts such events should be common near the critical curves (CCs) of galaxy clusters ("near region"), but some are found relatively far away from these CCs ("far region"). We consider the possibility that substructure on milliarcsecond scales (few parsecs in the lens plane) is boosting the microlensing signal in the far region. We study the combined magnification from the macrolens, millilenses, and microlenses ("3M lensing"), when the macromodel magnification is relatively low (common in the far region). After considering realistic populations of millilenses and microlenses, we conclude that the enhanced microlensing rate around millilenses is not sufficient to explain the high fraction of observed events in the far region. Instead, we find that the shape of the luminosity function (LF) of the lensed stars combined with the amount of substructure in the lens plane determines the number of microlensing events found near and far from the CC. By measuring β (the exponent of the adopted power law LF, $dN/dL = \phi(L) \propto (1/L^{\beta})$), and the number density of microlensing events at each location, one can create a pseudoimage of the underlying distribution of mass on small scales. We identify two regimes: (i) positive-imaging regime where $\beta > 2$ and the number density of events is greater around substructures, and (ii) negative-imaging regime where $\beta < 2$ and the number density of microlensing events is reduced around substructures. This technique opens a new window to map the distribution of dark-matter substructure down to $\sim 10^3\,M_\odot$. We study the particular case of seven microlensing events found in the Flashlights program in the Dragon arc ($z = 0.725$). A population of supergiant stars having a steep LF with $\beta = 2.55^{+0.72}_{-0.56}$ fits the distribution of these events in the far and near regions. We also find that the new microlensing events from JWST observations in this arc imply a surface mass density substructure of $\Sigma_* = 54\,M_\odot\,\mathrm{pc}^{-2}$, consistent with the expected population of stars from the intracluster medium. We identify a small region of high density of microlensing events, and interpret it as evidence of a possible invisible substructure, for which we derive a mass of $\sim 1.3 \times 10^8\,M_\odot$ (within its Einstein radius) in the galaxy cluster.

Key words. gravitational lensing: strong – gravitational lensing: micro – supergiants – dark matter

1. Introduction

Galaxy clusters are the most powerful lenses in the universe. At the critical curves (CCs hereafter), and ignoring microlenses,

---

[*] Corresponding author; jdiego@ifca.unican.es
[**] Brinson Prize Fellow.

small sources can be magnified by very large factors, with the maximum magnification for a source of radius R, $\mu_{\max} = \mu_o/\sqrt{R}$, where μ_o is a constant related to the smoothness of the lensing potential. For galaxy clusters, μ_o can be of order 10 when R is expressed in arcseconds. At the caustics of these clusters, stars with sizes a few times $R_\odot$ (that is, $R \approx 10^{-11}$ arcseconds

at redshift $z \approx 1$) can reach theoretical extreme magnification factors exceeding 10^6 (Miralda-Escude 1991). In practice, the ubiquitous presence of microlenses from the intracluster medium (ICM) reduces the maximum magnification for these stars to $<10^5$ (Venumadhav et al. 2017; Diego et al. 2018). Despite this reduction in the maximum magnification due to microlenses, the flux from massive lensed stars at $z \approx 1$ that are at a fraction of a parsec from a cluster caustic can be boosted by ~7–10 mag and be detected with current telescopes reaching a depth of 28 mag (Kelly et al. 2018; Golubchik et al. 2023; Diego et al. 2024a).

The extreme magnification near the critical curves of clusters has allowed the discovery of distant stars that would otherwise remain undetected. The first such star, Icarus at $z = 1.49$, was discovered (Kelly et al. 2018) with the *Hubble* Space Telescope (HST), and was quickly followed by many others also observed with HST (Rodney et al. 2018; Chen et al. 2019; Kaurov et al. 2019; Diego et al. 2022; Welch et al. 2022; Kelly et al. 2022; Meena et al. 2023a). The farthest star discovered to date with HST through this technique is Earendel at a record breaking $z \approx 6$ (Welch et al. 2022). In total, HST has already discovered several dozen lensed-star candidates at $0.725 < z < 6$ Kelly et al. (2022), most of them believed to be blue supergiants (BSGs) and luminous blue variable stars (LBVs). HST has passed the torch to the new *James Webb* Telescope (JWST), which in a short time has already discovered over a dozen lensed-star candidates (Chen et al. 2022; Diego et al. 2023a; Meena et al. 2023b; Furtak et al. 2024; Diego et al. 2023b; Yan et al. 2023). Among these, several are believed to be red supergiants (RSGs), which are difficult to detect with HST (Diego et al. 2023a,b, 2024a; Yan et al. 2023). JWST will extend the search for distant stars to even higher redshifts and also to fainter stars. With a little luck, JWST will even directly observe the first generation of stars (Pop III) in caustic crossing high-redshift galaxies (Windhorst et al. 2018).

Some of these transients are believed to be due not to microlensing, but to intrinsic variability of LBVs that can increase their brightness by several magnitudes (Weis & Bomans 2020). They have up to 5 mag variations on decade-long timescales, and smaller amplitudes on shorter timescales of months to years that are typical of supergiants. LBVs are luminous enough that they can be observed even at modest magnification factors ($\mu \approx 20$) if they are at $z < 1$ (a star at $z = 1$ with $L = 10^6 L_\odot$ would have apparent magnitude ~28.7 at $\mu = 20$). Owing to their variable nature, they can be identified as transients in difference images between two epochs. At higher redshift, even these bright stars would become undetectable unless they are magnified by larger factors (a star at $z = 2$ with $L = 10^6 L_\odot$ would have apparent magnitude ~32.4 at $\mu = 20$ and ~28.2 if $\mu = 1000$). Although most of the lensed stars are found in regions near cluster CCs, a significant fraction of these stars have been observed farther from the CCs where the magnification from the cluster is relatively small ($\mu < 100$). Examples include the off-caustic event described by Meena et al. (2023a) or some of the events reported by Kelly et al. (2022) and Yan et al. (2023).

The outbursts of LBVs can be confused with genuine microlensing events, especially if the observations are separated by long periods that do not allow us to distinguish a microlensing event from an LBV outburst based on the light curve. A genuine microlensing event (in the optically thin regime) near the microcaustic (or maximum magnification) has a well-defined shape for the light curve since the luminosity changes as $1/\sqrt{(t-t_o)}$, where t is time and t_o is the time at which the background star touches the microcaustic (t_o is a free parameter). LBVs are very

rare compared with the more numerous but fainter supergiant stars, and since we can only identify them through their outbursts (or active phase), active LBVs are even rarer, so we expect to see only a few of them. The specific number depends on their abundance in the host galaxy, driven primarily by the recent star-formation history of that galaxy; hence, we expect to see them in very blue portions of lensed galaxies. Despite their scarcity, but because of their high luminosity and varying nature, LBVs are good candidates for transient events that take place in regions of low magnification.

In the lens plane, the magnification at a short distance, d, from the CC can be well approximated by $\mu \approx \Theta('')/d('')$ (Schneider et al. 1992). In this expression $\Theta('')$ is related to the inverse of the derivative of the lensing potential at that position. For a symmetric lens, $\Theta('') = $ constant, and for an isothermal profile, it is exactly the Einstein radius, but for real non-symmetric lenses with elliptically shaped CCs, Θ varies along the CC, with maximum values at the cusps of caustics. For massive clusters where lensed stars have been discovered, $\Theta('')$ takes values between ~50'' and ~100''. Then, for these clusters, and at distances $d \gtrsim 1''$, the magnification from the cluster typically drops below 100. In these regions of the lens plane with $\mu < 100$, the combined effect from the macrolens and the microlenses is often subcritical, $\mu \times \Sigma_* \lesssim \Sigma_{\rm crit}$, for typical values of the surface mass density of microlenses found near CCs, $\Sigma_* < 20 M_\odot$. Near the CC, even for small values of Σ_* there is always a region around the CC in which microlensing supercriticality is achieved, $\mu \times \Sigma_* > \Sigma_{\rm crit}$. In this region, the probability of microlensing events is expected to be maximum (Diego et al. 2018; Palencia et al. 2023). We refer to this portion of the lens plane as the "near region". In contrast, outside this region and away from the CC, d increases and μ decreases with $\mu \times \Sigma_* < \Sigma_{\rm crit}$. Here we are in the microlensing subcritical portion of the lens plane where microlensing events are more rare. We refer to this portion of the lens plane as the "far region" corresponding to the regions inside and outside the corrugated network of small critical curves around the galaxy cluster CC (see, for instance, Diego et al. 2018, for a description of this corrugated network).

It is in principle difficult to explain the apparently high number of events found in the far region. This begs the question of whether a significant fraction of these events are active LBVs which can be more easily observed in the far region, or the cluster lens model is inaccurate on small scales, lacking substructures in the far region that can boost the magnification, and hence become supercritical around the substructures. Perturbations in the mass distribution on scales comparable to small satellites in the cluster (millilenses) can create pockets of relatively high magnification on angular scales of several milliarcseconds at distances of few arcseconds from the CCs. These pockets of high magnification become islands of supercriticality where microlenses along the line of sight can now create more frequent microlensing events. The combined lensing effect of a galaxy cluster scale lens, with its swarm of small satellites and the myriad of microlenses from the matter associated with the ICM, has not been studied previously in detail. We refer to this effect as 3M lensing (macromodel lenses, millilenses, and microlenses), and it constitutes one of the foci of this work.

Recent observations made with JWST of some of these cluster lenses have revealed a wealth of unresolved structures in the ICM (Lee et al. 2022; Faisst et al. 2022; Harris & Reina-Campos 2023). Some of these objects are expected to be globular clusters (GCs) that are stripped away from their host galaxies by strong tidal forces from the clus-

ter. These are the same forces that strip stars away from the infalling galaxies and into the ICM. An extended population of GCs in the ICM has been found, for example, in the rich lensing cluster Abell 2744 at $z = 0.3$ with JWST imaging (Harris & Reina-Campos 2023). In addition to GCs, the inner regions of galactic cores in small galaxies can survive tidal forces and appear as GC-like objects. These ultracompact dwarf galaxies (UCDs) tend to be more massive than GCs and possibly harbor a supermassive black hole (SMBH) in their center. For simplicity we refer from now on to all these unresolved objects as GCs, keeping in mind that other types of objects may fall into this category.

The number and distribution of these GCs is in agreement with observations made at lower redshift and also with expectations from numerical N-body simulations. Thousands of GCs with masses in the range $10^5 M_\odot < M < 10^7 M_\odot$ are expected to be found within the critical curve of these clusters (Faisst et al. 2022; Lee et al. 2022; Diego et al. 2024b; Harris & Reina-Campos 2023). These GCs can act as millilenses, whose lensing effect is magnified by the macrolens (Gilman et al. 2017; Dai et al. 2018; Williams et al. 2024). In the vicinity of GCs, pockets of high magnification are created which, combined with the ubiquitous microlenses, can result in an increased rate of microlensing events around the millilenses, and near their small CCs around them, typically spanning a few milliarcseconds in the image plane. That is, for the smallest millilenses the increased rate of events would appear to originate from the same HST or JWST pixel (30 milliarcseconds for NIRCam short-wavelength detectors). Small dark matter (DM) structures can also act as millilenses, since these are predicted by many DM models (Kolb & Tkachev 1993; Graham et al. 2016; Visinelli et al. 2018; Arvanitaki et al. 2020; Gilman et al. 2021; Gorghetto et al. 2022). Microlenses overlapping with these small-scale DM structures make transient events more likely around them, serving as signposts of small-scale fluctuations in the distribution of DM. This is discussed in detail in this work.

The small CCs around these millilenses (or in general small DM structures) will inevitably overlap with the microlenses from the same ICM. The net lensing effect is a combination of the macrolens, the millilens, and the numerous microlenses. The effect of large macromodel magnifications plus microlenses has been studied in detail in earlier work (Venumadhav et al. 2017; Diego et al. 2018; Diego 2019; Palencia et al. 2023). The combined effect of large macromodel magnification plus millilenses was studied over two decades ago by (for example) Mao & Schneider (1998) and Metcalf & Madau (2001), and more recently by many others (e.g., Hezaveh et al. 2016; Gilman et al. 2017, 2018; Dai et al. 2018; Cyr-Racine et al. 2019; Gilman et al. 2019, 2020; Powell et al. 2023; Gilman et al. 2024; Williams et al. 2024; Tsang et al. 2024). The combination of the three effects has not been considered in detail so far, and to the best of our knowledge is presented here for the first time.

In this work we pursue four goals: (i) study the macro+milli+micro lensing ("3M lensing") effect over stars at cosmological distances and near cluster CCs in order to provide context for recent and future discoveries of lensed stars where 3M lensing is likely taking place, (ii) address the question of whether the millilensing effect from the numerous millilenses is sufficient to explain the transient events found at distances $d > 1''$ from cluster CCs, (iii) study the relation between the number of observed microlensing events, the amount of substructure on small scales in the lens plane, and the luminosity function of the background population of high-redshift stars,

and (iv) apply our results to recent observations, in particular to the case of the seven alleged microlensing events found by HST in the Dragon galaxy at $z = 0.725$ as part of the Flashlights program (Kelly et al. 2022). This arc was originally known as the Giant Arc or A370 Arc01 (Soucail et al. 1987, 1988; Lynds & Petrosian 1989; Grossman & Narayan 1989; Smail et al. 1993, 1996), and rebranded as the Dragon arc after new images were obtained following the HST Servicing Mission 4 update of the ACS in 2009[1].

The paper is organized as follows. Section 2 presents a series of definitions that are used throughout and gives examples of typical scales appearing in lensing that become useful in later portions of the paper. The simulations of the 3M lensing effect used in this work are presented in Section 3 . We focus in Section 4 on the probability of magnification in 3M lensing. Section 5 discusses the scaling of the effect with millilens mass and macromodel magnification. In Section 6 we describe how to compute the contribution, from a given mass function of GCs, to the area in the source plane (which can be interpreted as a probability) where microlensing effects are expected to be maximum. Section 7 estimates the probability of microlensing events in the far region around millilenses, while Section 8 estimates the probability of microlensing events anywhere in the far region, not just near millilenses. In Section 9 we discuss how to apply the previous results to map the distribution of DM on small scales, and apply our results to the particular case of the Flashlight microlensing events in the Dragon arc. The Dragon arc has been observed by the VLT/MUSE, providing resolved spectral information along the arc (Patrício et al. 2018). We discuss our results in Section 10 and conclude in Section 11. An Appendix contains details of the lens model for the particular example used to illustrate this work.

2. Definitions and useful numbers

We use several definitions throughout the paper, which for convenience we summarize here. Critical curves (CCs) are the regions in the image plane (also known as lens plane or plane of the sky) where magnification formally diverges. The image plane and observer (or source) plane are connected through the lens equation, $\beta = \theta - \alpha(\theta, M)$, where β are positions in the source plane, θ are positions in the image plane and $\alpha(\theta, M)$ is the deflection angle that depends on the distribution of mass of the lens. Through this equation, we can map the CCs into the corresponding curves in the source (or observer) frame, which are called caustics. A caustic region is the portion of the source plane which is bounded by the caustic curves. The near region is defined as the portion of the lens plane close to the cluster CC where the rate of microlensing events is maximized. This region is defined in terms of the cluster magnification and the surface mass density of microlenses. It is a band around the cluster CC where the cluster magnification is above the critical value, $\mu \gtrsim \mu_{crit} = \Sigma_{crit}/\Sigma_*$, where Σ_{crit} and Σ_* are the critical surface mass density for lensing and the surface mass density of microlenses, respectively (Diego et al. 2018). An example is shown in Figure 1, where the near region is contained within the two thin cyan curves. Similarly, the far region is the portion of the lens plane where the macromodel magnification is $\mu < \mu_{crit}$, and in the same figure it would be the region outside the band defined by the two cyan curves.

---

[1] https://www.newscientist.com/article/dn17765-upgraded-hubble-telescope-spies-cosmic-dragon/
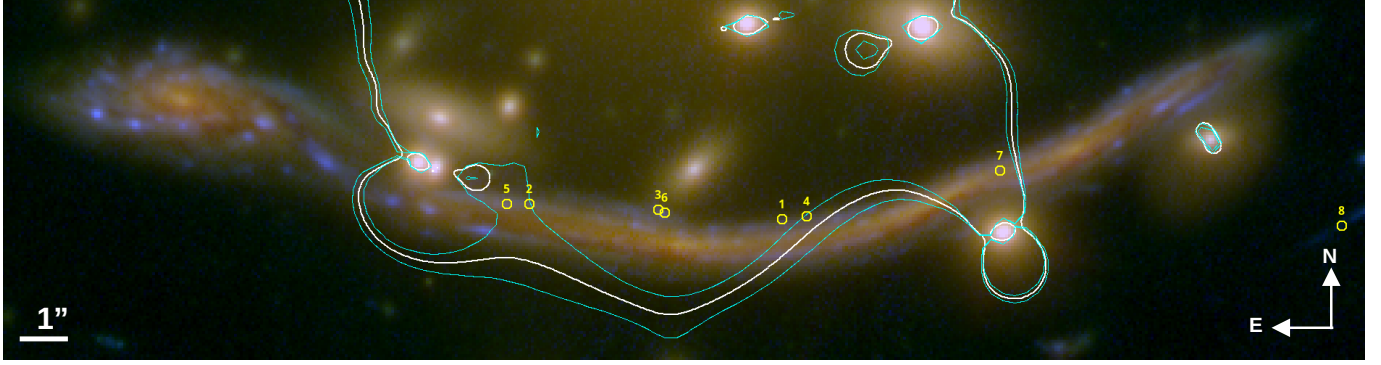
Fig. 1. Dragon arc as seen by HST (blue = $F435W$, green = $F814W$, red = $F160W$). The eight transients (seven in the arc) identified by Kelly et al. (2022) are marked with circles. Labels are the same as in the original reference. The white curve is the CC from our lens model (see Appendix) at the redshift of the arc. The two cyan curves mark the boundary region between macromodel magnification above and below 100. The arc covers ~1150 kpc^2 in the lens plane. Out of this, 190 kpc^2 is within the cyan curves (near region) and 960 kpc^2 is outside the cyan lines (far region).

Following standard practice (e.g., Treu 2010), the term macrolens is used when referring to the galaxy cluster scale lens, and the term millilens is used when referring to GCs or in general unresolved structures such as galactic core remnants, dwarf galaxies in the ICM, satellites in general, small DM halos, or intermediate-mass primordial black holes (Dike et al. 2023). These systems are expected to have Einstein radii of order milliarcseconds, hence the term millilensing. The term microlens is used for stars or stellar remnants in the ICM, which have Einstein radii of order microarcseconds. Some DM candidates such as primordial black holes with masses comparable to stellar objects would also fit in this category (see, for instance, Diego et al. 2018; Oguri et al. 2018; Vall Müller & Miralda-Escudé 2024).

For example, the Einstein radius of a 1 $M_\odot$ microlens at $z = 0.375$ and for a source at $z = 0.725$ (the redshifts of the cluster lens and Dragon galaxy, respectively) is 1.8 microarcseconds (μas) before accounting for the effect of the macrolens or millilens. For the same redshifts, a millilens with mass $10^5 M_\odot$ would have an Einstein radius of 0.57 milliarcseconds (mas), also before accounting for macromodel effects. For any other mass, M, at the same redshift, the Einstein radius would be $\theta_E \approx 1\,\text{mas} \times \sqrt{M/(3.1 \times 10^5 M_\odot)}$. In general, when embedded in a macromodel potential with magnification μ, the CC around the millilens or microlens with mass M behaves as a larger millilens or microlens with effective mass $\mu_t \times M$ (Diego et al. 2018; Oguri et al. 2018), where μ_t is the tangential macromodel magnification (μ_r would be the radial component and $\mu = \mu_t \mu_r$). For the particular case of a microlens near a millilens, the same scaling with magnification applies, only in this case the magnification μ_t is from the combined effect of the macromodel plus the millilens.

The CCs associated with these types of lenses are macro-CCs, milli-CCs, and micro-CCs. Similarly, we use the terms macrocaustic, millicaustic, and microcaustic when referring to the corresponding caustics. We refer to the macromodel magnification as μ_{1m}, while we use the term μ_{2m} when referring to the magnification from the combined macromodel plus millilens, and μ_{3m} (the 3M lensing magnification) when referring to the magnification of all three components (macrolens plus millilens plus microlenses).

In Section 9 we define the luminosity function (LF) of stars as $dN/dL = \phi(L) \propto (1/L)^\beta$, which gives the number of stars per luminosity bin and unit area. This "classic definition" is useful when working with nonmagnified and uniform distributions (or sources in the source plane before magnification is applied), since in this case the properties of the LF are independent of the region being considered. But when dealing with lensed sources, there is a strong dependence on the magnification. Because of this, we also use a different definition for the lensed luminosity function, or $\hat{\phi}(L)$, which gives the number of stars per luminosity bin and in a given area in the source plane (not per unit area). This alternative definition is useful when we are considering the number of stars in a particular region with macromodel magnification μ, in the interval $\mu_{min} < \mu < \mu_{max}$. That is, in this case $\hat{\phi}(L)$ means $\hat{\phi}(L, \mu_{min}^{max})$, where μ_{min}^{max} is all macromodel magnifications in the interval of magnification, but for convenience we simply use the expression $\hat{\phi}(L)$.

Below we present a few useful numbers for the particular case of the Dragon arc, which holds the record for the number of transients discovered as part of the Flashlights program Kelly et al. (2022). The location of these events in relation to the CC is shown in Figure 1. This arc contains seven high-significance transients, with at least two of them found in the far region (see Figure 1) and good candidates to be stars impacted by 3M lensing (Kelly et al. 2022). For the particular case of the Dragon galaxy, the redshift of the lens is 0.375 (A370 cluster), and the redshift of the lensed galaxy is 0.725. We adopt a flat-universe cosmology with $\Omega_m = 0.3$ and $h = 0.7$. For this model, the angular diameter distances to the lens at $z = 0.375$, the source at $z = 0.725$, and from the lens to the source are 1066 Mpc, 1495 Mpc, and 646 Mpc, respectively. We assume all matter in the lens is located in the lens plane at $z = 0.375$ and all sources are in the same plane at $z = 0.725$. In practice, the lens extends along the line of sight by $c \times \delta z \approx 1000\,\text{km s}^{-1}$. A smaller dispersion is expected in the source plane, but these corrections are at the percent level and are ignored here. For the same cosmology, 1″ subtends 5.16 kpc at $z = 0.375$ and 7.24 kpc at $z = 0.725$. The critical surface mass density for these redshifts is $\Sigma_c = 3640 M_\odot\,\text{pc}^{-2}$ and the distance modulus to $z = 0.725$ is 43.24 mag. For illustration purposes, a star with absolute magnitude -7 (color corrected in a given filter) and magnified by a factor of 100 would have apparent magnitude 31.2, still out of reach of JWST with 1 hr integration in one of the wide filters. However, the same star during a microlensing event (lasting typically a few days to a few weeks depending on the mass of the microlens, relative speed, and direction of motion with respect to the microcaustic), can be temporarily magnified by a factor of ~1000, and would appear ~2.5 mag brighter (i.e., ~28.7 mag) during that period. This would be detectable in a 1 h integration time with JWST and be interpreted as a transient.

Finally, we use the term "detectable through microlensing" (or DTM) stars to refer to all the stars that have detectable changes in brightness due to microlensing. These are either stars that (i) are detected in several epochs but between two epochs change their brightness by some amount (due to a microlensing event), or (ii) are detected in only one epoch because microlensing is temporarily boosting their flux. In general, we assume that the second type of DTM stars are ∼2 mag below the detection threshold before microlensing. Most of the stars at $z > 0.5$ that can be detected through microlensing belong to the supergiant type. During a microlensing event, these stars typically gain 1–3 mag (Kelly et al. 2018; Diego et al. 2018). We adopt a value of 2 mag as a compromise (see also Section 10.1 for a more detailed discussion). During a microlensing event, any DTM star increases (or decreases) its brightness by ∼2 mag and can be recognized as a transient.

3. 3M lensing

To study the 3M lensing effect, we rely on simulations that combine all three mass ranges (macro, milli, and micro). Since our focus is to study microlensing events around millilenses (embedded in a macrolens potential), we set the simulation parameters to match the scale of millilenses but at the same time resolve the microlenses. As mentioned in Section 2, before macromodel effects, the scale of a $10^5 M_\odot$ millilens is typically ∼1 mas in the image plane. On that scale, the effect of the macromodel can be very well approximated as a smooth gradient with a slope that is roughly the inverse of the Einstein radius of the macrolens (Diego et al. 2018). The scale of microcaustics is ∼1 μas. To properly resolve microcaustics, the pixel size needs to be much smaller than 1 μas. A pixel size of 10 nanoarcseconds (nas) in the source plane is sufficient to resolve the microcaustics from the smallest microlenses. After accounting for macromodel effects, the critical region of the millilens grows as the macromodel magnification. Hence, the simulation needs to span several milliarcseconds in the image plane if the macromodel magnification is $\mu_{1m} > 10$. To simulate several milliarcseconds in the image plane with a resolution of 10 nas in the source plane would require a prohibitive number of ∼10^{12} pixels. We can significantly reduce this by simulating a smaller millilens, since the number of pixels needed scales approximately with the mass of the millilens. Luckily, the magnification properties of 3M lensing for larger millilenses can be extrapolated by simply rescaling the results derived with smaller millilenses (see Section 5). In particular, we consider very small millilenses with masses of order $10^3 M_\odot$ and later study how our results scale with the millilens mass.

Hence, to explore 3M lensing in a wide range of scenarios, we define a fiducial model that is used for the main calculations and later study the scaling with macromodel, millilens, and microlenses around this fiducial model. For the fiducial model we adopt a macromodel magnification $\mu_m = \mu_t \times \mu_r = \pm 10 \times 2.3$, where μ_t and μ_r are respectively the tangential and radial magnifications from the macromodel. The magnification can be positive or negative depending on what side of the critical curve we are considering. The side with positive magnification is also the side with positive parity (counterimages have the same orientation as the original source). In contrast, when the magnification is negative the counterimage has negative parity (inverted in relation to the original source). The assumed macromodel magnification is small enough such that the microcaustics do not usually overlap and microlensing is a rare event. For the millilens we assume a fiducial model with a relatively small GC having a mass of $2 \times 10^3 M_\odot$, with a truncated core power-law density

profile

$$\rho(r)_{3D} \propto \frac{1}{(R_c + r)^\alpha} \, , \qquad (1)$$

where R_c is the core radius and the profile is truncated at some radius R_{max}. This millilens is representative of a small and compact GC that would survive the strong tidal forces in clusters. The core radii of the density profiles of millilenses can be approximately estimated from dwarf galaxies. Typical radii of $10^9 M_\odot$ galaxies in the LITTLE THINGS galaxy survey are about 300 pc, with substantial variation between individual galaxies (Table 2 of Oh et al. 2015). This size is consistent with those presented by Wolf et al. (2010). Motivated by the virial condition and assuming that the concentration parameter is the same for all subhalos, $r_{core} \propto m^{1/3}$. This relation can be scaled to smaller masses (Williams et al. 2024). For instance, for very small millilenses with mass $2 \times 10^3 M_\odot$, the core radii should be a factor of ∼80 smaller than for the $10^9 M_\odot$ halo, or ∼3.7 pc. For our calculations we adopt the most optimistic scenario where millilenses are most lensing-efficient, and therefore we assume a much smaller core radius of $r_c = 0.15$ pc. For the truncation radius we take ∼10 times the core radius. Very compact structures in the Milky Way, such as the central region of R136 in the Large Magellanic Cloud (LMC), would have a similar scale (diameter ≈1 pc from Massey & Hunter 1998). Our core and truncation radii for the small $2 \times 10^3 M_\odot$ GC are also consistent (after extrapolation to smaller masses) with the radii of the more massive GCs found in the Milky Way by Baumgardt & Hilker (2018), who find typical half-mass radii of ∼5 pc for GCs with mass ∼$10^5 M_\odot$. A small core radius also accounts for the fact that we expect the more compact structures to be the ones surviving in denser environments (Moliné et al. 2017).

The specific shape of the profile and truncation radius play a role in the lensing effect since they define the mass contained within the Einstein radius of the millilens. As mentioned above, in this work we consider the most favorable condition where the millilenses are very compact and most of their mass is contained within the Einstein radius. This is satisfied when $\alpha = 2$ or greater. In this situation, the dependence on the profile is very weak. Only for shallow profiles ($\alpha \lesssim 1.3$) and large cores, the millilens may be subcritical and not able to produce large magnification factors. A visual comparison of the millicaustics for four different millilens models is shown in Figure 2. The macromodel magnification for the four millilenses in the figure is set $\mu_{1m} = \mu_t \times \mu_r = -10 \times 2.3 = -23$. The two smallest millilenses have exactly the same mass and produce millicaustics that are nearly identical, despite the two millilenses having different core sizes and truncation radii (but the same α). The two millilenses with larger mass have a correspondingly larger millicaustic area. For the large millilens with $\alpha = 2$, the gap between the caustic regions (demagnification region) increases as the square root of the mass when compared to the smaller millilenses with the same α, so a millilens 100 times more massive can demagnify a region ten times larger in diameter. For the millilens with identical mass but a shallower profile ($\alpha = 1.5$), we observe a reduction in the lensing probability (or area with magnification greater than some value) of ∼25%. On the other hand, a steeper slope with $\alpha = 3$ (and consistent with N-body simulations of subhalos; Moliné et al. 2017) increases the lensing probability, but only by ∼2%, so our choice of $\alpha = 2$ is valid to represent even more compact millilenses with $\alpha > 2$.

Finally, to complete our fiducial model for the 3M lensing simulations, for the microlenses we consider a surface number density of $\Sigma_* = 50 M_\odot \, pc^{-2}$. This is close to the expected value
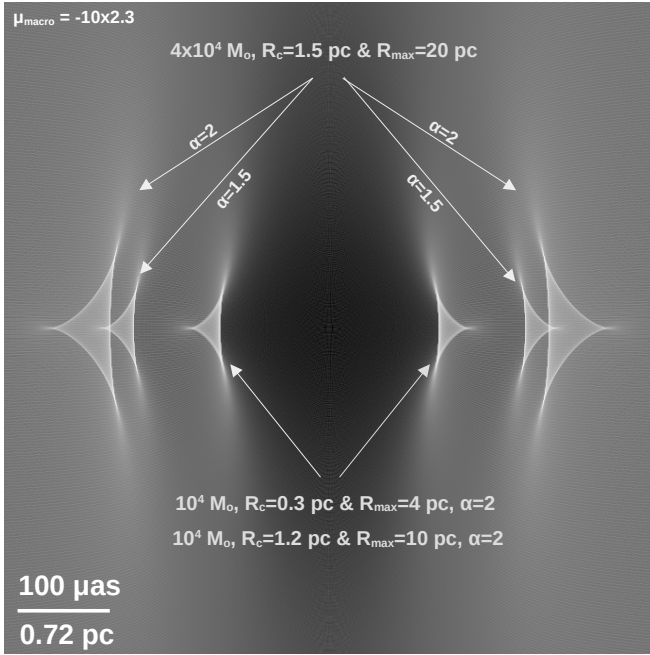
Fig. 2. Effect of the millilens profile. Comparison of millicaustics for four millilenses under the influence of the same macromodel magnification ($|\mu_{1m}| = 23$) but for different mass, core size R_c, truncation radius R_{max}, and exponent α. The profile is defined as $\rho(r) \propto (R_c + r)^{-\alpha}$. The image shown in grayscale is the sum of the four magnifications from the four millilenses. The caustics for the two millilenses with mass $10^4\,M_\odot$ and slope $\alpha = 2$ are nearly identical and fall on top of each other, indicating that the mass is the main driver defining the size of the caustic region. The largest millicaustic corresponds to a millilens with four times more mass, and larger core and truncation radii, but the same slope α. The area above $\mu = 100$ is a factor of four larger than in the smaller millilenses. A third millilens with the same mass, R_c, and R_{max} but a shallower profile ($\alpha = 1.5$) behaves as the larger millilens with $\alpha = 2$ but a mass of $2.93 \times 10^4\,M_\odot$, owing to the reduction in mass within the Einstein radius. Even shallower profiles ($\alpha \lesssim 1$) with large cores result in subcritical millilenses (no caustics or cusps). On the other hand, a steeper profile with $\alpha = 3$ or greater produces a millicaustic almost indistinguishable from the one obtained when $\alpha = 2$.

around the Dragon arc, if one assumes that stars in the ICM contribute ~2% to the total projected mass at this position. The value is also consistent with direct estimates of the surface mass density of stars in the intracluster light (ICL) from recent JWST data in massive clusters and at distances between 50 kpc and 70 kpc from the center of the cluster (Montes & Trujillo 2022), the distance at which our case study (the Dragon arc) is from the center of A370.

Originally, the pixel scale is set to 30 nas, and in the lens plane we distribute the microlenses randomly in a circular region of radius 1.2 mas until we reach the desired surface mass density of $\Sigma_* = 50\,M_\odot\,\mathrm{pc}^{-2}$. For the mass function of the microlenses we adopt a Chabrier (2003) model with a lower mass of 0.1 $M_\odot$. The specific model for the mass function plays a secondary role, since the most relevant parameter for microlensing is the value of Σ_*. The considered circular area is sufficiently large to easily accommodate the small millilens of the fiducial model. A second higher resolution simulation is later done around the cusps of the millicaustics and with a smaller pixel size of 10 nas that resolves the microcaustics even better.

Since the macromodel magnification can be positive (outside the cluster CC or positive-parity region; Blandford & Narayan

1986) or negative (interior to the cluster CC or negative-parity region), and the millilens caustics behave very differently depending on the parity, we simulate both parities but keep the absolute value of the macromodel magnification constant. When simulating the two parities, we only change the tangential component of the macromodel magnification – that is, we take the two values $\mu_t = \pm 10$. In tangential critical curves, the tangential magnification changes rapidly as one gets closer to (or farther away from) the CC, while the radial component of the macromodel magnification changes very slowly. A value of $\mu_t = \pm 10$ is representative of scenarios similar to the far region, where macro+microlensing alone is unlikely to produce transient events but the combined 3M lensing effect can boost the probability of transients around millilenses in the far region.

The magnification in the observer plane (caustics) is computed using standard ray tracing. We show the result for the fiducial model in Figure 3, and for the two parities, that is for $\mu_t = 10$ (positive parity) and $\mu_t = -10$ (negative parity). The radial magnification is identical in both cases, $\mu_r = 2.3$. The left panels show the caustic region with the 30 nas pixel size while the right panels display the higher resolution simulation with 10 nas per pixel and around two of the cusps of the millilens caustics. In all cases, the magnification (grayscale) is shown in log scale to better appreciate the details. The numbers in yellow indicate the typical magnification (from the macrolens and millilens) outside the caustic region and near the center of the caustic region.

At the caustics the magnification can be very large. For these simulations the maximum magnification is limited by the nonzero size of the pixel but still results in magnification factors of ~1000 at the caustics for the 30 nas pixel and a few thousand for the 10 nas pixel. A large star at $z = 0.725$ with $R \approx 100\,R_\odot$ would be ~33 times smaller than this pixel size, and the maximum magnification at the caustic would be ~6 times larger.

The case with positive parity (top-left panel) shows the classic diamond-shaped caustic. In the simulations, the larger tangential magnification from the macromodel goes in the horizontal direction, resulting in a caustic that is more stretched in the vertical direction. The magnification near the center of the caustic is almost twice the magnification of the macromodel, so most of the inner-caustic region provides a relatively modest boost in relation to the macromodel value. Only in the small regions near the four cusps of the caustic, and very close to the caustics themselves, the magnification from the millilens alone can be sufficiently large to make luminous stars at $z = 0.725$ detectable. Immediately outside the caustic region the most common value for the magnification is below the macromodel value. In this outer region the effect of the millilens is to slightly demagnify sources, hence compensating the larger magnification inside the millicaustic region, and ensuring that the average magnification over sufficiently large areas equals the macromodel value (flux conservation). A source which is significantly larger than the millilens caustic, for instance a star-forming region several parsecs in size, would have an average magnification very close to the macromodel value and thus insensitive to the presence of the millilens. Only very small objects within such a source, for instance stars, can attain large magnification values when they are near the millicusps or millicaustics.

For the case of negative parity (bottom-left panel) we observe some significant differences, with two small triangle-shaped high-magnification regions bracketing a larger low-magnification region. This is a well-known configuration for caustics in negative-parity regions (Chang & Refsdal 1979, 1984). The magnification between the two triangular-shaped caustic regions can be very small, of order 1. A small object with a size of 1 pc or less
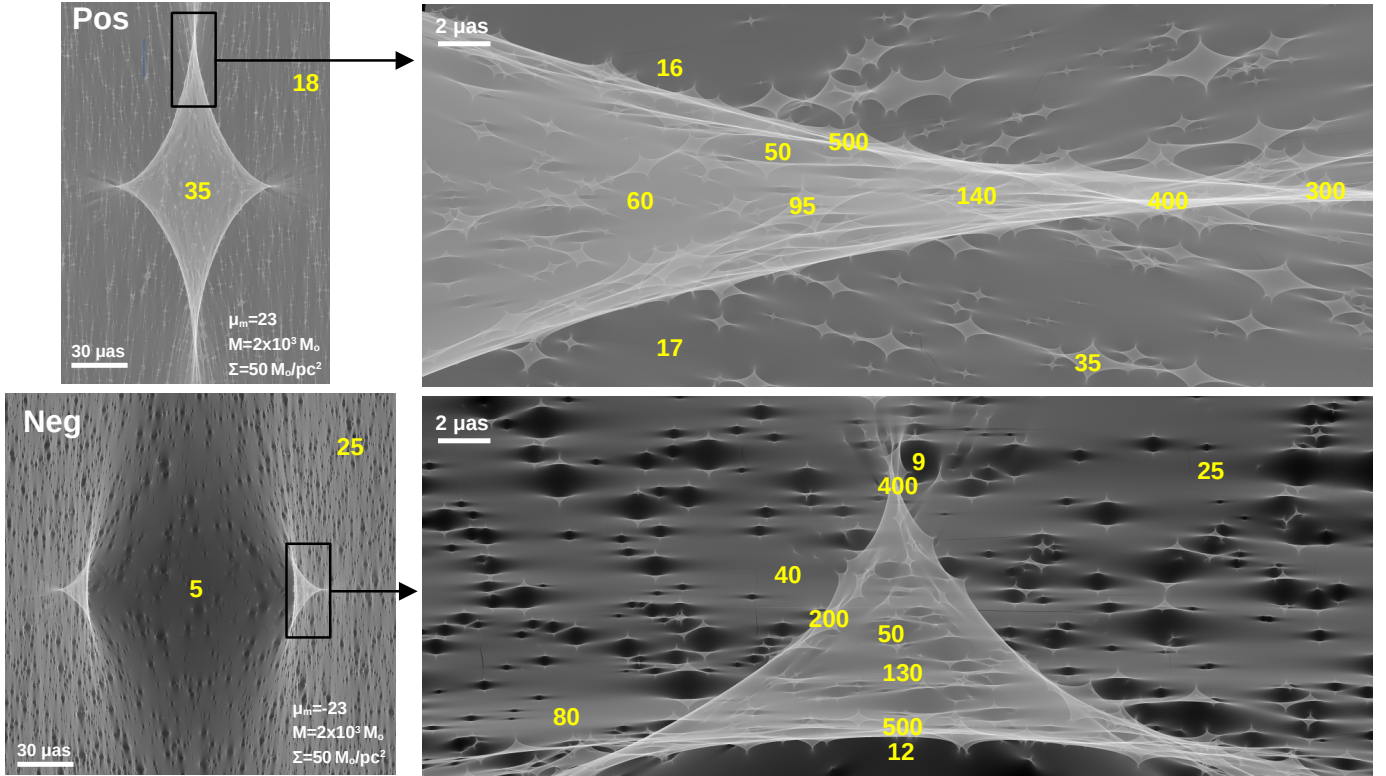
Fig. 3. Simulated magnification maps of 3M lensing. *Left panels.* Shown as grayscale is the log of the magnification in the observer plane (caustics) around a millilens with mass $2 \times 10^3 \, M_\odot$, in two regions (positive and negative parities) where the macromodel magnification is ± 23 and with a surface mass density of microlenses $\Sigma = 50 \, M_\odot \mathrm{pc}^{-2}$. The numbers in yellow indicate typical magnification values at these positions. For millilenses in regions with positive parity, outside the millicaustic region the magnification is typically below the macromodel value (but higher at the microcaustic regions). In the center of the millicaustic region, the typical magnification is ~50% higher than the macromodel value. This situation is reversed in the region with negative parity. *Right panels.* Zoom-in around the regions of highest magnification at the millicaustics and marked with black rectangles in the left panel. Near the millicusps and millicaustics, microcaustics always overlap one another at magnifications greater than 100, hence maximizing the occurrence of microlensing events.

placed in this inner region would be demagnified by the millilens, making its detection more difficult. This scale would be larger for heavier millilenses or larger macromodel magnification values. Hence, it is possible that sources a few pc in size such as GCs or small star-forming regions in the lensed galaxy get demagnified by a millilens and remain undetected if their lensed counterimage is in a negative-parity region behind a millilens. This cannot happen for counterimages behind the millilens in the portion of the lens plane with positive parity, where demagnification more than a few percent cannot take place. Since sources near a cluster caustic form two highly magnified counterimages near the CC, one counterimage with positive parity and one counterimage with negative parity, objects as small as a star or a small group of stars may appear highly magnified on one side of the CC (positive parity) and remain undetected on the other side of the CC (negative parity). This mechanism could explain the lack of asymmetry between the positive- and negative-parity images of stars or groups of stars recently observed in highly magnified galaxies (Diego et al. 2023b, 2022; Adamo et al. 2024). On smaller scales, a similar mechanism but involving microlenses was used to explain the lack of counterimages of lensed stars such as Icarus (Kelly et al. 2018).

At the smaller microlens level, we show in the right panels a zoomed-in version of the high magnification near a diamond-shaped cusp (positive parity) and a triangular-shaped caustic region (negative parity). In both cases we see how microcaustics adopt similar shapes (diamonds and triangles) and have a ten-

dency to align with the millicaustics. In some cases, microlenses around the millilens on the side with positive parity behave as microlenses with negative parity and vice versa for the side with negative parity. These rare exceptions can be appreciated near the cusp regions, where locally the parity can be inverted owing to the influence of the millilens. As expected, the number density of microcaustics increases in the cusps and near the caustics. This is due to the larger magnification of the millilens that concentrates more microcaustics in these regions. The size of the microcaustics also grows with the millilens magnification in a fashion similar to that near the caustics of galaxy clusters. The highest probability of observing microcaustic crossings is then near the cusps of millilens caustics. As described in earlier work (Diego et al. 2018; Palencia et al. 2023), when the effective surface mass density of microlenses approaches the critical value, Σ_{crit}, microlensing effects are maximized (in particular, fluctuations in the observed flux). For our fiducial model, this happens when the combined magnification from the macrolens and millilens is $\mu_{2m} \gtrsim \Sigma_{\mathrm{crit}}/\Sigma_* \approx 75$. The right panel of Figure 3 shows this effect near the cusps. For convenience, the baseline magnification (that is, μ_{2m}) is marked at different positions. In the case of millilens cusps in positive-parity regions (top-right panel), the magnification just outside the cusp is typically 25% smaller than the macromodel value. Inside the cusp region the magnification is higher than the macromodel and can exceed values of $\mu = 100$ near the cusps and caustics. There are also small areas around a few microcaustics where the parity is inverted and
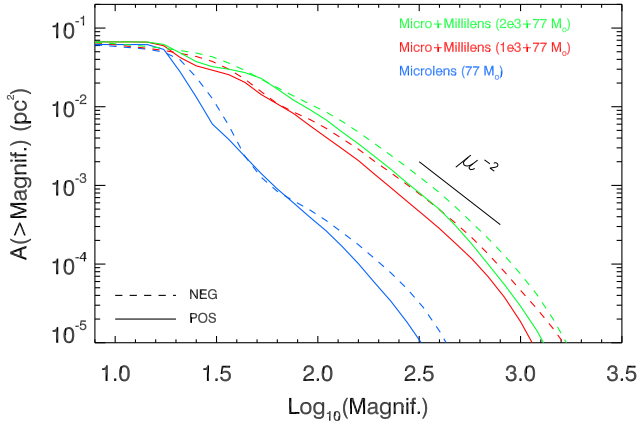
Fig. 4. Probability of magnification in 3M lensing. Blue lines are for macromodel plus microlenses only, while red and green lines are for 3M lensing and for two millilens masses. Dashed lines correspond to negative parity and solid lines to positive parity. The probability scales as the total small-scale mass (millilens plus microlenses).

the magnification can be relatively smaller. One such example is marked with the magnification value 50 in the top-right panel.

For millilenses in regions with negative macromodel parity (bottom-right panel), the most striking difference is the regions with significant demagnification. Outside the caustic region (or triangle), microcaustics can demagnify (with respect to the macromodel) regions as big as $R = 0.01$ pc, for instance the optical portion of a quasar accretion disk or a supernova photosphere months after the explosion[2]. In regions not containing a microcaustic, outside the caustic region the typical magnification ranges between $\mu \approx 25$ and $\mu \approx 40$ – that is, between ~10% and ~75% higher than the macromodel value – again compensating the lower magnification between the two caustic regions. Inside the caustic region the typical magnification is higher, especially near the cusps of the caustics. There is a sharp transition between the main caustic in the bottom portion of the figure where the magnification changes rapidly between extreme values to values of order 10. Inside the caustic region we also observe local changes in the parity, for instance around the microlens marked with magnification 200. As in the previous examples, near the cusps the microcaustics overlap, filling the space, and the probability of microlensing is maximum. In both examples we see this effect when $\mu_{2m} \gtrsim 100$, close to the value $\mu_{2m} \approx 75$ derived above. We adopt this value ($\mu_{2m} = 100$) as the critical magnification above which microcaustics are constantly overlapping and microlensing effects are maximum – that is, in what follows we assume $\mu_{crit} = 100$. In Section 10 we discuss how our results depend on this choice.

4. Statistics of 3M lensing magnification near a single millilens

The magnification pattern discussed in the previous section is interesting to interpret some events, and in particular to explain the lack of symmetry between pairs of images close to CCs when small objects, up to a few pc in size, are multiply imaged. In this work we are interested in the regime where the macromodel magnification is not that large, farther away from the CC, and in particular on the probability of having microlensing events

---

[2] This area would be even larger in regions with higher macromodel magnification.

around millilenses. For this it is useful to compute the area in the source plane having magnification above μ_{crit}, or $A(\mu > \mu_{crit})$, since stars in this area are the most likely to show microlensing effects.

We compute $A(\mu > \mu_{crit})$ in the two regions shown on the right side of Figure 3, for the two parities and around the cusps of the millilenses. We compare with the area computed in the same region and for the same configuration of microlenses but removing the millilens. The result is shown in Figure 4. Dashed lines refer to the area computed in portions of the lens plane with negative parity (bottom-right panel of Figure 3), while solid lines are for positive parity (top-right panel of Figure 3). The green lines are for a millilens with mass $2 \times 10^3 \, M_\odot$ plus microlenses, while the blue curves are for the case where only microlenses are included in the simulation. For comparison, we show as red lines the case where the mass of the millilens is reduced by a factor of 2. As in the case of microlensing near caustics explored in earlier work, the probability of high magnification is slightly larger in areas with negative parity (dashed lines). In these regions significant demagnification can take place in relatively large areas, that is compensated by the larger magnifications of the cusps.

In all cases, the probability of magnification scales as the expected μ^{-2} power law. The departure from this scaling at $\mu > 1000$ is mostly an artifact due to the nonzero pixel size, although at larger magnification factors of $\mu > 10000$ many microcaustics overlap and the magnification is expected to fall faster than μ^{-2} and become a log-normal distribution (Diego 2019; Palencia et al. 2023). The ratio of the green to the blue curves corresponds approximately to the ratio of masses between the millilens and the stellar mass in the same region. For this particular area the stellar mass in the right panels of the figure is roughly the fiducial value times the area of the two right panels and times the macromodel magnification (to transform the source area into image area): $M_* = (50 \, M_\odot \, \text{pc}^{-2}) \times (0.163 \, \text{pc}) \times (0.41 \, \text{pc}) \times 23 = 77 \, M_\odot$. Dividing the millilens mass ($2 \times 10^3 \, M_\odot$) by this mass gives a ratio of 27, which is roughly the ratio between the green and blue lines. Similarly, reducing the mass of the millilens by a factor of two results in a reduction in the probability by approximately the same factor (red curves).

Although not shown in the figure, the corresponding probability for the case where microlenses are ignored would be very similar to the fiducial model but a bit below the green lines owing to the small reduction in mass due to the absence of microlenses. Hence, if we are interested in the probability of having magnification $\mu_{3m} > 100$, this is basically determined by the millilens and the macromodel. In this situation, microlenses play the role of providing the temporary boost in flux to the lensed stars moving across the dense web of microcaustics to promote them beyond the detection limit and hence appear as transients. The problem can then be reduced to studying the contribution from a population of millilenses to the probability of having $\mu_{2m} > 100$ and across an area in the image plane where the macromodel takes different values of μ_{1m}.

5. Scaling with millilens mass and macromodel magnification

Having established that the most interesting 3M lensing effects concentrate around the cusps of the millilenses, and that we can reduce the problem we seek to solve to computing the probability that the macrolens plus millilens produce magnification greater than some value μ_{2m}, we now focus on the scaling of $A(> \mu_{2m})$ with the mass of the millilens (M_{mil}) and macromodel magnification (μ_{1m}).
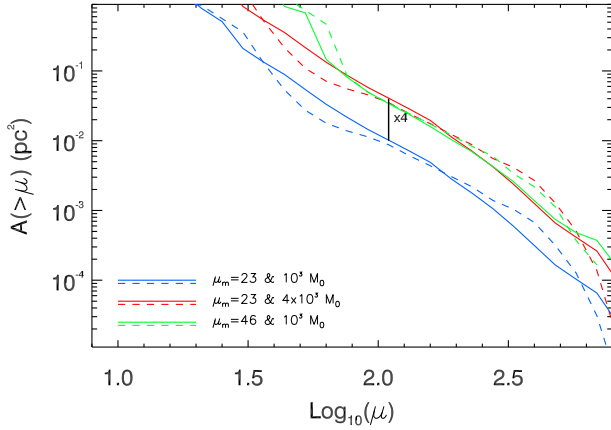
Fig. 5. Scaling of probability of magnification. The curves show the area in the source plane with magnification greater than a certain value due to millilenses in the lens plane. Solid lines are for millilenses in regions of the lens plane where the macromodel magnification is positive (positive parity), while dashed lines are for millilenses in regions of the lens plane with negative macromodel magnification (negative parity). Blue curves are for a millilens with mass $10^3 M_\odot$ and macromodel magnification ± 23, red curves are for millilenses with mass $4 \times 10^3 M_\odot$ and macromodel magnification ± 23. Green curves are for a millilens with mass $10^3 M_\odot$ and macromodel magnification ± 46. The black vertical line indicates a factor of four difference. The probability of magnification scales linearly with the mass of the millilens and quadratically with the macromodel magnification.

We characterize this probability by fitting the tail of the magnification with the canonical law $A(> \mu) = A_o/\mu^2$. The parameter A_o defines the strength of the millilens and contains the scaling we seek. Figure 5 shows an example with two masses for the millilens and two values for the macromodel magnification. As in Figure 4, dashed lines indicate negative parity and solid lines are for positive parity. The black vertical line shows a multiplicative factor of 4. This factor corresponds to the difference in mass and to the square of the difference in macromodel magnifications. Hence, the parameter A_o scales with mass as $A_o \propto M$ and with macromodel magnification as $A_o \propto \mu_{1m}^2$. A similar result is found in earlier work for microlenses (Diego et al. 2018; Palencia et al. 2023).

By fitting the different curves we find the scaling of the probability with the millilens mass (M_{mil}) and macromodel magnification (μ_{1m}),

$$A_{2m}(> \mu) = 0.19 \left(\frac{M_{mil}}{10^3 M_\odot} \right) \left(\frac{\mu_{1m}}{\mu} \right)^2 \text{ pc}^2 . \tag{2}$$

This scaling is almost insensitive to the particular values of the tangential and radial values of the macromodel magnification, and the probability depends only on their product, or μ_{1m}. In rare situations where $\mu_r \approx \mu_t$, the caustic shape morphs into a singular point, but the probability of magnification is still given by the same scaling and depends only on the product $\mu_t \times \mu_r = \mu_{1m}$. The law above is derived for the redshifts of the Dragon arc and cluster A370, but it can be rescaled for other redshifts simply by correcting for the factor $D_{ds}/D_d D_s$. Also, the scaling in Equation 2 appears to work for individual microlenses. We tested the scaling with a single simulation of a $2 M_\odot$ microlens in a potential with $\mu_{1m} = 20$ at a resolution of 2 nas per pixel, and the scaling in Equation 2 holds even at this low mass. One can even extrapolate this relation to cluster-scale lenses by considering $\mu_{1m} = 1$, since cluster lenses are generally in large-scale potentials with magnification $\mu_{1m} \approx 1$. The prediction for the area

above $\mu = 30$ for a cluster at $z = 0.375$ with mass $10^{15} M_\odot$, a source at $z = 0.725$, and $\mu_{1m} = 1$ is $A(\mu > 30) \approx 210 \text{ kpc}^2$, while for five well-modeled clusters in Vega-Ferrero et al. (2019) with masses $\sim 10^{15} M_\odot$ (excluding the supermassive MACS0717 cluster), the area $A(> \mu = 30)$ for these redshifts ranges between $\sim 1300 \text{ kpc}^2$ and $\sim 3700 \text{ kpc}^2$, corresponding to a factor of ~ 6 to ~ 20 more. Despite this disagreement, it is still remarkable that the prediction comes to within one order of magnitude, considering there is a 12 orders of magnitude difference in mass between a small $10^3 M_\odot$ millilens and a massive $10^{15} M_\odot$ galaxy cluster, and the latter are highly irregular, rich in substructure, and with shallower potentials (that are more efficient at increasing the area in the source plane with high magnification).

6. Probability of 3M lensing far from the cluster CCs from a population of millilenses.

Evolved GCs have masses in the range $\sim 10^3$–$10^6 M_\odot$ and are baryon dominated with mass-to-light ratios of a few (Goudfrooij & Fall 2016; Harris et al. 2017; Bragaglia et al. 2017; Baumgardt & Hilker 2018). Puffy or low-mass GCs are less resilient against disruption from tidal forces in the galaxy cluster, which together with two-body interactions can lead to their complete dissolution. Almost the entire range of GC luminosities has been measured in the Virgo and Fornax Cluster galaxies (Jordán et al. 2007; Villegas et al. 2010), where it is found that the luminosity functions (LFs) of evolved GCs are well matched by a log-normal distribution Harris et al. (2014). At higher redshifts, it is expected that the faint end of the LF will be boosted, since young low-mass clusters will not have been disrupted yet (Reina-Campos et al. 2022). Dynamical disruption mechanisms also affect massive clusters, thus lowering the maximum mass, but the presence of ultracompact dwarf galaxies (UCDs) in the observed samples would prevent detecting differences in this regime. Since colors and luminosities alone are not sufficient to disentangle these two populations, and both would produce the millilensing effect considered in this paper, we consider them both indistinctly. Recent work based on JWST has revealed a population of massive GC-like objects in galaxy cluster environments at intermediate redshifts, $z \approx 0.2$–0.4 (Faisst et al. 2022; Lee et al. 2022; Harris & Reina-Campos 2023). The high masses of some of these objects, exceeding in some cases $10^7 M_\odot$, are larger than those for massive GCs and are suspected to be the stripped galactic cores of dwarf galaxies (Faisst et al. 2022). The population of GC-like objects in galaxy clusters is then probably a combination of true GCs and UCDs.

To describe the mass function of GCs, we adopt a log-normal LF (Harris et al. 2014; Harris & Reina-Campos 2023). Assuming a constant mass-to-light ratio, the mass function should be similar to the LF (given as a function of magnitude in that reference). For the log-normal shape, we assume three parameters: (i) the peak, M_o, of the log-normal, which depends on the effect of dynamical disruption processes, as well as on the detection of the faintest and harder to detect GCs, (ii) the dispersion, σ, of the log-normal, and (iii) the number of GCs which we parameterize as a number density of GCs (the total area covered in the lens plane by the Dragon arc is $\sim 1150 \text{ kpc}^2$, out of which 960 kpc^2 are in the far region). The GC mass function takes the form (see Eq. (1) of Harris et al. 2014)

$$\frac{dN}{d \log_{10} M} = \frac{N}{\sqrt{2\pi}\sigma} \exp \left[-\frac{(\log_{10} M - \log_{10} M_o)^2}{2\sigma^2} \right], \tag{3}$$

where N is a normalization constant. We consider two alternative models that are shown in Figure 6; each one has a different
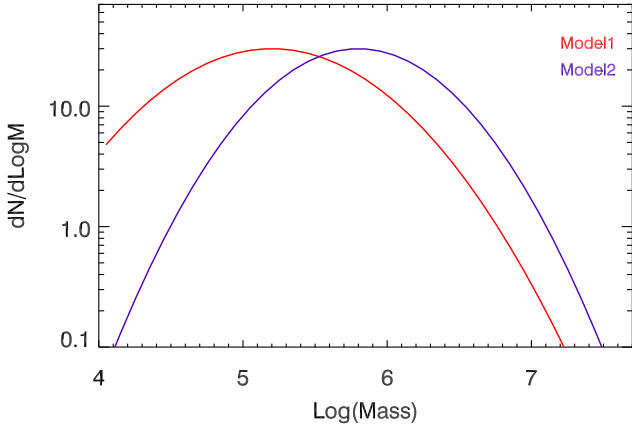
Fig. 6. GC mass function. The two colored lines show the two log-normal models used in this work (see text).

value of M_o and σ. Model 1 (with $\log_{10}(M_o) = 5.2$ and $\sigma = 0.6$) is our reference model and corresponds to the expected mass function of GCs from numerical simulations of star-cluster populations within cosmological zoom-in Milky-Way-mass simulations (Reina-Campos et al. 2022). In contrast, Model 2 (with $\log_{10}(M_o) = 5.8$ and $\sigma = 0.5$) is an alternative and top-heavy mass function that we use to check the dependency of our results with the GC mass function. The value of σ in these models is comparable to the universal value derived for the LF by Harris et al. (2014).

We can now combine all ingredients and compute the area in the source plane with magnification $\mu_{2m} > \mu_{crit}$ created by millilenses in the far region. Above μ_{crit}, microcaustics are constantly overlapping in the source plane and the probability of microlensing saturates at its maximum. As discussed earlier, we adopt $\mu_{crit} = 100$, which satisfies the supercritical condition $\Sigma_{eff} = \mu_{crit}\Sigma_* \gtrsim \Sigma_{crit}$ when $\Sigma_* \approx 50 \, M_\odot \, pc^{-2}$.

The area in the source plane where microlensing is most likely to take place is computed as the integral over the region in the lens plane with macromodel magnification $\mu_{1m} < \mu_{crit} = 100$ and the mass functions of GCs,

$$A_{far}(\mu_{2m} > 100) = \int d\mu_{1m} \int \frac{dN}{dM} P(\mu_{1m}) A_{2m}(> \mu) \, dM \,, \quad (4)$$

where $A_{2m}(> \mu)$ is given by Equation 2, the magnification is integrated between 1 and 100, and $P(\mu_{1m})$ is the probability for the macromodel magnification (or area with magnification μ) in the lens plane. This probability goes as μ^{-1} when taking logarithmic bins in μ. That is, we take $P(\mu_{1m}) = dA/d\log(\mu) = P_o/\mu_{1m}$ and determine P_o with the constraint $\int d\mu P(\mu_{1m}) = 1$.

7. Expected vs. observed number of transients near millilenses in the far region

With Equation 4, we can compute the area in the source plane around millilenses in the far region with magnification $\mu_{2m} > 100$, or $A_{far}(\mu_{2m} > 100)$, but we want to compare this area with the area in the near region satisfying $\mu_{1m} > 100$, or $A_{near}(\mu_{1m} > 100)$. Most microlensing events are expected to take place in these two areas. Microlensing can in principle take place with similar probability in both areas, provided the number density of stars is the same in both regions[3].

[3] In Section 9 we see how this also depends on the LF of the background stars.

The area in the near region is determined by the lens model for the galaxy cluster. We use the free-form WSLAP+ model derived for this cluster with the latest constraints from HST (see appendix). Based on the WSLAP+ model, we first compute the area in the source plane (from the macromodel) with magnification >100 and that overlaps with the Dragon arc. This area can be computed in the image plane, then divided by a factor of 100 to transform it into source-plane area, and finally divided by an additional factor of two to account for the two parities. This results in $0.95 \, kpc^2$ in the near region of the source plane where the macromodel should produce two counterimages with magnification $\mu > 100$ each. The two counterimages should appear in the corresponding near region in the image plane (band determined by the two cyan curves in Figure 1). Alternatively, the area above a certain magnification can be computed directly in the source plane with ray-tracing methods. In this case we obtain the total magnification of a source that gets multiply imaged into N counterimages. At large magnification factors, usually two of the counterimages carry most of the amplification (this happens when the source is very close to a cluster caustic). In this situation one can approximate the total magnification as twice the magnification from each counterimage. To account for this effect we then need to compute the area in the source plane with magnification >200, resulting in an estimate of $0.57 \, kpc^2$ in the source plane. Neither method is perfect when addressing global properties of an entire galaxy, especially in the case of the Dragon arc where multiple cluster caustics intersect the background galaxy but the range 0.57–$0.95 \, kpc^2$ should be a good approximation to the truth (within a factor of 2). This range for the area $A_{near}(\mu_{1m} > 100)$ is shown as an orange horizontal band in Figure 7. The luminous stars in this area are the most likely to experience microlensing near the cluster CC.

Before computing the result of Equation 4, we confirm that the macromodel probability of the WSLAP+ model does indeed scale as $P(\mu_{1m}) = dA/d\log(\mu) = P_o/\mu_{1m}$. This is demonstrated in Figure A.1 in the Appendix. The ordinate in Figure 7 shows Equation 4 computed in the far region and for the two GC mass-function models shown in Figure 6. That is, for each model we show the total area near millicaustics in the source plane with magnification greater than $\mu_{crit} = 100$, and as a function of the number density of GCs overlapping with the Dragon arc in the far region ($\mu_{1m} < 100$). Any star in the background galaxy that falls within this area in the source plane will have the same probability of experiencing a microlensing event (creating counterimages in the far region of the image plane) than stars with similar brightness in the near region of the source plane and with an estimated area of 0.57–$0.95 \, kpc^2$ (counterimages would form in the near region of the image plane).

The area in the near region (source plane) is shown as a horizontal orange band at the top of the figure. Clearly, the prediction (solid lines) is below the orange band for any reasonable number density of GCs (vertical blue band). Microlenses overlapping with the millilenses would increase this only by a small amount since the stellar mass from the ICL overlapping with the millilenses is much smaller (see Figure 4), so the contribution from microlenses overlapping with the millilens to the area, $A(\mu > 100)$, is very small.

The abscissa in Figure 7 shows the number density of GCs. The total number of GCs can be obtained after multiplying by the area contained in the far region of the Dragon arc in the lens plane ($960 \, kpc^2$). The average mass of a GC after integrating the GC mass function (normalized to $\int dN/dM = 1$ GC) is close to the peak of the log-normal, so the abscissa also can be transformed into surface mass density by simply multiplying by this
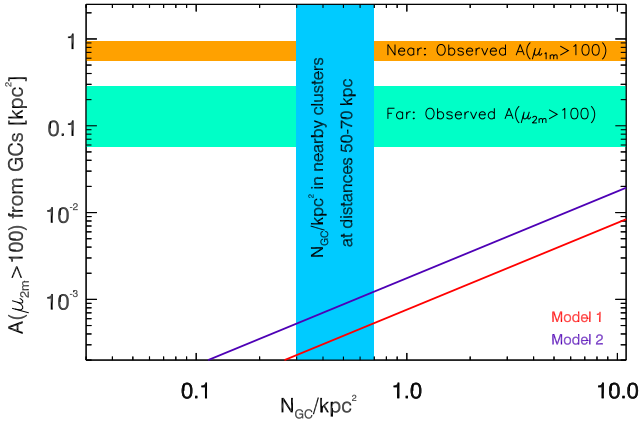
Fig. 7. Expected and observed area in the source plane with magnification $\mu > 100$. The red and blue solid lines show $A_{far}(\mu_{2m} > 100)$, the expected area in the source plane with magnification greater than 100 around millilenses in the far region for the two different mass functions shown in Figure 6. This area is computed as a function of the number density of millilenses, N_{GC}, and later rescaled to the area in the far region (960 kpc^2). The blue vertical band shows the typical range of number density of GCs at the distance of the Dragon arc from nearby clusters. The horizontal orange band shows $A_{near}(\mu_{1m} > 100)$, the area in the source plane with macromodel magnification $\mu_{1m} > 100$. The green horizontal band represents the fraction of microlensing events found in the far region with respect to the near region (~0.1 to 0.5 times the number of events found in the near region).

number. Since the Dragon arc is at a distance of ~50–70 kpc from the brightest cluster galaxy (BCG), we can compare this number density with the one observed in nearby clusters (such as Coma or Virgo). The blue vertical band in Figure 7 is the observed number density in the local universe from Peng et al. (2011) and for distances in the range 50–70 kpc.

Recent observations from the HST Flashlights program suggest that the observed rate in the far region is almost comparable to the number of events in the near region. The Dragon arc holds the record for the largest number of transient events reported so far in an individual galaxy. Kelly et al. (2022) find seven transients in this arc after comparing two deep epochs in very wide filters taken with HST. Six of these events have estimated macromodel magnifications below 100 (from two lens models), indicating a clear preference for these events to appear in regions of the lens model where the macromodel magnification is not extreme. The uncertainty in the magnification of these events is relatively high, especially near the CCs, but even adopting a more conservative value for the critical magnification of $\mu_{crit} = 30$, three of the events have magnifications below 30 in the two lens models considered by Kelly et al. (2022). From our lens model, at least two events are clearly in the far region (see Figure 1). As a conservative and generous range, we assume that the ratio of far-to-near events is between 0.1 and 0.5 times the lower bound of the orange band. This range is represented by the green band in Figure 7.

A similar result is found in the Warhol galaxy ($z = 0.94$) but with JWST observations (Yan et al. 2023). Seven transient events are found, with three in regions having macromodel magnification below 100 (and as low as $\mu \approx 30$). Interestingly, all three events peak their emission at wavelengths $\lambda_{peak} > 2\,\mu m$, suggesting these are cool stars. For the case of Warhol the rate of far-to-near events would then be close to 0.5, and given the very red nature of these transients, the LBV hypothesis seems less likely. The smaller number of events is partially due to the

fact that the galaxy is farther away, so it requires even more extreme magnification factors to detect the same star, disfavoring the LBV hypothesis for these events. Also, the cross section of the cluster caustics with the background galaxy is substantially smaller than for the Dragon arc, hence reducing the chance of finding stars near high-magnification regions. On the other hand, Warhol is at half the distance from the BCG than the Dragon arc, so the number density of GCs and the probability of microlensing near GCs in Warhol should be at least double the probability of the Dragon galaxy, but still far too small to explain the observed ratio of more than 0.1. In the same work (Yan et al. 2023), four additional events are reported in the galaxy Spock, at a slightly larger redshift, $z = 1.0054$. One out of the four events was found in a region with predicted macromodel magnification below 100, which would put the rate of far-to-near events at ~1/3, again orders of magnitude higher than expected. As in the case of Warhol, this transient is also very red, $\lambda_{peak} > 2\,\mu m$, making the LBV interpretation equally unlikely. For the Spock galaxy the high ratio of events far away from the CC is even more striking since this galaxy is in a portion of the cluster with an estimated surface number density of microlenses lower than for Warhol and the Dragon arc, so the amount of magnification needed (from the macromodel) to achieve the critical surface number density is higher. A detailed treatment for the Warhol and Spock galaxies is beyond the scope of this paper (see, however, Diego et al. 2024a, where two of the HST microlensing events in the Spock arc are studied in more detail). Here we simply use them as additional examples of an apparently high ratio of events in the far to near regions.

If the number density of stars that can be detected during a microlensing event is the same in the far and near region, the ratio of the areas in the near and far regions translates directly into the expected rate of microlensing events in the near and far regions. In Section 9 we subsequently see how the LF of the background stars plays also an important role in determining the final number of microlensing events, but here we can anticipate that for any reasonable LF, and at GC densities of 1 GC kpc^{-2}, the area from millilenses in the source plane (and hence the relative probability between the far to near events) is far below what is needed to produce a significant number of microlensing events in the far region (horizontal green band).

8. Transients from microlenses alone (no millilenses) in the far region

So far we have focused all our attention on the possible role played by millilenses at explaining the $0.7 < z < 1$ transient events observed in the far region of cluster CCs. Since in this region the macromodel magnification is relatively small, microcaustics from stars contributing to the ICM do not overlap in the source plane and the probability of microlensing is greatly diminished, but this does not mean microlensing in areas with lower magnification μ_{2m} cannot take place.

One fundamental difference between microlenses and millilenses is that microlenses have a much higher number density. At the distance from the BCG of the Dragon arc, the surface mass density of microlenses in our fiducial model is $\Sigma_* = 50\,M_\odot\,pc^{-2}$. This estimate is consistent with measurements based on the ICL at similar distances (Montes & Trujillo 2022). In the 960 kpc^2 occupied by the Dragon arc, this surface mass density translates into a total mass of $4.8 \times 10^{10}\,M_\odot$. This is a factor of 50 larger than the mass from GCs assuming a number density of 1 GC per kpc^2 and a mean mass of $10^6\,M_\odot$ per GC (Model 1).
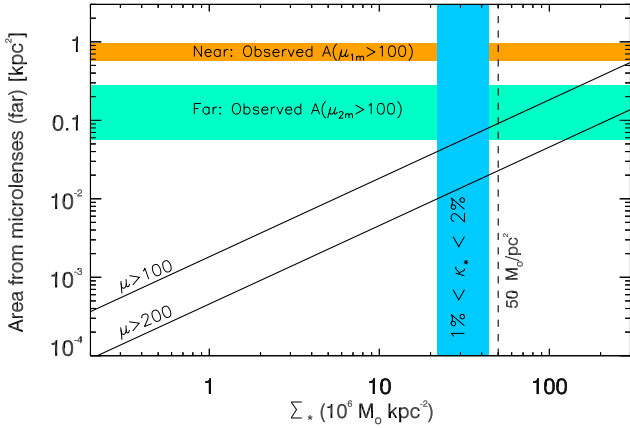
Fig. 8. Contribution from microlenses in the far region to high magnification. This result is similar to Figure 7, but considers only the effect of macromodel magnification and microlenses in the far region. The two solid black curves show the area in the far region of the source plane where microlenses create magnification factors >100 and 200, and as a function of the surface mass density of microlenses. Above $\mu > 100$ in the far region, stars are already very close to a microcaustic and can reach it in a few months. For $\mu > 200$ more stars can be detected, but they reach the microcaustic on a shorter timescale. The vertical blue band shows the expected range of surface mass densities for microlenses that constitute 1% and 2% of the total projected mass. (The convergence at the redshift of the Dragon arc from the macromodel in the Dragon arc region ranges between 0.58 and 0.62, so we adopt the mean value 0.6, while the shear ranges from 0.36 to 0.38.) The vertical dashed line is the fiducial microlensing model.

It is then natural to expect that microlenses alone should play a bigger role than millilenses. We repeat the calculation done for the GCs, but this time as a function of the surface mass density of microlenses and ignoring the contribution from millilenses. Since the GCs assumed earlier are very compact, with masses contained within their effective Einstein radius, they behave as point masses, so we can use the scaling law in Equation 2 by simply replacing the millilens mass by the surface mass density of microlenses. This extrapolation can be tested against the simulation result shown in Figure 4, where for 77 $M_\odot$ we find an area above $\mu = 100$ of $\sim 5 \times 10^{-4}$ pc^2, while for the same mass and the scaling in Equation 2 we expect 7.7×10^{-4} pc^2 (in both cases $\mu_{1m} = 23$).

The result for microlenses alone is shown in Figure 8, where we compare the area in the far region of the source plane with magnifications $\mu > 100$ and $\mu > 200$. As expected from their larger surface mass density, the contribution from microlenses is substantially more than from millilenses. In the figure, we mark with a vertical dashed line the surface mass density of our fiducial model. The microlenses in this model are sufficient to explain the elevated rate of events in the far region. The blue vertical band marks the range of surface mass densities corresponding to convergence from the stellar component between 1% and 2% of the total convergence of the cluster at the position of the Dragon arc. Our fiducial model corresponds to $\kappa_* = 2.3\%$, a reasonable value for distances between 50 and 70 kpc to the center of the cluster. In Figure 8 we do not include the effect of millilenses. They would contribute $\sim 10\%$ to the area above magnification $\mu > 100$. This is a modest increase in the probability of microlensing. However, this increase would not be uniformly distributed; it would concentrate around the position of the millilenses, hence introducing clustering in the distribution

of observed microlensing events that can be measured and used to reveal the presence of the millilenses.

The results presented so far have not taken into account the LF of the background stars, since we are simply looking at the ratio of areas (or relative probabilities) between the far and near regions where $\mu_{2m} > 100$ (or $\mu_{3m} > 100$ for the far region) and $\mu_{1m} > 100$ (for the near region). The probability of microlensing is proportional to these areas, but the number of stars that can be detected through microlensing (as mentioned earlier, we refer to this group of stars as DTM stars) depends strongly on the LF as we shall see in the next section, where we also discuss the key elements that makes imaging DM substructure with lensed stars possible.

9. Mapping dark matter substructures with microlensing events

We have seen how millilenses are not the most likely explanation for the high fraction of events (proportional to the area with high magnification) found in the far region, but microlenses (and LBVs) offer a more likely explanation. However, we have also seen how the rate of microlensing events is enhanced around millilenses (or in general perturbations in the small-scale distribution of mass). Figures 7 and 15 of Williams et al. (2024) show that the number of highly magnified images is proportional to the length of millilens CCs. This offers the interesting prospect of using distant stars as backlights and microlensing events as the markers of substructure that is influencing the number of detected microlensing events. We can then map the location of microlensing events, and use them to learn about the distribution of matter along the line of sight. This is analogous to using a photographic plate to trace the distribution of photons crossing an imperfect glass with nonuniform thickness, where in our case the photons are the distant stars being microlensed and the irregularities in the glass screen are the small perturbations in the lensing potential from DM substructures. The analogy with photographic plates will be made more evident later in this section.

So far we have ignored the role played by the LF of the population of lensed stars, but the distribution of microlensing events depends on the distribution of matter and the specific form of the LF. Many of the microlensed stars are ~ 1–2 (apparent) magnitudes below the detection threshold before microlensing. In the following sections, we assume all microlensing events provide a boost of ~ 2 mag (on average), so the DTM stars would be the ones that during a microlensing event can be detected. The specific amount of magnification provided by microlenses is irrelevant for our calculations. All that matters is that microlensing can promote fainter stars beyond the detection threshold and make them vary in flux between two epochs, so they can be recognized as transients in the difference of images taken with the same filter. Most of the DTM stars would be undetectable without microlensing, but some may be already detectable with just the boost provided by the macrolens (and the millilens, if one happens to be nearby) and before microlensing, but all of them would appear as transients during a microlensing event.

9.1. Number density of DTM stars

The number density of DTM stars, and in a region with magnification μ, is given by

$$\rho(\mu, \beta) = \frac{A_s(\mu)}{A_i(\mu)} \int_{L_1(\mu)}^{L_{max}} \phi(L) \, dL, \tag{5}$$

where $\phi(L) \propto (1/L)^\beta$ is the classic (per unit area) LF of the background population of stars, $L_1(\mu) = L_{\min}/\mu$ with $L_{\min}$ the minimum luminosity that could be detected at the redshift of the background galaxy at magnification $\mu = 1$ and is set by the depth of the observations, $L_{\min} = 10^{0.4(m_{thr}-dm(z))}$, with m_{thr} the limiting magnitude of the observations, $dm(z)$ the distance modulus to redshift z, and for simplicity we ignore color corrections. In reality, since we are interested in stars that can be detected through microlensing, $L_{\min}$ is smaller by a factor of $10^{2/2.5} = 6.31$, or 2 mag fainter, and during a microlensing event, a star with luminosity $L_{\min}/6.31$ will be magnified enough to be detected. Since these are DTM stars, they can be detected when experiencing a microlensing event, and eventually all of them would be detected if one could monitor the area for a sufficiently long time. $L_{\max}$ is the most luminous star in the area considered and depends on the assumed shape of the LF, or the existence of a limiting luminosity for the stars, such as the Humphreys-Davidson (HD) limit for RSGs (Humphreys 1978). For our purposes, we assume the observations are deep enough such that we can see microlensing from stars much less luminous (before magnification) than the most luminous star in the portion of the galaxy being magnified. In the expression above we assume that the probability of magnification when millilenses and microlenses are added is similar to the one given by the macromodel alone, so we can simply rely on the macromodel magnification. This is a very good approximation since millilenses and microlenses do not modify the probability of magnification significantly (when computed over areas much larger than the scale of the micro or millilenses), but rather they borrow magnification from surrounding regions and redistribute it around the millilenses and microlenses (see Figures 7 and 8 of Diego 2019, where the probabilities of magnification for the smooth model and the smooth model plus microlenses are very similar). For moderate magnification factors and typical depths, $L_1(\mu) > L_{\max}$ and the integral is zero in Equation 5, but for sufficiently large μ, $L_1(\mu) < L_{\max}$ and the number of DTM stars (and consequently microlensing detections) is greater than zero. The areas in the image and source plane are A_i and A_s (respectively), and they are related by $A_s = A_i/\mu$. Taking as unit area $A_s = 1$ and replacing $\phi(L)$ by $(1/L)^\beta$, we find

$$\rho(\mu, \beta) = \left.\frac{L^{1-\beta}}{\mu(1-\beta)}\right|_{L_1(\mu)}^{L_{\max}} \tag{6}$$

except for if $\beta = 1$, in which case the number density is

$$\rho(\mu, \beta) = \left.\frac{\log(L)}{\mu}\right|_{L_1(\mu)}^{L_{\max}}. \tag{7}$$

For a steep LF with $\beta = 3$ and sufficiently large values of μ, we find that $\rho(\mu) \propto \mu$, where we have assumed that $L_1(\mu) \ll L_{\max}$. The number density of DTM stars from such a population of background stars would then directly trace the magnification, and with it the distribution of mass on small scales. If microlenses are present in this area, they will make these DTM stars detectable, with a rate of microlensing events that increases with the abundance of microlenses as shown in Section 8. The tight relation between the LF and the number of microlensing events is discussed later in Section 9.3.

In general, for any $\beta > 1$ and for $L_{\min}/\mu \ll L_{\max}$, the number density of DTM stars scales with μ as

$$\rho(\mu, \beta) \propto \mu^{(\beta-2)}. \tag{8}$$

Based on this, we can define two distinct regimes, which we identify with traditional photographic plate imaging, where photons are crossing a glass with nonuniform thickness. For steep LF with $\beta > 2$, we are in the positive-imaging regime. Here, the number density of DTM stars (the photons that reach the photographic plate) is larger around substructures with larger magnification factors (or in our analogy, when photons are crossing portions of the glass that focus the light more into the photographic plate). For shallow LF with $\beta < 2$ we are in the negative-imaging regime, where the number density of DTM stars is reduced around substructures. In our photographic-plate analogy, this would correspond to the negative of the photographic plate, where the silver particles have absorbed more photons behind the small-scale structures. For the particular case of an LF with $\beta = 2$, we expect the number density of events to be uniform and independent of magnification. The photographic analogy would be the superposition of the positive image and the negative plate, leaving as a result a homogeneous image.

To better illustrate this specific case, we discuss a simple experiment. For $\beta = 2$, and considering two logarithmic bins in magnification, we can think of two regions, A and B, with the same LF but different magnifications; $10 < \mu_B < 100$ and $100 < \mu_A < 1000$. Region A has mean magnification a factor of ten larger than region B. In the image plane, region A is ten times smaller than region B (true for logarithmic bins in μ), while in the source plane, region A is 10^2 times smaller than region B. The number density of DTM stars in the source plane of region A is ten times larger than in region B ($\int_{L_1}^{\infty} L^{-2} dL \propto L_1^{-1}$, with L_1 ten times smaller in A than in B), but the area is 10^2 smaller so there are ten times fewer objects detected in the image plane in region A than in B (see Figure 9), but since the area in the image plane of B is ten times larger, the number density of DTM stars in the image plane is the same in A and B.

The case of a shallow LF, $\beta < 2$, is counterintuitive since we expect to see a smaller number density of DTM stars in regions of higher magnification. This shallow LF resembles the faint end of the LF of quasistellar objects (QSOs) at high redshift. Most lensed QSOs are found in regions with moderate magnification factors, in agreement with Equation 8 above. This is similar to the enhancement-dilution effect, or magnification bias, discussed in the context of distant lensed galaxies and QSOs (Canizares 1981; Narayan 1989; Borgeest et al. 1991; Narayan & Wallington 1993; Broadhurst et al. 1995; Umetsu et al. 2014).

9.2. The observed luminosity function.

The LF of the observed events, $\hat{\phi}(L)$, can be directly related to the LF of the background stars (before magnification). For an LF with $\beta \neq 3$ and considering a region with minimum magnification μ_1, we have

$$\hat{\phi}(L) = \int_{\mu_1}^{\mu_{\max}} \phi(L/\mu) A_s(\mu) \frac{d\mu}{\mu} \propto \mu_1^{\beta-3} \phi(L), \tag{9}$$

where $\mu_{\max}$ is the maximum magnification for a star, $\mu_{\max} \approx 10^4$ for supergiant stars (see Section 10.1), and we have assumed $\mu_{\max} \gg \mu_1$. The extra term $1/\mu$ inside the integral is the reduction in luminosity bin size at magnification μ, and $A_s(\mu) \propto \mu^{-3}$. For the particular case of $\beta = 3$ we have

$$\hat{\phi}(L) \propto (\log(\mu_{\max}) - \log(\mu_1)) \times \phi(L), \tag{10}$$

which has a weak dependence on μ_1. As discussed in Section 2, the luminosity functions in Equations 9 and 10 do not conform with the classic definition of number density per luminosity and unit area, but instead correspond to regions in the image plane

with an area that depends on μ_1 and $\mu_{\max}$. The LF maintains its form, but its amplitude (compared with the amplitude of $\phi(L)$), scales as $\mu_1^{\beta-3}$ for $\beta \neq 3$, and remains virtually independent of the magnification when $\beta \approx 3$. For a LF with $\beta \approx 3$, the reduction in area in the source plane ($A_s(\mu) \propto \mu^{-3}$) is almost perfectly compensated by the increase in the number of objects with smaller luminosity L/μ. In this case, the number of lensed objects per logarithmic interval in magnification is the same at all magnifications (a visual example of this constancy in μ is shown in Section 9.4).

Similarly, we can define the probability of magnification of the lensed stars, $\hat{\phi}(\mu)$ as

$$\hat{\phi}(\mu) = \int_{L_{\min}}^{\infty} \phi(L/\mu) A_s(\mu)\, dL \propto \frac{\mu^{\beta-3}}{L_{\min}^{\beta-1}}. \tag{11}$$

For an LF with $\beta = 3$, all magnifications have similar probability, so the observed population of lensed stars will have equal fractions of fainter and luminous stars. A shallower LF with $\beta < 3$, will be dominated by low-magnification events far from the CC (i.e., from intrinsically very luminous stars), while a steeper LF with $\beta > 3$ will produce mostly high-magnification events near the CC (i.e., low intrinsic luminosity stars).

9.3. Connecting the distribution of transients with the LF and the amount of substructure.

Interestingly, the observed number density of events in Flashlights traces the magnification (higher concentration of events in the near region), so this points (in principle) to a population of DTM with $\beta > 2$ (Equation 8). But this would be true only if substructure is not present. The relation between the LF and substructure adds complexity to this interpretation, as we have seen in Sections 7 and 8, where the probability of having transient events (proportional to the area where large magnification factors are possible) in the far region depends also on the amount of substructure (micro- and millilenses). To estimate the ratio of microlensing events in the far and near regions, one needs to take into account both the LF and the amount of substructure.

We begin by computing the number of DTM stars in a logarithmic bin in magnification. For the sake of clarity, we derive the scaling with μ both in the image and source planes, and show how they are both equivalent. To show this scaling with μ, we consider two areas in the image plane, A_i and B_i, with magnifications $\mu_A < \mu_B$. For simplicity, the widths in magnification of areas A_i and B_i are the same in logarithmic scale. In particular, we will be considering two bins in magnification (in log scale), $10 < \mu < 100$ for the far region and $100 < \mu < 1000$ for the near region, or respectively A_i and B_i. When considering logarithmic bins, the area of A_i is μ_B/μ_A times bigger than the area of B_i when computed in the image plane. In the source plane, the area of B_s is reduced in size by an extra factor μ_B/μ_A since $A_s = A_i/\mu_A$ and $B_s = B_i/\mu_B$. That is, when computing areas in the source plane the area of A_s is $(\mu_B/\mu_A)^2$ times bigger than the area of B as expected (Schneider et al. 1992).

Image-plane interpretation. We count DTM stars in the image plane in areas A_i and B_i with magnification μ_A and μ_B. The number of DTM stars are the ones that are found in the smaller areas $A_s = A_i/\mu_A$ and $B_s = B_i/\mu_B$ in the source plane, and above the luminosity $L_1(\mu) = L_{\min}/\mu$, where $\mu = \mu_A$ or $\mu = \mu_B$. Since $B_i = A_i\mu_A/\mu_B$, then the area of B_s in the source plane is $B_s = B_i/\mu_B \propto \mu_B^{-2}$, and the number of DTM stars scales with

μ as

$$\frac{dN_{DTM}}{d\log(\mu)} \propto \frac{1}{\mu^2} \int_{L_1(\mu)}^{L_{\max}} \phi(L) dL \propto \frac{\mu^{\beta-3}}{L_{\min}^{\beta-1}}. \tag{12}$$

Source-plane interpretation. We count stars that fall in the source plane in areas A_s and B_s with magnifications μ_A and μ_B. In this case, the calculation is simplified since we can work directly with the area in the source plane which scales as $\propto \mu^{-2}$. The number of DTM stars above luminosities $L_1(\mu) = L_{min}/\mu$ is then

$$\frac{dN_{DTM}}{d\log(\mu)} \propto \frac{1}{\mu^2} \int_{L_1(\mu)}^{L_{\max}} \phi(L) dL, \propto \frac{\mu^{\beta-3}}{L_{\min}^{\beta-1}} \tag{13}$$

and thus equivalent to Equation 12. In the above equations, where we have made the usual approach that $L_{\max} >> L_1(\mu) = L_{\min}/\mu$. For DTM stars, we have seen how $L_{\min}$ is approximately 6.31 times below the luminosity corresponding to the detection limit ($2.5\log_{10}(6.31) \approx 2$ mag). We have also ignored the multiplicity of counterimages, but this cancels out when considering the ratio of events in the near and far region, assuming the multiplicity is the same in both regions. From Equation 13, and for an LF with $\beta = 3$, we expect the same number of DTM stars per logarithmic bin in magnification (see also Equation 11). In the image plane, the area per logarithmic bin in magnification scales as μ^{-1}, so the number density of DTM stars for this case would go as μ and trace the magnification, in agreement with Equation 8.

The total number of stars that experience microlensing in the far region is given by the number of DTM stars in that region (Eq. 13) times the probability of each star to experience microlensing. This probability is proportional to the black solid line in Figure 8:

$$N_{\text{far}} \propto \frac{dN_{\text{DTM}}}{d\log(\mu)} \times A(\mu_{2m} > 100, \Sigma_*). \tag{14}$$

For the near region, we have a similar expression, but replacing $A(\mu_{2m} > 100, \Sigma_*)$ by $A(\mu_{1m} > 100)$, which is given by the orange band in Figure 8. Here we ignore millilenses since we have established in Section 8 that the dominant effect is coming from microlenses.

We can now express the ratio of events in the near and far regions:

$$\frac{N_{\text{near}}}{N_{\text{far}}} = \left(\frac{\hat{\mu}_{\text{near}}}{\hat{\mu}_{\text{far}}}\right)^{\beta-3} \frac{A(\mu_{1m} > 100)}{A(\mu_{2m} > 100, \Sigma_*)}. \tag{15}$$

In the expression above, we made the simplification that in the far and near regions the number of events can be expressed in terms of their corresponding average magnification, $\hat{\mu}$, computed as the mean magnification in the source-plane region with $\mu > \mu_1$,

$$\hat{\mu} = \frac{\int_{\mu_1}^{\infty} \mu A_s d\mu}{\int_{\mu_1}^{\infty} A_s d\mu} = 2\mu_1. \tag{16}$$

Here we adopt $\mu_1 = 10$ for the far region of the Dragon arc and $\mu_1 = \mu_{\text{crit}} = 100$ for the near region.

The value of β can be obtained by inverting the equation above,

$$\beta = 3 + \frac{\log(N_{\text{near}}/N_{\text{far}}) - \log(R_A(\Sigma_*))}{\log(\hat{\mu}_{\text{near}}/\hat{\mu}_{\text{far}})}, \tag{17}$$

where R_A is the ratio of areas,

$$R_A(\Sigma_*) = \frac{A(\mu_{1m} > 100)}{A(\mu_{2m} > 100, \Sigma_*)} = A(\mu_{1m} > 100)\frac{500}{\Sigma_*(M_\odot\,\mathrm{pc}^{-2})}\,, \quad (18)$$

where from Figure 8) we have $A(\mu_{2m} > 100, \Sigma_*) = \Sigma_*(M_\odot\,\mathrm{pc}^{-2})/500$. For the near region, we take $A(\mu_{1m} > 100) \approx 0.7\,\mathrm{kpc}^2$, which is approximately in the middle of the orange band in Figure 8. That is, $R_A = 7$ for $\Sigma_* = 50\,M_\odot\,\mathrm{pc}^{-2}$. Only the ratio of areas in the far and near regions is relevant for this calculation, and this ratio of areas is independent of the value of μ_{crit} (see Section 10.2 below). The number of microlensing events is $N_{\mathrm{far}} = 2$, (labelled 3 and 6 in Figure 1) and $N_{\mathrm{near}} = 5$, while the ratio $\hat{\mu}_{\mathrm{near}}/\hat{\mu}_{\mathrm{far}} = 10$. Replacing these numbers in Equation 17, we finally obtain $\beta = 2.55^{+0.69}_{-0.48}\,{}^{+0.18}_{-0.29} = 2.55^{+0.72}_{-0.56}$, where the first error comes from Poissonian uncertainty in N_{near} and N_{far}, and the second error corresponds to the range $25\,M_\odot\,\mathrm{pc}^{-2} < \Sigma_* < 75\,M_\odot\,\mathrm{pc}^{-2}$. The final error bar is obtained after adding in quadrature the first two uncertainties.

With Equation 17, β can be quickly calculated for any arbitrary amount of substructure. Even though we have expressed Equation 17 as a function of the stellar surface mass density, Σ_*, in truth this value represents all substructure that contributes to the area in the source plane with magnification $\mu > 100$. For a larger value of $\Sigma_* = 140\,M_\odot\,\mathrm{pc}^{-2}$ we find $\beta = 3$ and for $\Sigma_* = 14\,M_\odot\,\mathrm{pc}^{-2}$ we find $\beta = 2$, or $\beta = 1$ for an unrealistically low $\Sigma_* = 1.4\,M_\odot\,\mathrm{pc}^{-2}$. Since β can be measured directly through the observed LF (when sufficient events are available), one can invert Equation 17 and derive Σ_*, or in general the surface mass density of substructure since all small substructure contributes to Σ_*.

Equation 17 and Figure 10 summarize the intricate relationship between the number of observed microlensing events, the LF, and the amount of substructure. The same ratio of events in the far and near regions can be obtained by (i) reducing β (hence increasing the relative number of DTM stars in the far region) and reducing Σ_* (compensating the increase due to lowering β by reducing the amount of substructure producing microlenses and millilenses), or (ii) increasing β, which increases the number of DTM stars in the near region in relation to the number of available DTM stars in the far region, but increasing Σ_* as well, thus compensating the reduction of DTM stars in the far region by increasing the chance of a microlensing event.

A visual version of Equation 17 is shown in Figure 10, where we invert the equation to show the ratio of events in the near and far regions as a function of β and the amount of substructure Σ_*. In the figure we highlight with white lines two possible combinations of the parameters Σ_* and β that produce equal ratios of events in the far and near regions, or five times more events in the near region than in the far region. The measured ratio using Flashlights data (2.5) falls in between these two lines. Future observations of this fascinating galaxy will improve the constraints of the observed ratio of events near-to-far and the exponent β. This can later be used to derive the amount of substructure, Σ_*, needed to make the observed ratio compatible with the observed β.

Recently, Fudamoto et al. (2024) report 46 alleged microlensing events in the Dragon arc based on the difference between two observations made with JWST and separated by $\sim$8.5 months. 45 out of these 46 events are in the Dragon arc. Together with the seven events from the Flashlights program in the Dragon arc, there are 52 known events so far in this arc. As shown by Broadhurst et al. (2024, who use the same lens model as in this paper, and hence the same definition of the far and near regions), 17 of these events are in the far region, and hence the
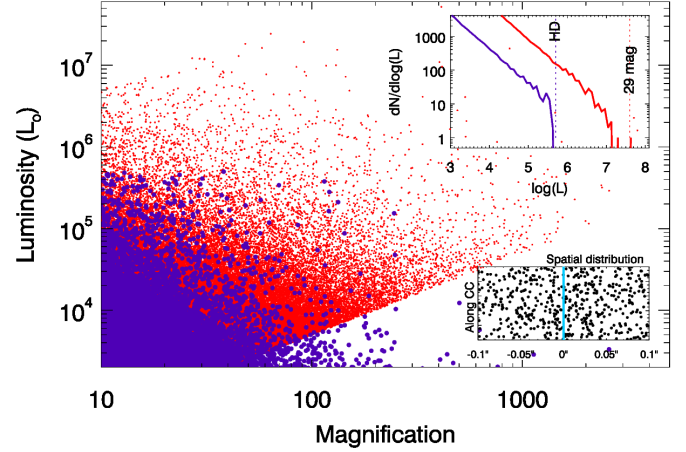

Fig. 9. Monte Carlo realization of lensed stars for a case with $\beta = 2$. For this example, half a million stars with luminosities in the range $50\,L_\odot < L < 5 \times 10^5\,L_\odot$ are placed in a region with magnification $\mu_{3m} > 10$. The maximum luminosity corresponds to the observed HD limit for supergiant stars in the local universe Humphreys (1978). The diagonal cut in the red points is just $L_{\mathrm{obs}} = 50 \times \mu\,L_\odot$. The simulation is complete above $L_{\mathrm{obs}} \approx 10^6\,L_\odot$. The LF of stars has a slope of $\beta = 2$. Blue dots represent stars before magnification and red dots are the magnified stars. In the top-right inset we show the LF of the stars before magnification (blue curve) and the observed LF after magnification (red curve). In the bottom-right inset we show the spatial distribution of the 500 brightest events and for a model with magnification $\mu = 1''/d$ where d is the distance to the CC (marked with a vertical light blue line at $d = 0''$) in arcseconds.

ratio of events $N_{\mathrm{near}}/N_{\mathrm{far}} = 2.05$. Since the lensed LF maintains the slope β (see Eq. 9), one can derive β from the observed LF of the microlensed events. From the new events of Fudamoto et al. (2024), Broadhurst et al. (2024) derive the observed slope of the LF and finds $\beta \approx 2.5$. From Equation 17, we can substitute the ratio $N_{\mathrm{near}}/N_{\mathrm{far}} \approx 2.05$ and the observed $\beta \approx 2.5$ to derive a value for the amount of substructure of $\Sigma_* = 54\,M_\odot\,\mathrm{pc}^{-2}$, very close to our fiducial value for microlenses, $\Sigma_* \approx 50\,M_\odot\,\mathrm{pc}^{-2}$. Thus, we can conclude that microlensing from intracluster stars with $\Sigma_* = 54\,M_\odot\,\mathrm{pc}^{-2}$ offers an excellent explanation for the distribution of the large number of events in this arc reported by Fudamoto et al. (2024). As discussed earlier, millilenses are expected to contribute $\sim$10% to the number of events, so the higher value for Σ_* derived above is consistent with the small contribution coming from millilenses. Future observations of this arc with JWST will reveal even more events that will better show not only the difference between the far and near regions, but also the clustering of microlensing events around the few millilenses that are expected to overlap with this arc. Some of this clustering is already evident in the spatial distribution of events shown by Fudamoto et al. (2024), but a detailed analysis is needed to distinguish between the clustering due to millilenses in the lens plane and intrinsic clustering of luminous stars in the background galaxy.

9.4. Validation with Monte Carlo simulations

In order to test the validity of Equations 8, 9, 11, and 13, we perform Monte Carlo simulations (see Kelly et al. 2018, where they also used simulated data to study the particular case of Icarus). One example is shown in Figure 9 for the particular case of $\beta = 2$ and computed in a region where the magnification is > 10 (i.e., $\mu_1 = 10$). We create a sample of half a million stars from the
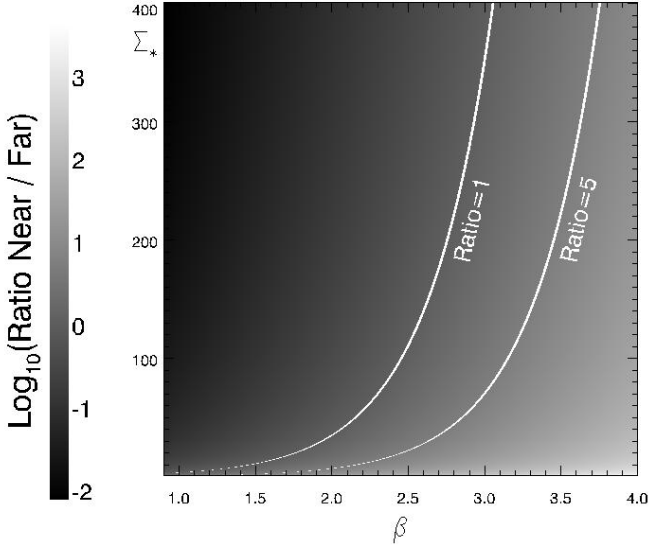
Fig. 10. Ratio N_{near}/N_{far} as a function of β and Σ_* (expressed in units of $M_\odot$ pc^{-2}). The white lines show the combination of β and Σ_* that predict the same number of events in the far and near regions (Ratio = 1) or five time more events in the near region than in the far region (Ratio = 5).

Fig. 11. Similar to Figure 9 but for the case where $\beta = 3$. In this case, all magnifications have comparable probability as predicted by Equation 11. The lensed LF (thick red line) lies a factor of $\hat{\mu}^{\beta-1}$ above the nonlensed LF (blue line). The smaller number of events at high luminosity when compared to Figure 9 is due to the fact that the number of stars in the simulation is the same in both cases, but for the steeper LF there are more stars with lower luminosities. The number density if events concentrates around the CC, tracing the magnification.

LF with luminosities $L > 50\,L_\odot$. This sample is shown as thick blue dots, with the bulk of the blue points falling below $10^3\,L_\odot$ and hence not shown in the plot. For this simulation we have set an upper limit to the intrinsic luminosity equal to the HD limit of $5 \times 10^5\,L_\odot$, so no blue dots are found above this value. The LF from this sample is shown as a blue line in the top-right panel. To each star we assign a random magnification in the interval $\mu_1 < \mu < \mu_{max}$ and following the canonical probability $A_s(\mu) \propto \mu^{-3}$. After multiplying the luminosity by the magnification, the magnified stars are shown as red dots, with the magnification for each star indicated on the abscissa. We compute the lensed LF from the same sample of half a million stars (red curve in the top-right panel). In a realistic situation, the number of stars in an area with magnification $\mu_1 < \mu < \mu_{max}$ should be a factor $\hat{\mu}^2$ smaller, but we use the same sample for convenience. In these conditions, the amplitude of the red line scales as $\hat{\mu}^{\beta-1}$, instead of the expected $\hat{\mu}^{\beta-3}$ from Equation 9, so for this particular case of $\beta = 1$, the red line is above the blue one by a factor $\hat{\mu} = 20$. Only one star is above a detection threshold of 29 mag, but a few dozen have apparent luminosities (after magnification) above several million and are in regions where the magnification already exceeds the critical value, $\mu_{crit} = 100$, so they are good candidates to move toward a microcaustic and be promoted beyond the detection threshold.

The diagonal cut in the red points corresponds to the smallest luminosity considered in the simulation, $50\,L_\odot$. In the simulation, there is no star brighter than the HD limit of $L_{max} \approx 5 \times 10^5\,L_\odot$, while in the lensed sample, we can reach apparent luminosities exceeding $10^7\,L_\odot$. The most luminous stars in the lensed sample correspond in this case to relatively moderate magnification factors, $\mu \approx 400$ (before microlensing).

Regarding the number density, Equation 8, the uniform number density of observed events when $\beta = 2$, is also well reproduced by the Monte Carlo, as shown in the bottom-right inset plot of Figure 9. The abscissa shows the inverse of the magnification of the observed events, which can be transformed into a distance to the CC for an spherical isothermal lens model with a small Einstein radius of $\sim 1''$. A small reduction in the number density is observed at larger distances (small magnifications).
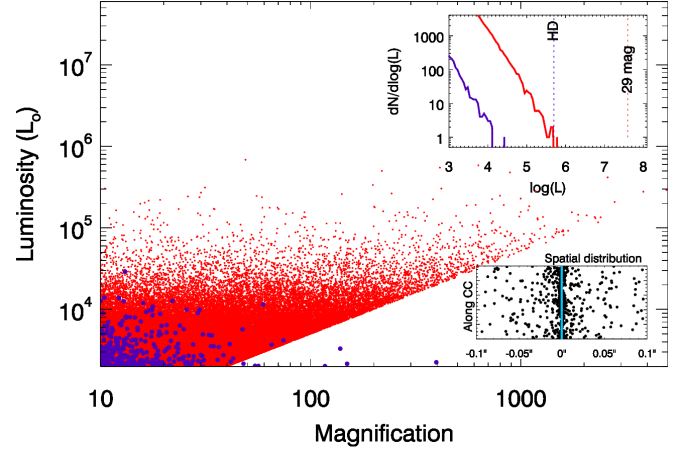
This reduction in number is due to the imposed HD limit in the Monte Carlo simulation.

The dependence on the exponent β is made more evident when we compare the previous result with the Monte Carlo simulation for the case with $\beta = 3$, and shown in Figure 11. Again, the bulk of the blue points falls below luminosities $10^4\,L_\odot$ and is not shown in the plot. The red points have a uniform distribution in magnification, as predicted by Equation 11. As before, the observed LF (thick red line in the upper-right inset plot) lies above the blue curve by a factor $\hat{\mu}^{\beta-1} = (2 \times 10)^2$. The spatial distribution shows a much different distribution than in the case with $\beta = 2$, with the number density directly tracing the magnification (Equation 8).

9.5. A possible invisible millilens in the Dragon arc

Since β can be estimated directly from the observed $\phi(\hat{L})$, or from the distribution of events as discussed above, combining the spatial distribution of the number density with the observed LF it is then possible to identify deviations that can be attributed to local departures from the macromodel magnification or regions in the source plane with a different LF. Departures from a smooth distribution in the number density can be taken as evidence for substructure, which can locally increase the number density of DTM stars (and transient events) according to $\rho(\mu,\beta) \propto \mu^{(\beta-2)}$. The number density of microlensing events is then a direct tracer of substructure, and can be used to map the underlying structure of DM fluctuations on subarcsecond scales and down to the milliarcsecond scale (see also the CC structure around millilenses in Williams et al. 2024, Fig.3).

In Section 4 we discussed how a millilens with mass as small as $2 \times 10^3\,M_\odot$ can boost the probability of microlensing by at least an order of magnitude when compared to the case of microlenses only (see Figure 4). Such a millilens and its associated Einstein ring would be too small to be resolved even with JWST, so all the microlensing events near the cusps of the millilens would seem to originate from the same pixel. Since the timescale for a single star with $\mu \approx 100$ to reach the closest microcaustic is about

1 yr (see next section), repeated observations with high cadence (weeks) should reveal the population of bright stars behind the microlens as each one crosses one of the multiple microcaustics around the millilens cusp. As discussed above, the LF of all events coming from this single pixel should be proportional to the LF of the background population of luminous stars, and with the same exponent β.

For larger millilens masses, and $\beta > 2$, the microlensing events around the millilens will take place in neighboring pixels and form a cluster of events. The clustering of detected events can be used to trace the underlying mass distribution of millilenses. Since microlensing events are most likely in supercritical regions, if a substructure in the far region becomes supercritical, microlensing events will more likely be detected around that region than in nearby subcritical regions. If enough events are detected in a lensed galaxy, a pattern emerges with clusters of events at the position of these substructures. We can approximate the size of a supercritical region of a substructure with mass $M_{\rm sub}$ as the area contained within its observed Einstein radius, which is given by $\Theta_{\rm obs} \approx \sqrt{\mu_{\rm 1m}} \times \Theta_E$ (Diego et al. 2018). If the substructure has circular symmetry, $\Theta_E = \sqrt{(4GM_{\rm sub}/c^2)(D_{ds}/(D_d D_s))}$, and the mass of the substructure can be obtained as

$$M_{\rm sub} = \Theta_{\rm obs}^2 \frac{c^2}{4G\mu_{\rm 1m}} \frac{D_d D_s}{D_{ds}}. \qquad (19)$$

Events 3 and 6 in Figure 1 are located in the far region but separated by only 0.15″. At this position $\mu_{\rm 1m} = 13$ from our lens model. If we assume the observed Einstein radius is half the separation between the events, from the relation above we would obtain a mass for the possible undetected substructure along the line of sight to these events of $\sim 1.3 \times 10^8\,M_\odot$ within the Einstein radius of the substructure. The virial mass of the substructure could be significantly higher if it is not concentrated enough to contain most of its mass within its Einstein radius. From the N-body simulations discussed in Section 10.4, we expect of order 1 satellite galaxy with this mass and overlapping with the Dragon arc. This is a tantalizing result, but it cannot be taken too seriously because it lacks statistical significance, and the two events mentioned above could simply be a chance occurrence of two microlensing events that happen to take place near each other. However, if substructures in this mass range exist in the cluster in the far region of the Dragon arc, they will become more evident with future observations of this arc, since new events will have a higher tendency to appear in positions near previous events. Future JWST observations may also reveal the hidden substructure that is increasing the microlensing rate at this location, or alternatively there is an overdensity of luminous stars at this position in the Dragon arc that can also increase the rate of microlensing events there. LBV can be distinguished from genuine microlensing events through their light curves, since they have light curves that depart from the $1/\sqrt{t - t_o}$ behaviour expected for microlensing events (t is time and t_o is the crossing time if the microcaustic).

10. Discussion

It is important to discuss some of the approximations in this work and consider aspects that have not been treated in the previous sections but affect some of our conclusions.
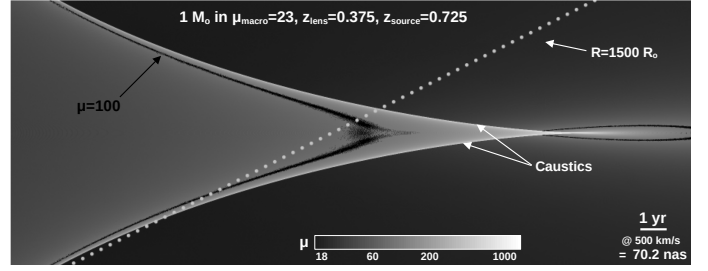

Fig. 12. Cusp region around a microcaustic with $1\,M_\odot$ in a macromodel with magnification $\mu = 23$. The gray scale shows the magnification. The region near the caustic with magnification ~ 100 is marked in black. The white bar in the bottom right is the distance moved by a background star at $z = 0.725$ in 1 yr when the relative velocity is $500\,{\rm km\,s^{-1}}$. The maximum magnification is 3700 at the tip of the cusp. The pixel size is $320\,R_\odot$, so a supergiant star with diameter $80\,R_\odot$ would reach twice this magnification at the maximum, or $\mu_{\rm max}^{\rm tip} \approx 7400$. At the fold caustics, the maximum magnification for a $80\,R_\odot$ diameter star would be smaller, $\mu_{\rm max}^{\rm fold} \approx 1000$. The straight line with dots shows the track of a hypothetical star moving across the caustic. The size of the dot corresponds to a star with radius $1500\,R_\odot$.

10.1. Duration of microlensing events

At $\mu = 100$, a star with absolute magnitude -7 would still be undetected within the far and near regions (apparent magnitude 31.24 for $\mu = 100$). A few of these undetected stars will be close enough to a microcaustic in the source plane. Approximately half of these stars will be moving away from the microcaustic and hence remain undetected in the near future, but the other half will be moving closer to the microcaustic and become brighter over time. At magnification $\mu = 1000$, the same star with absolute magnitude -7 can now be detected easily with JWST in exposures of 1 hr. For a background star near a microcaustic, the time it takes to move from $\mu = 100$ to $\mu = 1000$ depends on several factors such as the mass of the microlens, the macromodel magnification, the relative velocity between the background star and the microcaustic, the direction of motion relative to the microcaustic, and the point of crossing of the microcaustic. To get a sense of this timescale, we assume a microlens with mass $1\,M_\odot$, the same macromodel magnification of the fiducial model $\mu_{\rm 1m} = 23$, and a relative velocity of $v\cos\alpha = 500\,{\rm km\,s^{-1}}$, where α is the angle between the direction of motion and the microcaustic. In this situation, when the background star is at magnification $\mu = 100$, the microcaustic is $\sim 0.1\,\mu$as away as shown in Figure 12.

At a velocity of $v\cos\alpha = 500\,{\rm km\,s^{-1}}$, a star with absolute magnitude -7, within a microcusp and with magnification $\mu \approx 100$, would take ~ 1 yr to reach the caustic and become detectable. At the tip of the cusp, the magnification for a supergiant star with diameter $80\,R_\odot$ reaches a maximum of ~ 7500, and ~ 1000 at the fold caustics. At this velocity, this maximum magnification can be maintained for ~ 1.5 days Miralda-Escude (1991), after which the magnification will drop to a factor of a few and the star will no longer be detectable. Approximately 1/3 of the background stars in the far region with absolute magnitude -7, and near a microcaustic with magnification $\mu \approx 100$, will move toward the caustic and become detectable after 0.5–1 yr of observation (or similarly, they are detectable now and will disappear behind the microcaustic after 1 yr or move away from the caustic). The exact same reasoning applies for the more numerous stars with double the magnification, or absolute magnitude $-7 + 2.5\log_{10}(2) = -6.25$, but in an area eight times
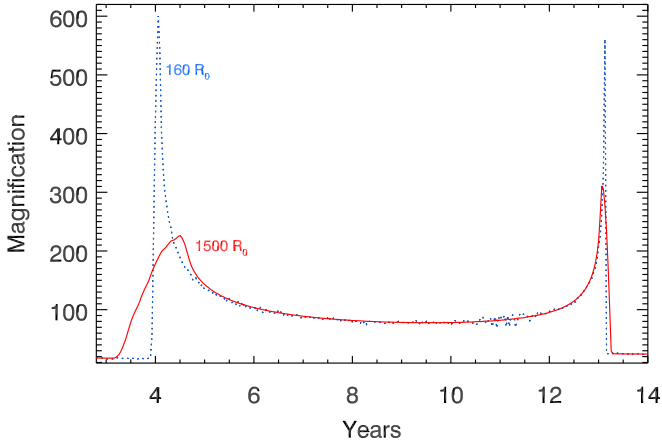
Fig. 13. Light curve of a star moving at $500\,\mathrm{km\,s^{-1}}$ along the track shown in Figure 12. The blue dotted line corresponds to a star with the same diameter as the pixel of the simulation ($320\,R_{\odot}$), while the red solid line corresponds to a much larger and more luminous star with $R = 1500\,R_{\odot}$. The star crosses the entire caustic region in ~9 yr, with the first peak grazing the caustic and producing a wide peak.

smaller in the source plane ($A/d\mu \propto \mu^{-3}$) and with magnification $\mu > 200$. But in this case, the distance to the microcaustic is four times smaller, so the cadence should be higher in order to detect these stars before they cross the caustic and become undetectable again.

One factor to keep in mind is the very large radius of the most luminous RSGs, that can reach radii of ~$1500\,R_{\odot}$ (Meynet et al. 2015). Since the maximum magnification is lower for larger stars, these very luminous RSGs would generally have smaller magnification factors, typically a few hundred as shown in Figure 13. When addressing the detectability of these stars, care needs to be taken to account for the smaller maximum magnification of large RSGs. Interestingly, the large radii of massive RSGs should correlate with where they are observed. If an RSG has a radius of $1500\,R_{\odot}$ and it can only be magnified by factors of a few hundred in the far region, only those RSGs exceeding $10^5\,L_{\odot}$ can be observed at $z = 0.725$ during peak magnification. Less luminous RSGs of similar size can still be observed in the near region since multiple microcaustics overlap and the magnification can be $\mu > 2000$ in this case (or 2.5 mag deeper). The most luminous RSG is expected to have a maximum luminosity close to the HD limit, $L_{\max} \approx 5 \times 10^5\,L_{\odot}$ or absolute magnitude ~ -9.5, which at magnification 100 can still be detected at $z = 0.725$ with apparent magnitude ~28.7, so anywhere in the far region provided they are close enough to a microlens or millilens. BSGs, on the other hand, can be even more luminous and smaller ($R \lesssim 25\,R_{\odot}$), so during a microlensing event they can be magnified by factors $\mu > 1000$ in the far region, allowing us to see the fainter (but more magnified) BSGs or the brighter but with more moderate and more likely magnification factors ($\mu < 50$).

10.2. Critical magnification

The definition of the far-to-near ratio depends on the ratio of areas near to and far from the cluster CC, which in turn depends on our choice for the critical magnification. Taking a larger $\mu_{\rm crit}$ by a factor of three would lower the orange (and green) bands in Figure 7 and Figure 8 by a factor of 9, but it would also lower the solid lines by the same factor of 9, leaving our conclusions unchanged. Hence, our conclusions are independent of the par-

ticular value of $\mu_{\rm crit}$. There is, however, a relatively small dependence on $\mu_{\rm crit}$ impacting our findings. The results in Figure 7 and Figure 8 are normalized to the area in the far region, which for $\mu_{\rm crit} = 100$ represents 83.4% of the area covered by the Dragon arc. Lowering $\mu_{\rm crit}$ would reduce this fraction of area where far events from millilenses can take place. The dependence of the fraction of area in the near region scales almost linearly with $\mu_{\rm crit}$, so we can approximate this as $F_{\rm far} \approx 1 - 16.7/\mu_{\rm crit}$, where 16.7 is the fraction of area in the near region when $\mu_{macro} = 100$. For example, from this law we find $F = 66.6\%$ and $F = 91.6\%$ for $\mu_{\rm crit} = 50$ and $\mu_{\rm crit} = 200$, respectively, while from the lens model we find 69.3% and 93.2%. Hence, in the conservative case where $\mu_{\rm crit} = 50$ (this would require a very high $\Sigma_{\rm crit} > 100\,M_{\odot}\,{\rm pc}^{-2}$), we find that the lines in Figures 7 and 8 would be corrected by a factor of $66.6/83.4 = 0.8$, while for $\mu_{\rm crit} = 200$, the same curves would move upward by a small amount $91.6/83.4 = 1.1$, leaving our results virtually unchanged.

10.3. Slope of the lensing potential

Related to the previous point, another source of uncertainty impacting our results is the specific properties of the cluster lens model, in particular the slope of the lensing potential. The macromodel magnification enters in Equation 2 quadratically. If the macromodel magnification in the far region is three times larger, this would increase the amplitude of the solid lines in Figure 7 by a factor of 9, bringing the prediction from millilenses and the observation to better agreement. The median and mean magnifications of the WSLAP+ model in the far region of the Dragon arc are 19.2 and 27.7, respectively. Increasing the magnification by a factor of three would bring the most common values of μ_{1m} in Equation 2 close to the value of $\mu_{\rm crit}$, resulting in a very uniform distribution of microlensing events along the Dragon arc. Such a lens model would require a very shallow lensing potential, possibly in conflict with lensing constraints. Comparing our lens model magnification in the far region with the predicted magnification from the lens models in the same region of Li et al. (2024), we find that on average those models predict $24\% \pm 22\%$ more magnification than our lens model in the far region. Based on this, and taking the upper limit (46% increase), we expect the solid colored lines in Figure 7 to increase by a factor of ~2.1, still insufficient to explain the low rate of predicted events.

10.4. Number density of millilenses

N-body simulations show a tight correlation between the virial mass of the cluster and the number of GCs, $M_{\rm vir} = 5 \times 10^9\,M_{\odot} \times N_{\rm GC}$ (Burkert & Forbes 2020; Valenzuela et al. 2021). The galaxy clusters in which transient stars have been found are all very massive, with virial masses ~$10^{15}\,M_{\odot}$. Hence, we expect ~2×10^5 GCs in each of these clusters. It is difficult to estimate with precision the expected number density of GCs (detected and nondetected) at the positions of the transients, but we can get an order-of-magnitude estimate and see if it is in agreement with the observed densities in nearby clusters.

If we assume that the distribution of GCs follows a cored isothermal profile, then the number density of GCs falls with distance to the center as ~$(R_c + R)^{-1}$. Assuming all ~2×10^5 GCs in the cluster are within a radius of 1 Mpc with this profile, we find that the number densities at $R = 50\,\mathrm{kpc}$ and $R = 70\,\mathrm{kpc}$ (measured from the BCG in the south of A370, which is the center of the South group in this merging cluster) vary between ~0.76 and

~1.13 per kpc^2 when the core radius (R_c) varies between 0 and 10 kpc. This is within a factor of two of what was assumed in Figure 7, and hence this higher estimate of the number density is still insufficient to explain with millilenses the anomalously high observed ratio of far-to-near microlensing events. Modifying the radial profile to a steeper one with number density scaling as $(R_c + R)^{-2}$ (similar to what would be expected if the distribution of GCs follows the profile from Navarro et al. 1996), and for the most favorable scenario with $R_c = 0$, the number density increases to ~1.6 and ~2.9 GCs per kpc^2 at 70 kpc and 50 kpc (respectively) from the center of the halo. Alternatively, we measure the number density of compact unresolved objects in the vicinity of the Dragon arc using the most sensitive F200W image of the new JWST data of this cluster. We find a number density of ~0.1 object per kpc^2 (see also Berkheimer et al. 2024). These compact objects correspond to the massive end of the tail of the distribution in Figure 14, with masses above a few times $10^6\,M_\odot$ (see the masses derived by Faisst et al. 2022, for similar objects). From the same figure, we estimate (conservatively) that the number of missing objects is a factor of ~10 larger, which would bring the number density of GCs near the Dragon arc to ~1 object per kpc^2. The very same conclusion is reached by Saifollahi et al. (2024), based on Euclid observations of the much closer Fornax cluster; they reach high completeness up to ~24 mag in Euclid's Y, J, and H bands (Fornax is at just 20 Mpc, so this depth is equivalent to reaching ~29 mag with JWST), finding a maximum number density of GCs of ~1 GC per kpc^2. This is still insufficient, since it would place the ratio of far-to-near events one order of magnitude below the observed rate.

The mass function in Figure 6 excludes halos more massive than a few times $10^7\,M_\odot$. Naturally, we expect halos in this mass range to still contribute as millilenses, but it is unclear how many of those exist, since their potential must be shallow enough to not contain dense concentrations of stars in their central regions, hence evading direct detection.

Nevertheless, dwarf galaxies, or small satellites in general, are expected to be numerous in cluster environments and introduce perturbations in the magnification in the far region (and also in the near region). From the lensing point of view, their cored structures and relatively low mass make many of them subcritical (that is, they do not produce CCs). However, the fraction of critical to subcritical halos remains unknown in cluster environments, so it is difficult to accurately predict their contribution to the probability of high magnification. Even if they do not reach criticality, at distances of ~1″ from the cluster CC, the mass associated with a satellite may be enough to alter the inverse of the magnification $\mu_{2m}^{-1} = (1 - \kappa_{2m})^2 - \gamma_{2m}^2$ near the satellite, and bring it close to the small value needed for microlensing events to be maximized, $|\mu_{2m}^{-1}| \approx 10^{-2}$.

Here we rely on results from pseudo-analytical realizations based on state-of-the-art recipes calibrated using numerical N-body simulations to assess the contribution from undetected satellite galaxies. We employed a sample of 16 very-high-resolution realizations using the MOKA algorithm (Giocoli et al. 2012, 2016) assuming a mass of $10^{15}\,M_\odot$ at $z = 0.37$. The cluster-size halo is populated by Monte Carlo sampling the sub-halo mass function measured by (Giocoli et al. 2010) and extrapolating it to $1.5 \times 10^5\,M_\odot$. The halos have a triaxial model (Jing & Suto 2002; Despali et al. 2014) and subhalos are spatially distributed as calibrated by Gao et al. (2004).

The resulting mass function from satellites is shown in Figure 14, where we compare it with the mass function of GCs used in our main result (Model 1). For this comparison, we have
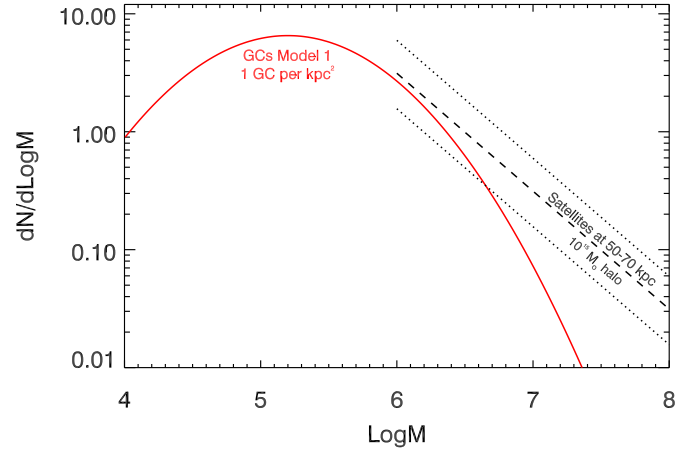

Fig. 14. Comparison of mass functions of GCs and satellites. The red solid line shows the total number of GCs in the area corresponding to the Dragon arc from the GC mass function, after normalizing it to a density of 1 GCs per kpc^2. The black dashed line shows the mass function from N-body simulations of a population of satellites in a cluster with $M_{vir} = 10^{15}\,M_\odot$ and computed in a similar area at distances between 50 and 70 kpc from the center of the halo (same distance of the Dragon arc). The dotted lines shows the dispersion in the number of satellites from 16 different realizations. In most realizations, no halos more massive than ~$10^8\,M_\odot$ are found within the region considered. The total mass from the GC mass function is $4.01 \times 10^8\,M_\odot$ and the total mass from the satellite mass function is $3.07 \times 10^8\,M_\odot$. Finally, the sum of mass from the GC mass function up to $10^6\,M_\odot$ and the mass from the satellite mass function above $10^6\,M_\odot$ is $5.01 \times 10^8\,M_\odot$.

computed the total number of GCs in the area of the Dragon arc (~960 kpc^2) and assuming the number density of GCs is 1 per kpc^2, roughly the upper limit of the vertical blue band in Figure 7. The mass function of satellites is also normalized to the same area covered by the Dragon arc, and it corresponds to the abundance of satellites with masses larger than $10^6\,M_\odot$ found at distances between 50 and 70 kpc from the cluster center in simulated clusters at $z \approx 0.37$, and with virial mass $10^{15}\,M_\odot$. Rather than repeating the calculation we did for GCs, we can estimate the contribution from these satellites in the most favorable situation. We assume that all satellites are supercritical and compact enough so they contribute to the magnification similarly to GCs – that is, they follow the scaling of Equation 2. This is an optimistic scenario because a fraction of these satellites will be subcritical. In fact, it is unlikely that a large fraction of them are supercritical since this would imply they have dense detectable cores, and none is clearly observed as a resolved source in the vicinity of the Dragon arc. Nevertheless, under the ideal assumption above, the upper-limit contribution to the area in Equation 4 from the satellites should be proportional to their total integrated mass. We compute this mass from the dashed-line model shown in Figure 14 and find a total mass of $3.07 \times 10^8\,M_\odot$ in the area occupied by the Dragon arc. Repeating the same calculation for the red solid curve in Figure 14, we find that the GCs contribute $4.01 \times 10^8\,M_\odot$ in the same area, or 30% more. Considering instead a combined mass function composed of the red curve up to $10^6\,M_\odot$ and the dashed black line above this mass, the total mass is $5.01 \times 10^8\,M_\odot$, or 25% more than the GC contribution. Translating these numbers into Figure 7, the red curve (Model 1) would move upward by only a factor of 1.25. Hence, even in the most optimistic case in which satellites are very compact and supercritical, the contribution from the undetected satellite galaxies is relatively minor.

10.5. Substructure along the line of sight

So far we have assumed the milli-lens substructure capable of promoting a micro-lensing event to detectability lies inside the virial radius of the cluster. However, CDM also predictions a sizeable population of halos along the line of sight (LOS), which can also contribute to the lensing perturbations (e.g., Gilman et al. 2019). On average, we expect substructures to contribute to the surface mass density an amount similar to the contribution from the mean density of the universe, $\bar{\rho} = \Omega_m \times \rho_{\rm crit}$. In CDM, the overwhelming majority of dark matter halos on the relevant mass scales are subcritical, and the lensing effects of these objects drop when placed close to the observer and source. Therefore, we consider LOS contributions from the redshift range $0.15 < z < 0.5$, assuming the source is the Dragon galaxy (at $z = 0.725$). In this interval, the critical density of the universe is, on average, 1.41 times higher than at $z = 0$, $\rho_{\rm crit}(z = 0) = 2.77 \times 10^{11} \times h^2\, M_\odot {\rm Mpc}^{-3}$. Projecting along the line of sight (1268 Mpc comoving), we then get an average contribution of $\Sigma_{LOS} \sim 242.7\, M_\odot\, {\rm pc}^{-2}$. Unlike subhalos of the cluster, dark matter halos outside the cluster environment are not affected by tidal forces.

To assess the contribution from halos outside the virial radius of the cluster, we calculate the expected number of dark matter halos in the mass range 10^5–$10^9\, {\rm M}_\odot$ using the mass function model presented by Sheth & Tormen (1999). In addition to the halos drawn from the Sheth-Tormen mass function, we account for correlated structure around the cluster. The mass of the cluster 10^{14}–$10^{15}\, {\rm M}_\odot$ causes a local enhancement to the density field that increases the number of dark matter halos within $\sim 5\, {\rm Mpc}^{-1}$; these objects are effectively at the same redshift as the cluster itself, but are not inside the virial radius or even necessarily bound to the cluster potential, and thus they are typically not included in satellite mass functions. We model the local enhancement through the two-halo term (Gilman et al. 2019), with the additional correction proposed by Lazar et al. (2021). The halos along the line of sight, including those corresponding to correlated structure around the cluster, contribute 100–$200\, {\rm M}_\odot\, {\rm pc}^{-2}$, depending on the assumed virial mass of the cluster. This constitutes a significant contribution that can potentially impact the results discussed in previous paragraphs. From Figure 10, an increase of this magnitude in Σ would make the amount of substructure in the far region very large, $\Sigma_{\rm Tot} = \Sigma_* + \Sigma_{\rm LOS} \approx 150\, M_\odot\, {\rm pc}^{-2}$, which would imply a ratio of events of ~ 1 in the far and near regions, and in conflict with the observed ratio of ~ 2.5. Alternatively, for a fixed ratio of events ~ 2.5, a larger value for $\Sigma_{\rm Tot}$ would imply a large value of $\beta \gtrsim 3$. However, in CDM dark matter halos on these scales are predicted to be subcritical for lensing, even when placed on top of a cluster convergence map, and thus they contribute subdominantly to the lensing magnification perturbations required to boost the signal from microlenses. We will be able to revisit this topic with future measurements of more events in the Dragon arc that will enable measurements of β directly, and provide a better estimation of the ratio of events in the near and far regions, resulting in a constraint on the contribution from microlenses and millilenses in the lens plane and in the LOS to $\Sigma_{\rm Tot}$.

10.6. Presence of hyperluminous stars

The events found in the Dragon galaxy by the Flashlights program are observed in only one filter, so unfortunately we lack color information to assess whether these events could be LBVs at moderate magnifications or microlensing events of the much more abundant but fainter supergiant stars. In low-redshift arcs such as the Dragon arc, hyperluminous stars with absolute magnitude -11 or brighter could be seen anywhere around the arc in regions with moderate magnification factors of $\mu_{1m} > 20$. Such high luminosity can be reached, for instance, during an outburst of an LBV reaching an absolute magnitude of -11 (Weis & Bomans 2020), or apparent magnitude ~ 29 with magnification $\mu \approx 20$. Outbursts as luminous as -14 mag have been recorded (Pastorello et al. 2010), or even brighter for the so-called "supernova impostors" that can be as luminous as supernovae (e.g., Kilpatrick et al. 2018). These superluminous LBVs are exceedingly rare and can be detected anywhere at this redshift without the help of magnification, so we do not consider them here. In the case of typical LBV outbursts, they would be more likely detected in the far region since this corresponds to a larger area in the source plane.

In the Dragon arc, Li et al. (2024) estimate ~ 3 LBVs should be present. We can independently estimate the number of LBVs if we assume the Dragon galaxy contains a similar number of LBVs as the number found in our neighborhood. The number of LBVs with absolute magnitude brighter than -10 found in the Milky Way plus LMC plus SMC is $N_{\rm LBV} \approx 100$ (Humphreys & Davidson 1979; Hamann et al. 2006; Crowther et al. 2010; Hainich et al. 2014). To see one of these stars without help from microlensing or millilensing, the magnification needs to be $\mu_{1m} \approx 50$. Only a portion of the Dragon arc has magnification > 50. We can estimate this by multiplying the upper bound of the orange region in Figure 7 by a factor of four – that is, $\sim 5\, {\rm kpc}^2$ in the Dragon arc are magnified by a factor of 50 or more. This corresponds to a fraction of $6.3\% \times (R_{\rm gal}/5{\rm kpc})^2$, where we have adopted the estimated radius of the Dragon galaxy from our lens model, $R_{\rm gal} \approx 5\, {\rm kpc}$. The number of expected LBVs in this area is then $N \approx 6$, close to the estimate from Li et al. (2024). Most of these LBVs likely will be in a quiescent phase and hence not detected as transients when comparing observations separated by 1 yr (or ~ 0.5 yr in the source frame), but for observations at two epochs separated by several years, a significant fraction of them will show measurable changes in flux and be identified as transients. We conclude that some of the events found in the far region of the Dragon arc may be LBVs, but without color information we cannot confirm this hypothesis.

At higher redshifts, larger magnification factors are needed to see outbursting LBVs, so the expectation in this case is to see mostly genuine microlensing events in the far region. An example (but also an exception) is Godzilla, a star which is believed to be an outbursting LBV at $z = 2.37$ with at least five counterimages (Diego et al. 2022), all of them (but one) are undetected at macromodel magnification $\mu_{1m} \approx 100$ (the example), but interestingly with one being detected thanks to the magnification boost provided by a millilens ($\mu_{2m} \gtrsim 2000$) (the exception). Diego et al. (2022) estimate that at any given point ~ 30 extremely magnified LBVs (EMBLVs) at $1 < z < 3$ and with magnification >1000 should be detectable in the sky and with apparent magnitudes as bright as 24. Large-scale high-cadence surveys such as LSST can reveal them and complete a census of EMBLVs up to $z \approx 3$.

10.7. Alternatives to ΛCDM

Although the combination of standard microlensing (that is, not involving a millilenses) and LBVs offers the simplest explanation for the high ratio of events found in the far region, it is interesting to consider other scenarios in which dark matter physics,
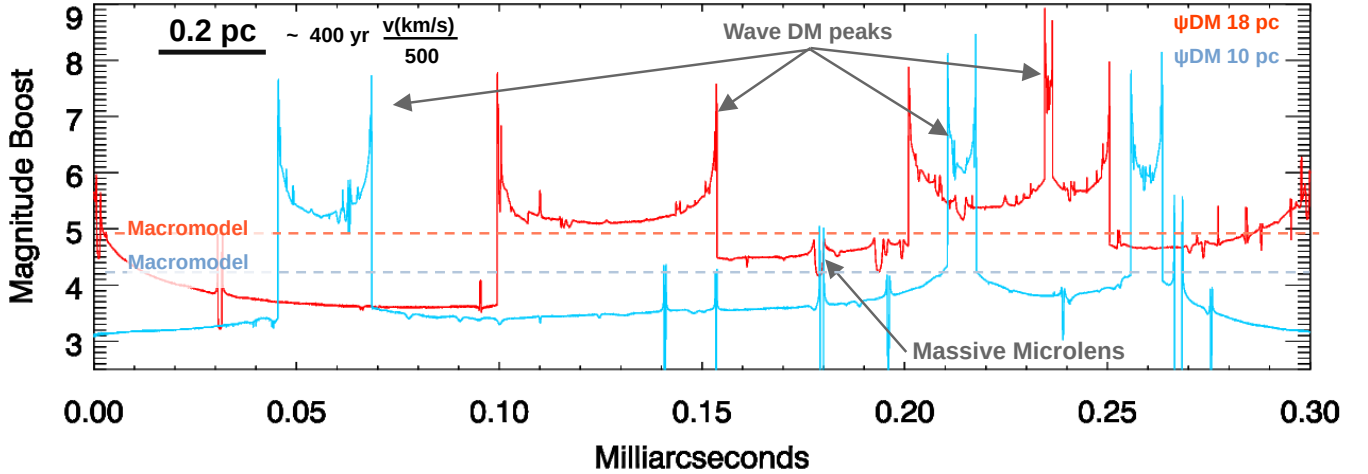
Fig. 15. Simulated magnification (in magnitudes) for a source traversing a caustic region. The lens model includes ψDM with de Broglie wavelength 18 pc (red) or 10 pc (blue). Two macromodel magnifications are shown: $\mu_{\mathrm{macro}} \approx 100$ (red) and $\mu_{\mathrm{macro}} \approx 50$ (blue). The effects of different lens components are marked. For this figure, the number density of microlenses was reduced to just $1\,M_\odot\,\mathrm{pc}^{-2}$ in order to better show the effect of ψDM. Even at moderate values of μ_{macro} (blue), the effects of ψDM are very different from standard CDM expectations.

or various baryonic effects (e.g., Ragagnin et al. 2024), alter the properties of halos. For example, warm DM models predict less substructure on subgalactic scales. Surviving halos in warm DM have lower concentrations than their CDM counterparts, and therefore have a suppressed lensing efficiency, lowering the contribution from millilenses to the lensing probability. On the other hand, self-interacting DM can cause halos to undergo core collapse, a process that dramatically raising their central density, potentially to a degree that causes them to become super-critical for lensing (Gilman et al. 2021). Alternatively, wave dark matter, (ψDM) is expected to increase the magnification in the far region. In this model, DM has density fluctuations at scales given by the de Broglie wavelength and the halo mass (Schive et al. 2016), from the dependence on momentum:

$$\lambda_{dB} = 15 \left(\frac{10^{-22}\,\mathrm{eV}}{m_\psi} \right) \left(\frac{10^{15}\,M_\odot}{M_{\mathrm{cluster}}} \right)^{1/3}\,\mathrm{pc}\,, \qquad (20)$$

where m_ψ is the mass of the ultralight axion-like particle (ALP). For masses $m_\psi \approx 10^{-22}\,\mathrm{eV}$ and a $10^{15}\,M_\odot$ cluster, this scale corresponds to 3 mas in the lens plane. This pervasive interference substructure causes the CC to become corrugated on the de Broglie scale (Chan et al. 2020; Laroche et al. 2022; Amruth et al. 2023), and increasingly so for more massive halos, with many detached islands where the magnification diverges at relatively large offsets from the cluster CC (Amruth et al. 2023; Laroche et al. 2022; Powell et al. 2023).

Similarly to microlenses, ψDM fluctuations are ubiquitous across the lens plane, and as in the case of microlenses and millilenses, these fluctuations get amplified near the CC by the macromodel. In Figure 15 we show the effect of ψDM over a small region in the observer plane. For this particular case the source is at $z = 1$, but the effect would be very similar for $z = 0.725$. The simulation of ψDM follows Amruth et al. (2023), and for this particular case the value of Σ_* has been decreased to $\Sigma_* = 1\,M_\odot\,\mathrm{pc}^{-2}$, to better appreciate the ψDM effect. Two models for ψDM are considered with $\lambda_{\mathrm{dB}} = 10\,\mathrm{pc}$ and $\lambda_{\mathrm{dB}} = 18\,\mathrm{pc}$, for cluster-scale lenses. We also consider two macromodel magnifications. As shown in the figure, ψDM introduces perturbations in the magnification pattern in the source plane at the sub-milliarcsecond level. This scale is similar to the scale of the caustics from the GCs considered earlier, and is consistent

with results from analyses of multiply-imaged quasars in which the effect from ψDM is shown to be comparable to the effect of population of halo millilenses (Laroche et al. 2022). Interestingly, smaller masses for the ALP (blue curve) result in more pronounced effects but over a smaller region. In both cases, a significant portion of the source plane can attain sufficiently high magnifications so the critical magnification is reached, maximizing the probability for microlensing effects to take place.

The de Broglie wavelength (hence the mass of the ALP) and the macromodel magnification determine the type of object that can exhibit different magnifications. For instance, in the $\lambda_{\mathrm{dB}} = 10\,\mathrm{pc}$ model and $\mu_{1m} = 50$ in Figure 15, the scale of the object needs to be typically larger than $\sim 0.1\,\mathrm{pc}$ in order to be insensitive to ψDM fluctuations, while for the $\lambda_{\mathrm{dB}} = 18\,\mathrm{pc}$ and $\mu_{1m} = 50$ model the source needs to be larger than $0.5\,\mathrm{pc}$ in order to not exhibit asymmetric fluxes. Future observations of the Dragon arc will reveal additional microlensing events, which are expected to form clusters of microlensing events around the strongest fluctuations in the boson field.

11. Conclusions

We study the 3M lensing effect from the combination of a macromodel, a millilens, and microlenses. The possibility is considered that microlensing events found at relatively large distances from the cluster CC in the Dragon arc, or far region, are aligned with millilenses in the lens plane that increase the probability of microlensing. We study the scaling of the area above certain magnification (or lensing probability) near a millilens, with the mass of the millilens and macromodel magnification, with and without adding microlenses. Near the cusps of millilenses, this probability scales with the mass of the millilens, and microlenses play a minor role. We consider a realistic population of millilenses and model their mass function with a log-normal function, then compute the total area in the far region of the source plane associated with this population of millilenses that has magnification greater than some critical value. We find that the contribution to this area from millilenses is less than the contribution from the far more numerous microlenses elsewhere in the source plane. Hence, the addition of millilenses does not appreciably increase the expected rate of microlensing events

far from the critical curve (which is given mostly by the more numerous microlenses). Other factors, such as the presence of LBVs, also contribute to the number of transient events in the far region, especially in lensed galaxies at low redshift where LBVs can be detected even at modest magnification factors.

We pay special attention to the spatial distribution of microlensing events and find that the number density of microlensing events also depends on the exponent of the LF, $\rho(\mu, \beta) \propto \mu^{\beta-2}$. We make the analogy of traditional photographic-plate imaging and identify two regimes: (i) positive-imaging regime when $\beta > 2$ and the number density of microlensing events is higher around massive substructures (high μ), and (ii) negative-imaging regime when $\beta < 2$ where microlensing and microlensing events have smaller number densities at the position of massive substructures (also high μ).

We discuss the intimate relationship between the abundance of DTM stars and the number of observed microlensing events where the second is proportional to the former. We demonstrate, both analytically and with Monte Carlo simulations, how the number density of DTM stars shows a strong dependence on the LF and the macromodel magnification. Once the population of DTM stars has been established (from the LF and the macromodel magnification), the problem of estimating the number of microlensing events can be reduced to studying a population of DTM stars as they move across the web of microcaustics, where the later depends not only on the amount of substructure (microlenses and millilenses), but also on the macromodel magnification. We use the observed density of events in the far and near regions of the Dragon arc to derive the slope of the LF, finding that a steep LF with $\beta = 2.55^{+0.72}_{-0.56}$ is consistent with the observations. Variation of the LF along the lensed Dragon Arc or absorption by dust are not considered in this work but they should add an additional element of uncertainty in the results. With future data, one can measure the slope β directly from the observed LF and confront it with our estimate of $\beta = 2.55^{+0.72}_{-0.56}$ derived from the spatial distribution of the number density of microlensing events.

We derive a relation between the slope of the LF, β, the amount of substructure, Σ, and the ratio of observed microlensing events in the near and far regions, $N_{\text{near}}/N_{\text{far}}$. We estimate the amount of substructure along the line of sight and, from the relation between β, Σ, and $N_{\text{near}}/N_{\text{far}}$, we argue that most of this substructure should be in the form of subcritical halos. Otherwise, the inferred values of β would be very high.

Small substructures in the far region of the CC can be mapped (imaged) by measuring this number density of microlensing events, which should correlate with the location of millilensing substructures. The clustering may also reveal a non-uniform distribution of the background stellar population that can equally show clustering. Repeated observations of the same arc may be and a detailed analysis of the photometry (or spectra if available) may be needed in order to clearly distinguish between the two scenarios. We apply this technique to two microlensing events forming a pair of local high density, and under the assumption of a uniform distribution of the background stars, find that if this peak in the density of microlensing events is due to a substructure, its mass is $\sim 1.3 \times 10^8\, M_\odot$ within its Einstein radius. This technique shall open a new window to map the distribution of mass on scales of milliarcseconds, including perturbations in the DM field. As an illustration, we consider the case of ψDM and argue that this type of model can be proven with repeated observations of low-redshift caustic-crossing arcs, such as the Dragon arc, thereby greatly increasing the statistics on the spatial distribution of microlensing

events and revealing the hidden nature of DM at subarcsecond scales.

At the time of submission of this paper, new JWST observations have revealed more than 40 microlensing candidates in the near and far regions of the Dragon arc. Most of these events are suspected to be due to RSGs at $z = 0.725$. These events are presented by Fudamoto et al. (2024). Here we do a simple estimation based on the new findings by Fudamoto et al. (2024) and conclude that microlenses with a surface mass density $\Sigma_* = 54\, M_\odot\, \text{pc}^{-2}$ offer a good explanation for the distribution of microlensing events in the new JWST data. Detailed analysis of these new events will be the subject of a future paper.

References

Adamo, A., Bradley, L. D., Vanzella, E., et al. 2024, ArXiv e-prints [arXiv:2401.03224]

Amruth, A., Broadhurst, T., Lim, J., et al. 2023, Nat. Astron., 7, 736

Arvanitaki, A., Dimopoulos, S., Galanis, M., et al. 2020, Phys. Rev. D, 101, 083014

Baumgardt, H., & Hilker, M. 2018, MNRAS, 478, 1520

Berkheimer, J. M., Carleton, T., Windhorst, R. A., et al. 2024, ApJ, 964, L29

Blandford, R., & Narayan, R. 1986, ApJ, 310, 568

Borgeest, U., von Linde, J., & Refsdal, S. 1991, A&A, 251, L35

Bragaglia, A., Carretta, E., D'Orazi, V., et al. 2017, A&A, 607, A44

Broadhurst, T. J., Taylor, A. N., & Peacock, J. A. 1995, ApJ, 438, 49

Broadhurst, T., Li, S. K., Alfred, A., et al. 2024, ArXiv e-prints [arXiv:2405.19422]

Burkert, A., & Forbes, D. A. 2020, AJ, 159, 56

Canizares, C. R. 1981, Nature, 291, 620

Chabrier, G. 2003, PASP, 115, 763

Chan, J. H. H., Schive, H.-Y., Wong, S.-K., Chiueh, T., & Broadhurst, T. 2020, Phys. Rev. Lett., 125, 111102

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, ApJ, 881, 8

Chen, W., Kelly, P. L., Treu, T., et al. 2022, ApJ, 940, L54

Crowther, P. A., Schnurr, O., Hirschi, R., et al. 2010, MNRAS, 408, 731

Cyr-Racine, F.-Y., Keeton, C. R., & Moustakas, L. A. 2019, Phys. Rev. D, 100, 023013

Dai, L., Venumadhav, T., Kaurov, A. A., & Miralda-Escude, J. 2018, ApJ, 867, 24

Despali, G., Giocoli, C., & Tormen, G. 2014, MNRAS, 443, 3208

Diego, J. M. 2019, A&A, 625, A84

Diego, J. M., Protopapas, P., Sandvik, H. B., & Tegmark, M. 2005, MNRAS, 360, 477

Diego, J. M., Tegmark, M., Protopapas, P., & Sandvik, H. B. 2007, MNRAS, 375, 958

Diego, J. M., Broadhurst, T., Chen, C., et al. 2016, MNRAS, 456, 356

Diego, J. M., Kaiser, N., Broadhurst, T., et al. 2018, ApJ, 857, 25

Diego, J. M., Pascale, M., Kavanagh, B. J., et al. 2022, A&A, 665, A134

Diego, J. M., Meena, A. K., Adams, N. J., et al. 2023a, A&A, 672, A3

Diego, J. M., Sun, B., Yan, H., et al. 2023b, A&A, 679, A31

Diego, J. M., Li, S. K., Meena, A. K., et al. 2024a, A&A, 681, A124

Diego, J. M., Adams, N. J., Willner, S. P., et al. 2024b, A&A, in press, https://doi.org/10.1051/0004-6361/202349119

Dike, V., Gilman, D., & Treu, T. 2023, MNRAS, 522, 5434

Faisst, A. L., Chary, R. R., Brammer, G., & Toft, S. 2022, ApJ, 941, L11

Fudamoto, Y., Sun, F., Diego, J. M., et al. 2024, ArXiv e-prints [arXiv:2404.08045]

Furtak, L. J., Meena, A. K., Zackrisson, E., et al. 2024, MNRAS, 527, L7

Gao, L., White, S. D. M., Jenkins, A., Stoehr, F., & Springel, V. 2004, MNRAS, 355, 819

Gilman, D., Agnello, A., Treu, T., Keeton, C. R., & Nierenberg, A. M. 2017, MNRAS, 467, 3970

Gilman, D., Birrer, S., Treu, T., Keeton, C. R., & Nierenberg, A. 2018, MNRAS, 481, 819

Gilman, D., Birrer, S., Treu, T., Nierenberg, A., & Benson, A. 2019, MNRAS, 487, 5721

Gilman, D., Birrer, S., Nierenberg, A., et al. 2020, MNRAS, 491, 6077

Gilman, D., Bovy, J., Treu, T., et al. 2021, MNRAS, 507, 2432

Gilman, D., Birrer, S., Nierenberg, A., & Oh, M. S. H. 2024, ArXiv e-prints [arXiv:2403.03253]

Giocoli, C., Tormen, G., Sheth, R. K., & van den Bosch, F. C. 2010, MNRAS, 404, 502

Giocoli, C., Meneghetti, M., Bartelmann, M., Moscardini, L., & Boldrin, M. 2012, MNRAS, 421, 3343

Giocoli, C., Bonamigo, M., Limousin, M., et al. 2016, MNRAS, 462, 167

Golubchik, M., Zitrin, A., Pierel, J., et al. 2023, MNRAS, 522, 4718

Gorghetto, M., Hardy, E., March-Russell, J., Song, N., & West, S. M. 2022, JCAP, 2022, 018

Goudfrooij, P., & Fall, S. M. 2016, ApJ, 833, 8

Graham, P. W., Mardon, J., & Rajendran, S. 2016, Phys. Rev. D, 93, 103520

Grossman, S. A., & Narayan, R. 1989, ApJ, 344, 637

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, W. E., & Reina-Campos, M. 2023, MNRAS, 526, 2696

Harris, W. E., Morningstar, W., Gnedin, O. Y., et al. 2014, ApJ, 797, 128

Harris, W. E., Blakeslee, J. P., & Harris, G. L. H. 2017, ApJ, 836, 67

Hezaveh, Y. D., Dalal, N., Marrone, D. P., et al. 2016, ApJ, 823, 37

Humphreys, R. M. 1978, ApJS, 38, 309

Humphreys, R. M., & Davidson, K. 1979, ApJ, 232, 409

Jing, Y. P., & Suto, Y. 2002, ApJ, 574, 538

Jordán, A., McLaughlin, D. E., Côté, P., et al. 2007, ApJS, 171, 101

Kaurov, A. A., Dai, L., Venumadhav, T., Miralda-Escudé, J., & Frye, B. 2019, ApJ, 880, 58

Kelly, P. L., Diego, J. M., Rodney, S., et al. 2018, Nat. Astron., 2, 334

Kelly, P. L., Chen, W., Alfred, A., et al. 2022, ArXiv e-prints [arXiv:2211.02670]

Kilpatrick, C. D., Foley, R. J., Drout, M. R., et al. 2018, MNRAS, 473, 4805

Kolb, E. W., & Tkachev, I. I. 1993, Phys. Rev. Lett., 71, 3051

Laroche, A., Gilman, D., Li, X., Bovy, J., & Du, X. 2022, MNRAS, 517, 1867

Lazar, A., Bullock, J. S., Boylan-Kolchin, M., et al. 2021, MNRAS, 502, 6064

Lee, M. G., Bae, J. H., & Jang, I. S. 2022, ApJ, 940, L19

Li, S. K., Diego, J. M., Kelly, P. L., et al. 2024, ArXiv e-prints [arXiv:2404.08571]

Lynds, R., & Petrosian, V. 1989, ApJ, 336, 1

Mao, S., & Schneider, P. 1998, MNRAS, 295, 587

Massey, P., & Hunter, D. A. 1998, ApJ, 493, 180

Meena, A. K., Chen, W., Zitrin, A., et al. 2023a, MNRAS, 521, 5224

Meena, A. K., Zitrin, A., Jiménez-Teja, Y., et al. 2023b, ApJ, 944, L6

Metcalf, R. B., & Madau, P. 2001, ApJ, 563, 9

Meynet, G., Chomienne, V., Ekström, S., et al. 2015, A&A, 575, A60

Miralda-Escude, J. 1991, ApJ, 379, 94

Moliné, Á., Sánchez-Conde, M. A., Palomares-Ruiz, S., & Prada, F. 2017, MNRAS, 466, 4974

Montes, M., & Trujillo, I. 2022, ApJ, 940, L51

Narayan, R. 1989, ApJ, 339, L53

Narayan, R., & Wallington, S. 1993, Liege Int. Astrophys. Colloquia, 31, 217

Navarro, J. F., Frenk, C. S., & White, S. D. M. 1996, ApJ, 462, 563

Niemiec, A., Jauzac, M., Eckert, D., et al. 2023, MNRAS, 524, 2883

Oguri, M., Diego, J. M., Kaiser, N., Kelly, P. L., & Broadhurst, T. 2018, Phys. Rev. D, 97, 023518

Oh, S.-H., Hunter, D. A., Brinks, E., et al. 2015, AJ, 149, 180

Palencia, J. M., Diego, J. M., Kavanagh, B. J., & Martinez, J. 2023, A&A, 687, A81

Pastorello, A., Botticella, M. T., Trundle, C., et al. 2010, MNRAS, 408, 181

Patrício, V., Richard, J., Carton, D., et al. 2018, MNRAS, 477, 18

Peng, E. W., Ferguson, H. C., Goudfrooij, P., et al. 2011, ApJ, 730, 23

Powell, D. M., Vegetti, S., McKean, J. P., et al. 2023, MNRAS, 524, L84

Ragagnin, A., Meneghetti, M., Calura, F., et al. 2024, A&A, 687, A270

Reina-Campos, M., Keller, B. W., Kruijssen, J. M. D., et al. 2022, MNRAS, 517, 3144

Rodney, S. A., Balestra, I., Bradac, M., et al. 2018, Nat. Astron., 2, 324

Saifollahi, T., Voggel, K., Lançon, A., et al. 2024, A&A, submitted [arXiv:2405.13500]

Schive, H.-Y., Chiueh, T., Broadhurst, T., & Huang, K.-W. 2016, ApJ, 818, 89

Schneider, P., Ehlers, J., & Falco, E. E. 1992, Gravitational Lenses (Springer-Verlag Berlin Heidelberg New York)

Sheth, R. K., & Tormen, G. 1999, MNRAS, 308, 119

Smail, I., Ellis, R. S., Aragon-Salamanca, A., et al. 1993, MNRAS, 263, 628

Smail, I., Dressler, A., Kneib, J.-P., et al. 1996, ApJ, 469, 508

Soucail, G., Fort, B., Mellier, Y., & Picat, J. P. 1987, A&A, 172, L14

Soucail, G., Mellier, Y., Fort, B., Mathez, G., & Cailloux, M. 1988, A&A, 191, L19

Steinhardt, C. L., Jauzac, M., Acebron, A., et al. 2020, ApJS, 247, 64

Treu, T. 2010, ARA&A, 48, 87

Tsang, A., Çağan Şengül, A., & Dvorkin, C. 2024, ArXiv e-prints [arXiv:2401.16624]

Umetsu, K., Medezinski, E., Nonino, M., et al. 2014, ApJ, 795, 163

Valenzuela, L. M., Moster, B. P., Remus, R.-S., O'Leary, J. A., & Burkert, A. 2021, MNRAS, 505, 5815

Vall Müller, C., & Miralda-Escudé, J. 2024, ArXiv e-prints [arXiv:2403.16989]

Vega-Ferrero, J., Diego, J. M., & Bernstein, G. M. 2019, MNRAS, 486, 5414

Venumadhav, T., Dai, L., & Miralda-Escudé, J. 2017, ApJ, 850, 49

Villegas, D., Jordán, A., Peng, E. W., et al. 2010, ApJ, 717, 603

Visinelli, L., Baum, S., Redondo, J., Freese, K., & Wilczek, F. 2018, Phys. Lett. B, 777, 64

Weis, K., & Bomans, D. J. 2020, Galaxies, 8, 20

Welch, B., Coe, D., Diego, J. M., et al. 2022, Nature, 603, 815

Williams, L. L. R., Kelly, P. L., Treu, T., et al. 2024, ApJ, 961, 200

Windhorst, R. A., Timmes, F. X., Wyithe, J. S. B., et al. 2018, ApJS, 234, 41

Wolf, J., Martinez, G. D., Bullock, J. S., et al. 2010, MNRAS, 406, 1220

Yan, H., Ma, Z., Sun, B., et al. 2023, ApJS, 269, 43

Appendix A: Lens Model

Details of the algorithm are provided by Diego et al. (2005, 2007, 2016). This modeling technique has been applied successfully to several clusters observed with HST and JWST (Diego et al. 2005, 2007, 2016, 2023a, 2024b).

The model for A370 is derived using 32 lensed galaxies with spectroscopic redshifts, and producing over 90 multiple images, or constraints. The model is derived as part of the Beyond the Ultradeep Frontier Fields and Legacy Observations (BUFFALO) project (GO-15117, PIs Steinhardt & Jauzac; Steinhardt et al. 2020), and also incorporates information from weak lensing measured with HST images. Details of the dataset are given by Niemiec et al. (2023).

This model incorporates all member galaxies detected by HST near the Dragon arc, so it includes all relevant deflectors at galactic scales and above. The CC predicted by our lens model in the Dragon arc is shown in Figure 1. For this work we are interested in the area in the source plane with magnification $\mu > \mu_{\mathrm{crit}}$ from millilenses that are in regions of the lens plane where $\mu_{1m} < \mu_{\mathrm{crit}}$ (or far region). We are also interested in a similar area in the source plane but from regions in the lens plane near the cluster CC where the macromodel alone can provide $\mu_{1m} > \mu_{\mathrm{crit}}$ needed for the probability of microlensing to be high enough. More precisely, we are interested in the ratio of the two areas, since this ratio will essentially correlate with the ratio of events found near the CC and far from the CC. The use of a different lens model should have a relatively small impact on our conclusions provided these lens models include all member galaxies near the critical region, since these member galaxies can alter the position of the CCs. The ratio of events should then remain more or less constant for most models, with a relatively small dependence on the slope of the lensing potential. This difference in slopes can account for a factor of ~ 2 in the ratio of areas and hence on the ratio of events between different
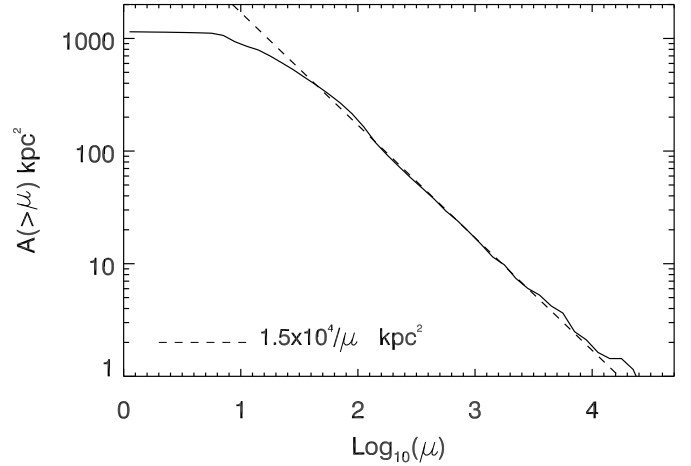

Fig. A.1. Area above a certain magnification in the Dragon arc. The solid line shows the area with magnification $> \mu$ computed in the image plane and in the region occupied by the Dragon arc. The dashed line is the simple power-law fit $A(> \mu) = 1.5 \times 10^4/\mu$ in kpc^2.

lens models. A level of uncertainty of a factor of ~ 2 should be kept in mind owing to uncertainty in the macro galaxy cluster model.

The area in the image plane above a magnification μ computed in the region of the Dragon arc is shown in Figure A.1. As expected, this area scales as the canonical $1/\mu$ scaling law. Above $\mu = 100$ there are $190\,\mathrm{kpc}^2$ in the image plane. Dividing by $\mu = 100$, this corresponds to $1.9\,\mathrm{kpc}^2$ in the source plane, and correcting for the multiplicity factor two (for every counter-image with magnification $\mu = 100$ in the image plane, there is another on the other side of the CC with similar magnification), we arrive at $0.95\,\mathrm{kpc}^2$, setting the upper boundary of the orange region in Figure 7.