

UC Berkeley

UC Berkeley Previously Published Works

Title

Statistical microlensing towards magnified high-redshift star clusters

Permalink

<https://escholarship.org/uc/item/0jq7h3z7>

Journal

Monthly Notices of the Royal Astronomical Society, 501(4)

ISSN

0035-8711

Author

Dai, Liang

Publication Date

2021-01-21

DOI

10.1093/mnras/stab017

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

Statistical Microlensing Toward Magnified High-Redshift Star Clusters

Liang Dai^{1,2}★

¹*Institute for Advanced Study, 1 Einstein Drive, Princeton, NJ 08540, USA*

²*University of California, Berkeley, Department of Physics, 366 LeConte Hall, Berkeley, CA 94720, USA.*

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT

We study light variability of gravitationally magnified high-redshift star clusters induced by a foreground population of microlenses. This arises as the incoherent superposition of light variations from many source stars traversing the random magnification pattern on the source plane. The light curve resembles a scale-invariant, Gaussian process on timescales of years to decades, while exhibits rapid and frequent micro-caustic crossing flares of larger amplitudes on timescales of days to months. For a concrete example, we study a young Lyman-continuum-leaking star cluster in the Sunburst Arc at $z = 2.37$. We show that one magnified image happens to be intervened by a foreground galaxy, and hence should exhibit a variable flux at the 1–2% level, which is measurable in space with ~ 1 –3 ks exposures on the Hubble Space Telescope and more easily with the James Webb Space Telescope, or even from the ground using a ~ 4 -meter telescope without adaptive optics. Detailed measurement of this variability can help determine the absolute macro magnification and hence the intrinsic mass and length scales of the star cluster, test synthetic stellar population models, and probe multiplicity of massive stars. Furthermore, monitoring the other lensed images of the star cluster, which are free from significant intervention by foreground microlenses, can allow us to probe planetary to stellar mass compact objects constituting as little as a few percent of the dark matter. Given the typical surface density of intracluster stars, we expect this phenomenon to be relevant for other extragalactic star clusters lensed by galaxy clusters.

Key words: gravitational lensing: micro – gravitational lensing: strong – dark matter – galaxies: clusters: individual: PSZ1 G311.65-18.48;

1 INTRODUCTION

Strong gravitational lensing offers extraordinary opportunities to probe star formation at high redshifts. Thanks to lensing magnification, distant and intrinsically small star-forming galaxies can be detectable behind galaxy or galaxy cluster lenses, and often in the resultant giant arcs candidates of massive young star clusters are seen (e.g. Vanzella et al. (2017); Rivera-Thorsen et al. (2019)). These impressive objects have recently aroused tremendous interest because they likely represent a major mode of star formation at Cosmic Noon in environments of low metallicity and high specific star-formation rate, and because their kind might be an efficient source of escaping ionizing radiation during the epoch of reionization (Ricotti 2002). It is therefore worth considering how by exploiting lensing more information about their astrophysical properties might be extracted beyond the resolution limit of current instruments.

When a population of compact objects intervene the line of sight, they act as microlenses and cast a random pattern of flux

magnification on the source plane (Gott 1981; Young 1981; Chang & Refsdal 1984; Wambsganss 2001). Under these circumstances, individual members of a stellar system independently flicker as they steadily traverse this pattern, and they collectively give rise to a variable flux integrated over the entire association which is often unresolved or marginally resolved. This concept of statistical microlensing was previously proposed for “pixel” microlensing toward crowded star fields in nearby galaxies (Crofts 1992; Crofts & Tomaney 1996; Alcock et al. 1999), and for surface brightness variability of distant galaxies (Lewis & Ibata 2001; Lewis et al. 2000; Tuntsov et al. 2004; Gil-Merino & Lewis 2006).

In this work, we consider the similar effect acting on highly magnified star clusters behind galaxy cluster lenses. We turn our attention to the existing populations of microlenses – stars either residing in intervening minor foreground galaxies or from the diffuse intracluster light (ICL) (Zwicky 1951; Lin & Mohr 2004; Zibetti et al. 2005). These populations normally have a small mean convergence $\kappa_\star \sim 0.01$, and hence, toward unlensed or moderately magnified ($|\mu_{\text{macro}}| \sim \text{a few}$) sources, would only cause insignificant microlensing under a low optical depth. Interestingly,

★ E-mail: liangdai@berkeley.edu

for sources under high magnifications $|\mu_{\text{macro}}| \sim \mathcal{O}(10)\text{--}\mathcal{O}(100)$, the strong external shear elevates microlensing into effectively the optical thick regime, with $\kappa_{\star} |\mu_{\text{macro}}| = \mathcal{O}(1)$ (Venumadhav et al. 2017; Diego et al. 2018; Oguri et al. 2018). Given the latest observational progress, it is timely to revisit statistical microlensing in this latter regime.

Recent years have seen important development in detecting individual superluminous stars in caustic-straddling arcs at $z \sim 1\text{--}1.5$ (Kelly et al. 2018; Chen et al. 2019; Kaurov et al. 2019), which are magnified by astounding factors in the hundreds and exhibited flux stochasticity due to microlensing. Arguably, observing statistical microlensing of a large association of bright stars can be complementary to, and perhaps in some aspects more advantageous, than observing similar effects on individually resolved stars. Foremost, highly magnified individual stars (Miralda-Escudé 1991) become exceedingly difficult to detect beyond $z \approx 2$ due to the Eddington limit on the intrinsic luminosity of massive stars (Dai et al. 2018), while young star clusters at Cosmic Noon redshifts with lensing-enhanced visual magnitudes $m \approx 22\text{--}25$ are much more easily detectable at high signal-to-noise ratios (SNRs).

Individual stars are most easily detectable under the large magnification boost during micro-caustic crossings (Diego 2019). However, the erratic nature of these rapid events render them rare to be caught at random observing epochs. Moreover, for any single star, micro-caustic crossings often occur rather infrequently, unless in the improbable case of an exceptionally large $|\mu_{\text{macro}}| \gtrsim 10^3$. Owing to the highly non-Gaussian nature of the light curve, an impractically long observing program would be necessary to accumulate a statistically meaningful amount of variability data. By contrast, these shortcomings can be significantly mitigated, if not completely overcome, by targeting a collective source consisting of a large number of stars. As we will show with examples, the collective variability resembles a scale-invariant, Gaussian process on the timescale of years, while on short timescales exhibits extremely frequent (e.g. a dozen of events per month) flares that arise from micro-caustic crossings of a large number of stars. The trade-off would be that a higher relative precision in photometry has to be achieved, which nevertheless might not be a problem — after all, the absolute flux fluctuation from an incoherent sum of, say $\sim N$ comparably bright stars, is $\sim \sqrt{N}$ times larger than that of any single star. We therefore propose, in conjunction with the search and monitoring of extremely magnified individual stars (e.g. Kelly et al. (2019)), that microlensing-induced variabilities of magnified extragalactic star clusters be sought and measured.

With the goal to develop general insight into this phenomenon, in this work we examine the specific example of a Cosmic Noon star cluster found in the gravitationally lensed Sunburst Arc at $z = 2.37$ behind a galaxy cluster lens (Dahle et al. 2016), which has been dubbed the “LyC Knot” because Lyman Continuum emission was detected from it (Rivera-Thorsen et al. 2019). This LyC Knot shows a total of 12 lensed macro images, many of which likely to have high magnifications in the range of tens (Vanzella et al. 2020b). Identified as a 3 Myr-old compact star cluster with a total mass $M_{\star} \approx 10^7 (50/|\mu_{\text{macro}}|)^{1/2} M_{\odot}$ (Chisholm et al. 2019), microlensing induced flux variability seen across a wide range of observed wavelengths must be dominated by $\sim 10^3\text{--}10^4$ O-type main-sequence stars and evolved B-type supergiants. Assuming the standard initial mass function (IMF) and neglecting stellar multiplicity, our stellar population modeling suggests that one of the lensed images of the LyC Knot, due to an intervening minor foreground galaxy G1 (which provides $\kappa_{\star} \approx 0.01$ as we estimate), should exhibit a variable flux at 0.8–2.5% level if it has a macro

magnification $|\mu_{\text{macro}}| = 20\text{--}50$. Given the remarkable apparent brightness of those lensed images, such percent-level variability should be measurable in the rest-frame UV bands for $\sim 1\text{--}3$ ks exposures with the HST (Rivera-Thorsen et al. 2019).

We will show that, as long as the magnified flux is held fixed by observation, the typical amplitude of flux variability depends only weakly on the IMF and metallicity, but is sensitive to the macro magnification μ_{macro} . If accurate modeling of the microlens surface density κ_{\star} is available, measuring the flux variability of a star cluster can determine its absolute mass scale by determining $|\mu_{\text{macro}}|$, which will be a valuable result since lens modeling may suffer from large uncertainty at the locations of high-magnification features (Priewe et al. 2017). Detailed statistical quantification of the flux variability under statistical microlensing and comparing data to theory will offer a new avenue to test stellar population synthesis in the context of high- z environments of high specific star-formation rate and low metallicity. We furthermore suggest that the fraction of massive stars in binary or multiple stellar systems (Sana et al. 2012, 2013), another crucial aspect of population synthesis particularly important for understanding ionizing fluxes (Stanway et al. 2016; Ma et al. 2016; Rosdahl et al. 2018), may also be probed with statistical microlensing, as a larger fraction will enhance the variability amplitude.

Many of the earlier works on statistical microlensing were originally motivated by the prospect to probe stellar or planetary mass compact objects which were thought to constitute the cosmological Dark Matter (DM). Since then, an $\mathcal{O}(1)$ mass fraction of these compact DM has been ruled out by microlensing surveys (Alcock et al. 2001; Griest et al. 2013; Tisserand et al. 2007; Niikura et al. 2019) and various other astrophysical (Brandt 2016) and cosmological tests (Mediavilla et al. 2017; Ali-Haïmoud & Kamionkowski 2017; Zumalacárregui & Seljak 2018). In spite of that, we propose that empirical constraints on percent-level variability seen in magnified star clusters (e.g. with the other lensed images of the LyC Knot in the Sunburst Arc) could tighten the limit on the mass fraction to $\mathcal{O}(10^{-2})$ across the mass range $\sim 10^{-6}\text{--}10^2 M_{\odot}$.

The remainder of this paper is organized as follows. In Sec. 2, we justify several simplifying assumptions underlying our statistical description of the collective microlensing phenomenon, and introduce the formalism to quantify the amplitude variance and temporal correlation of flux variability. Then in Sec. 3, we study the specific case of the LyC Knot of the Sunburst Arc. In Sec. 4, we discuss related issues regarding variable stars, stellar multiplicity, and compact dark matter. Finally, we summarize our conclusions in Sec. 5. Additional technical details are presented in Appendices: we discuss in App. A a simple mass profile model for the cluster lens in front of the Sunburst Arc, and estimate the abundance of intracluster stars; we characterize in App. B the stellar population of the foreground galaxy G1, from which we derive the abundance of microlenses toward one magnified image of the LyC Knot.

2 THEORETICAL FORMALISM

The overall flux of a star cluster can have temporal variability if individual member stars have variable fluxes due to microlensing (Venumadhav et al. 2017; Diego et al. 2018; Oguri et al. 2018). When the star cluster is unresolved, what is measurable is the integrated flux but not the fluxes from individual members (Crofts 1992; Lewis et al. 2000; Lewis & Ibata 2001). Recently, collective variability was explored in Dai et al. (2020) under the idealistic assumption of identical member stars, in the context of understanding highly magnified,

asymmetric image pairs of compact star clusters residing in the caustic straddling lensed galaxy SGAS J122651.3+215220 (Dahle et al. 2016).

In this work, we expand that study by accounting for a realistic source flux distribution based on synthetic stellar populations. A general theoretical framework to quantify statistical microlensing was presented in Tuntsov et al. (2004). In the following, we follow closely the reasoning of that reference and reproduce many of results there, albeit we will employ the method of characteristic function to understand Gaussianization of light variability.

We consider, in the projected vicinity of the line of sight toward the star cluster, randomly located microlenses, which are superimposed on top of some locally uniform coarse-grained convergence κ_0 and shear γ_0 from the smooth macro lens. The microlenses on average make a small but important contribution $\kappa_\star = \Sigma_\star/\Sigma_{\text{crit}} \sim 10^{-2}$ to κ_0 , where Σ_\star is their surface mass density, and $\Sigma_{\text{crit}} = (c^2/4\pi G)(D_S/D_L D_{LS})$ is the critical surface density (Blandford & Narayan 1986), with D_L , D_S and D_{LS} being the angular diameter distances to the lens plane, to the source plane, and from the lens plane to the source plane, respectively. The star cluster is subject to a (signed) macro magnification $\mu_{\text{macro}} = \mu_t \mu_r$, where

$$\mu_t = (1 - \kappa_0 - \gamma_0)^{-1}, \quad \text{and} \quad \mu_r = (1 - \kappa_0 + \gamma_0)^{-1} \quad (1)$$

are the one-dimensional magnifications along and perpendicular to the macro elongation, respectively. We will focus on the regime of large macro magnification near a tangential caustic, i.e. $|\mu_t| \gg 1$ and $\mu_r \sim \mathcal{O}(1)$.

2.1 Basic Assumptions

To begin with, we justify several simplifying assumptions we will make in this paper.

Firstly, we assume that individual member stars have uncorrelated variabilities. For stellar microlenses of masses $\sim \mathcal{O}(1 M_\odot)$, this should be a good approximation because the characteristic length scale of the magnification pattern cast onto the source plane, on the order of the Einstein scale $\sim \mathcal{O}(10^3 \text{ AU})$ (or more precisely a factor of $\sqrt{|\mu_t|}$ larger perpendicular to macro elongation (Oguri et al. 2018)), is smaller than the typical projected separations between the brightest member stars, say $\sim 10^4$ – 10^5 OB stars packed within a radius of a few pc. This assumption becomes invalid for tight binary or multiple stars, which are in fact not uncommon for massive stars (Sana & Evans 2011; Sana 2017). For now, we neglect stellar multiplicity, and will comment on this later.

Secondly, the source-plane magnification pattern is assumed to be statistically homogeneous. For microlenses in either the intracluster space or from an old stellar population in a foreground galaxy, their projected positions typically have randomized sufficiently. Also, the extent of the entire star cluster (a few pc across), when mapped onto image plane, probes typically $\sim 10^3$ – 10^4 microlenses simultaneously, adequately sampling the underlying statistics.

For a third premise, we assume that all member stars traverse the magnification pattern at the same transverse velocity, which we define to be at an angle ϕ with respect to the direction of macro elongation. This uniform velocity is essentially set by the transverse motion of the star cluster as a whole relative to the lens. After all, internal motions of the star cluster members $\sim \mathcal{O}(10 \text{ km/s})$ are much slower than the typical peculiar motion of the large scale structure $\sim \mathcal{O}(10^2$ – $10^3 \text{ km/s})$ (Miralda-Escudé 1991; Venumadhav et al. 2017). The velocity dispersion of the microlenses are also

negligible, as it is suppressed by a factor of $|\mu_t|$ when mapped onto the source plane.

Since in this context source stars are effectively point sources for magnifications $\lesssim 10^4$ (Venumadhav et al. 2017), we assume that all member stars, across a wide range of wavelengths, have the same statistical distributions for the fractional flux variability.

For one more simplifying assumption, we will neglect any intrinsic variability of the source stars including stellar pulsations, outbursts, or other explosive transients (see Sec. 4.1 for discussion). We will also neglect the effects of eclipses in multiple-star systems. With these assumptions, in the following we will analytically express the statistics of the collective flux variability in terms of that for individual member stars, following the standard results in statistics.

2.2 Individual Stars

Under microlensing, the flux F of a given individual star fluctuates around the mean value \bar{F} ,

$$F = \bar{F} (1 + \delta). \quad (2)$$

Consider the probability distribution function (PDF) for the fractional fluctuation δ , $P(\delta)$, which satisfies $\int d\delta P(\delta) = 1$ and $\langle \delta \rangle = \int d\delta P(\delta) \delta = 0$. The characteristic function (CF) is given by the Fourier transform of $P(\delta)$,

$$\Phi(\omega) \equiv \langle e^{i\omega\delta} \rangle = \int d\delta P(\delta) e^{i\omega\delta}. \quad (3)$$

The logarithm of the CF, $\Psi(\omega) = \ln \Phi(\omega)$, gives the cumulant generating function (CGF).

For an idealized point source, $P(\delta)$ peaks at $\delta \sim \mathcal{O}(1)$, but has a power-law tail $P(\delta) \propto \delta^{-3}$ for $\delta \gg 1$ (Peacock 1982; Schneider 1987) because fold singularities dominate the PDF in the high magnification regime (Blandford & Narayan 1986; Schneider & Weiss 1988). Formally, this results in a logarithmic divergence in the second-order cumulant $\langle \delta^2 \rangle_c = \langle \delta^2 \rangle - \langle \delta \rangle^2$, and power-law divergences in higher order cumulants. For a realistic stellar source, the finite source size truncates the high magnification tail and hence regularizes these divergences, but the cumulants may still have large numerical values, reflecting the strongly non-Gaussian nature of $P(\delta)$. When all cumulants $\langle \delta^n \rangle_c$, $n = 2, 3, \dots$ are finite, the CGF is analytic around $\omega = 0$, and the values of the cumulants can be read from the Taylor expansion coefficients:

$$\Psi(\omega) = \sum_{n=2}^{\infty} \frac{i^n}{n!} \langle \delta^n \rangle_c \omega^n. \quad (4)$$

In particular, $\langle \delta^2 \rangle_c$ is equivalent to the quantity ε_μ^2 defined by Tuntsov et al. (2004).

2.3 Integrated Flux of a Cluster: One-point statistics

We now consider the many members of a star cluster. The members can be partitioned into bins according to the mean flux \bar{F} , labeled by $I = 1, 2, \dots$. The I -th bin consists of N_I (nearly) identical stars, each of which has a mean flux \bar{F}_I . The flux integrated over the entire cluster therefore has a mean value $\bar{F}_{\text{cl}} = \sum_I N_I \bar{F}_I$. The integrated flux F_{cl} at a given epoch fluctuates around this value by a fractional amount Δ , defined through

$$F_{\text{cl}} = \bar{F}_{\text{cl}} (1 + \Delta). \quad (5)$$

Under our assumptions, Δ has a vanishing mean $\langle \Delta \rangle = 0$, and its own CGF is given by

$$\Psi_{\text{cl}}(\omega) = \ln \Phi_{\text{cl}}(\omega) = \ln \langle e^{i\omega\Delta} \rangle = \sum_I N_I \Psi \left(\frac{\bar{F}_I}{\bar{F}_{\text{cl}}} \omega \right). \quad (6)$$

Taking the limit of a continuous distribution of \bar{F} for individual stars, we introduce the flux distribution function $dN/d\bar{F}$, and derive

$$\bar{F}_{\text{cl}} = \int d\bar{F} \frac{dN}{d\bar{F}} \bar{F}, \quad (7)$$

and

$$\Psi_{\text{cl}}(\omega) = \int d\bar{F} \frac{dN}{d\bar{F}} \Psi \left(\frac{\bar{F}}{\bar{F}_{\text{cl}}} \omega \right) \quad (8)$$

Inserting Eq. (4) into Eq. (8), we have

$$\Psi_{\text{cl}}(\omega) = \sum_{n=1}^{\infty} \frac{i^n}{n!} \left(\frac{1}{\bar{F}_{\text{cl}}^n} \int d\bar{F} \frac{dN}{d\bar{F}} \bar{F}^n \right) \langle \delta^n \rangle_c \omega^n. \quad (9)$$

Therefore, Δ has cumulants

$$\langle \Delta^n \rangle_c = \frac{1}{N^{n-1}} \frac{\langle \bar{F}^n \rangle}{\langle \bar{F} \rangle^n} \langle \delta^n \rangle_c, \quad (10)$$

where we define the population averaged moments of the mean flux \bar{F} ,

$$\langle \bar{F}^n \rangle \equiv \frac{1}{N} \left(\int d\bar{F} \frac{dN}{d\bar{F}} \bar{F}^n \right), \quad (11)$$

and $N \equiv \int d\bar{F} (dN/d\bar{F})$ is the total number of stars. From Eq. (10), we find

$$\frac{[\langle \Delta^n \rangle_c]^{1/n}}{[\langle \Delta^2 \rangle_c]^{1/2}} = N^{\frac{1}{n}-\frac{1}{2}} \frac{\langle \bar{F}^n \rangle^{1/n}}{\langle \bar{F}^2 \rangle^{1/2}}. \quad (12)$$

If the distribution $dN/d\bar{F}$ is fixed but N increases, the higher order cumulants $\langle \Delta^n \rangle_c$ for $n = 3, 4, \dots$ are all suppressed relative to the second order cumulant $\langle \Delta^2 \rangle_c$. In the limit that all cumulants $\langle \Delta^n \rangle_c$ become negligibly small except for $n = 2$, $P(\Delta)$ approaches a normal distribution with a variance

$$\langle \Delta^2 \rangle_c = \epsilon_2^2 \langle \delta^2 \rangle_c, \quad (13)$$

where we introduce the dimensionless parameter independent of microlensing:

$$\epsilon_2 \equiv \left(\frac{1}{N} \frac{\langle \bar{F}^2 \rangle}{\langle \bar{F} \rangle^2} \right)^{1/2}. \quad (14)$$

Eq. (13) states that the fractional fluctuation in the cluster integrated flux F_{cl} has a root mean square (RMS) that is a factor ϵ_2 smaller than that of the individual stars. For a stellar population of fixed intensive properties, ϵ_2 scales with the total stellar mass M_\star as $\epsilon_2 \propto M_\star^{-1/2}$. This is the intuitive result that the highly non-Gaussian flux variability of a large number of cluster member stars collectively give rise to variability in the integrated flux that is more Gaussianized and has a reduced amplitude.

At low microlensing optical depths $\kappa_\star |\mu_{\text{macro}}| \ll 1$, the high magnification tail of $P(\delta)$ only makes a subdominant contribution to $\langle \delta^2 \rangle_c$, so the RMS of Δ is primarily determined by the width of the core around $\delta \approx 0$.

At high microlensing optical depths $\kappa_\star |\mu_{\text{macro}}| \gtrsim 1$, a situation we will demonstrate later, it is the heavy tail of the PDF $P(\delta)$ at $\delta \gg 1$ rather than the core at $\delta \approx 0$ that dominates $\langle \delta^2 \rangle_c$. In this situation, the RMS of Δ primarily depends on micro-caustic crossings of individual stars. By the same logic, any residual non-Gaussian behavior in Δ is sensitive to the high- δ cutoff in $P(\delta)$.

2.4 Integrated Flux of a Cluster: Two-point Statistics

The one-point PDF $P(\delta)$ or $P(\Delta)$ does not capture all statistical information. Additional information is encoded in multi-point statistics of δ , namely correlation between δ 's measured at multiple source-plane positions. These are directly translated into correlation between δ 's measured at different epochs under a uniform source motion, and hence provide the timescale information of the variability (Wyithe & Turner 2002). Previously, Lewis & Irwin (1996) and Neindorf (2003) studied the temporal correlation in the light curve of microlensed quasars. Following the same logic, we can study the temporal correlation of the integrated flux F_{cl} .

The simplest example is the two-point correlation function,

$$\xi(\vartheta) \equiv \langle \delta(\theta) \delta(\theta + \vartheta) \rangle, \quad (15)$$

also referred to as the structure function by Wyithe & Turner (2001) (and a related definition by Lewis & Irwin (1996)). By statistical homogeneity of the source-plane magnification pattern, it is only dependent on the source-plane separation ϑ along the source trajectory, or equivalently on the time separation under uniform source motion. Define the Fourier transform,

$$\tilde{\delta}(f) \equiv \int d\theta e^{i2\pi f \theta} \delta(\theta), \quad (16)$$

where f is the spatial frequency conjugate to θ . The Fourier transform has a power spectrum

$$\langle \tilde{\delta}(f) \tilde{\delta}^*(f') \rangle = P_\delta(f) \delta_D(f - f'), \quad (17)$$

where $P_\delta(f) = \int d\vartheta e^{i2\pi f \vartheta} \xi(\vartheta)$.

In a similar way, we can define the power spectrum $P_\Delta(f)$ for the integrated flux variability. Under the assumptions of independent variability and uniform source motions among individual stars, we have the analogy of Eq. (13) for the power spectra

$$P_\Delta(f) = \epsilon_2^2 P_\delta(f). \quad (18)$$

Without explicitly writing them down, we point out that the disconnected part of the higher-order multi-point statistics also satisfy various relations that resemble Eq. (10).

According to Eq. (18), $P_\Delta(f)$ inherits exactly the same functional shape as $P_\delta(f)$, despite the fact that flux variabilities of individual stars are highly non-Gaussian while that of the integrated flux Gaussianizes substantially. This feature implies, as long as our basic assumptions remain valid, that by measuring the shape of $P_\Delta(f)$ (and other disconnected multi-point correlation statistics) one can extract statistical information about the variability of individual stars, which reflect the population properties of the microlenses.

3 CASE STUDY: LYC KNOT IN SUNBURST ARC

We now apply our theoretical framework to an observed star cluster that we judge is subject to statistical microlensing. It is a compact young star cluster at $z = 2.37$, whose host galaxy is gravitationally distorted into a giant arc, dubbed the Sunburst Arc (Rivera-Thorsen et al. 2017), by a foreground galaxy cluster PSZ1 G311.65-18.48 (hereafter PSZ1-G311 for short) (Planck Collaboration et al. 2014). The star cluster is found to show an impressive set of 12 magnified images along the arc. Remarkably, redshifted LyC radiation was directly detected from each of these avatars (Rivera-Thorsen et al. 2019). Referred to as the LyC Knot, this star cluster is intriguing in its own right as it may give clues about the astrophysical properties of ionizing sources at high redshifts.

At the projected radii along the Sunburst Arc, we set a fiducial

value $\kappa_0 = 0.6$ based on a spherical Navarro-Frenk-White (NFW) model (Navarro et al. 1996, 1997) for the cluster lens profile. See App. A for details. Given the values of κ_0 and μ_{macro} , we separately determine μ_t and μ_r from Eq. (1).

Stellar population modeling indicates that, out to the rather large projected distance of the Sunburst Arc from the brightest cluster galaxy (BCG), ~ 170 kpc, the ICL contributes a small convergence from stellar microlenses $\kappa_{\star} \lesssim 0.002$ (see App. A), which as we can tell from later analysis is unimportant. Therefore, microlensing by intracluster stars is unlikely to play a significant role in the case of the Sunburst Arc, unlike other studied giant arcs with $\kappa_{\star} = O(10^{-2})$ in different massive lensing galaxy clusters, e.g. MACS J1149.5+2223 (Oguri et al. 2018), MACS J0416.1-2403 (Kaurov et al. 2019), and SDSS J1226+2152 (Dai et al. 2020).

Interestingly, a faint foreground galaxy, which we refer to as G1, happens to intervene the line of sight toward one of the 12 lensed images of the LyC Knot — Image 5 following the denotation of Rivera-Thorsen et al. (2019) (see Fig. B1). The expected surface number density of stellar microlenses in G1, $\kappa_{\star} \sim O(10^{-2})$, together with the (most likely) high macro magnification factor of Image 5, $|\mu_{\text{macro}}| \sim O(10)$, makes this a realistic case for statistical microlensing. Our goal will be to quantify the resultant flux variability, in terms of the variance as in Eq. (13), and the power spectrum as in Eq. (18).

3.1 Stellar Source Population of the LyC Knot

To evaluate one ingredient in Eq. (13), the parameter ϵ_2 , we have to model the population of source stars within the LyC Knot. Detailed rest-UV spectroscopy carried out by Chisholm et al. (2019) indicates that the LyC Knot is a Cosmic Noon cousin of the young star cluster R136 situated at the center of the star forming region 30 Doradus in the Large Magellanic Cloud (LMC) (Crowther et al. 2016). It is however one or two orders of magnitude more massive with a total stellar mass $M_{\star} \sim 10^6\text{--}10^7 M_{\odot}$, while still has a compact spatial extent $\lesssim 20\text{--}30$ pc (Vanzella et al. 2020b). Specifically, Chisholm et al. (2019) reported an age $t_{\text{SSP}} = 3.0\text{--}3.6$ Myr, metallicity $Z = 0.55\text{--}0.66 Z_{\odot}$, and inferred source-frame dust extinction $E(B-V) \simeq 0.15$.

At such a young age, the dominant contributors to the rest-frame UV/optical fluxes are massive O-type main-sequence stars and evolved B-type supergiants. The lower number of these massive stars compared to that of the low-mass stars has the favorable implication that the parameter ϵ_2 may not be tremendously suppressed to an observationally inaccessible level.

To evaluate ϵ_2 , we construct synthetic simple stellar populations (SSPs) at a uniform age t_{SSP} using the public code Flexible Stellar Population Synthesis (FSPS) (Conroy et al. 2009; Conroy & Gunn 2010). With FSPS, we adopt a double power-law model for the IMF including initial masses between $0.08 M_{\odot}$ and $100 M_{\odot}$. For initial masses between $0.08 M_{\odot}$ and $0.5 M_{\odot}$, we fix the slope to be $\alpha = 1.3$. For initial masses heavier than $0.5 M_{\odot}$, we consider different scenarios: (1) a standard Salpeter-like slope $\alpha = 2.3$ (Kroupa 2001a); (2) a top-heavy slope $\alpha = 1.6$, which has been suggested for intense star forming regions under extreme conditions (Jefábková et al. 2017). Besides, the parameter ϵ_2 is also dependent on metallicity. We also note that binaries or multiple stellar systems are not included in these models.

In Fig. 1, we show ϵ_2 as a function of the stellar age t_{SSP} . Since ϵ_2 is filter dependent, we evaluate ϵ_2 for several HST and JWST wide filters that span the observed wavelength range from UV to optical and near-IR, redshifted from $z_s = 2.37$. For the standard high-mass

IMF slope $\alpha = 2.3$, and $Z = 0.6 Z_{\odot}$ inferred for the LyC Knot, we find $\epsilon_2 \approx (1.5\text{--}3)\% (M_{\star}/10^6 M_{\odot})^{-1/2}$ for stellar populations younger than ~ 3 Myr and in a variety of (observer-frame) filters from UV to IR (left panel of Fig. 1). The value of ϵ_2 begins to evolve significantly when the star cluster age further. In the bluer filters F275W and F814W, ϵ_2 slowly transitions to a monotonic decrease for $t_{\text{SSP}} \gtrsim 6$ Myr, while in the redder filters F160W and F444W, ϵ_2 first reaches a peak of a factor of a few to ten larger at $t_{\text{SSP}} \approx 3\text{--}6$ Myr, and then steadily decreases for $t_{\text{SSP}} \gtrsim 10$ Myr.

Across a wide range of wavelengths, very young stellar populations with $t_{\text{SSP}} \lesssim 3$ Myr exhibit a stable value of ϵ_2 dominated by massive O stars during their main sequence stage. As the stellar population grow older, many massive stars become evolved ones and are subject to dramatic changes in their stellar radii and temperatures. Consequently, the flux distribution $dN/d\bar{F}$ significantly changes and so does the value of ϵ_2 .

The middle panel of Fig. 1 shows that ϵ_2 only mildly depends on the IMF slope for massive stars for a wide range of ages $t_{\text{SSP}} < 100$ Myr, as long as the integrated mean flux \bar{F}_{cl} of the star cluster is kept fixed (by observation). Between the standard scenario for the IMF slope and the more extreme top-heavy scenario, ϵ_2 varies by less than 20% at fixed \bar{F}_{cl} . This is because the flux, across a wide range of wavelengths, is dominated by the same brightest stars, which correspond to the massive end of the IMF. Varying the IMF slope at fixed \bar{F}_{cl} changes the relative abundance between high-mass and low-mass stars, but hardly changes the absolute number of the former. Placing a significant constraint on the IMF slope is possible but will necessarily require a fairly precise determination of ϵ_2 .

Similarly, the right panel of Fig. 1 shows that for young populations $t_{\text{SSP}} \lesssim 6$ Myr, ϵ_2 is mildly sensitive to the metallicity within the range $0.1 < Z/Z_{\odot} < 1$, except in the case of very low metallicity $Z = 0.1 Z_{\odot}$ and through in bluest filter F275W sensitive to ionizing radiation, for which the stellar model predicts that ϵ_2 in this filter can increase by a factor of two for $t_{\text{SSP}} \gtrsim 3$ Myr. We also note that ϵ_2 should be independent of dust reddening if all member stars are subject to homogeneous extinction.

Using the mass-magnification relation plotted in Figure 6 of Vanzella et al. (2020b), we predict for the LyC Knot with $t_{\text{SSP}} = 3$ Myr and $Z = 0.6 Z_{\odot}$ that $\epsilon_2 = (\mu_{\text{macro}}/20)^{1/2} (0.5\text{--}0.6)\%$, insensitive to the high-mass IMF slope and essentially in all filters we consider here.

The source stellar population can be reconstructed based on photometric and spectroscopic data. On the other hand, the value of ϵ_2 could be inferred (using Eq. (13)) from combining a direct measurement of $\langle \Delta^2 \rangle_c$ and a theoretical prediction of $\langle \delta^2 \rangle_c$ through modeling the macro lens and the intervening microlens population. We could then pin down the absolute macro magnification μ_{macro} needed to have the right value of ϵ_2 for the right source stellar population, which would allow us to determine the intrinsic mass and size of the star cluster.

Since multiple macro lensed images of the LyC Knot have the same ϵ_2 , one could directly measure $\langle \Delta^2 \rangle_c$'s between the various images, and translate that into the relative sizes of $\langle \delta^2 \rangle_c$. On the other hand, the relative μ_{macro} 's can be directly measured from photometry, and the signs can be inferred from the observed topology of the lensed image configuration. (Note that in the case of lensed images along a highly magnified, tangential arc with $|\mu_t| \gg |\mu_r|$, μ_r can be more easily extracted from a macro lens model without large uncertainties.) Since $\langle \delta^2 \rangle_c$ has nonlinear dependence on μ_{macro} , and the microlens population typically differs from one line of sight to another, we might be able to determine the correct absolute scale for the μ_{macro} 's without any accurate knowledge of ϵ_2 .

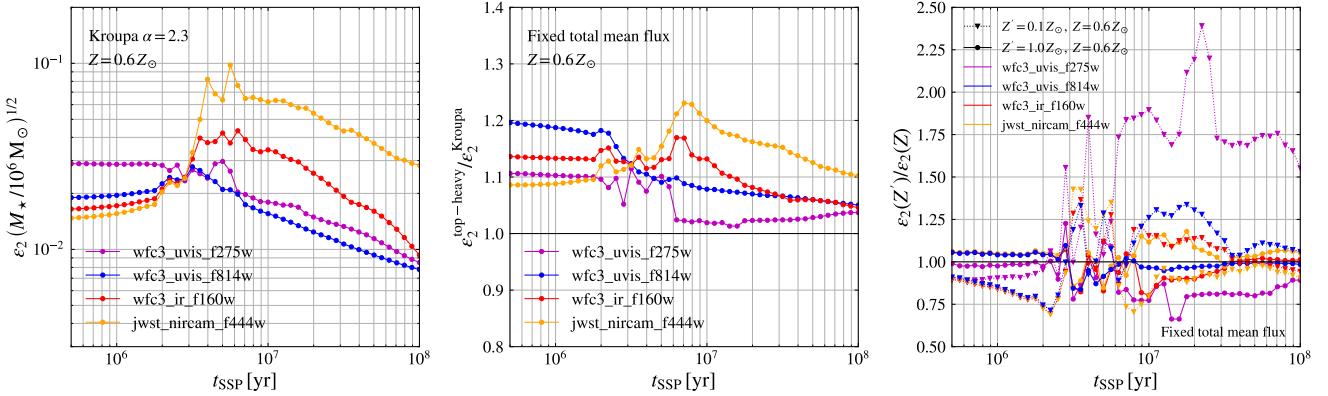


Figure 1. Theoretical values for the parameter ϵ_2 computed using the stellar population synthesis package FSPS (Conroy et al. 2009; Conroy & Gunn 2010), accounting only for single stellar evolution. We consider the evolution of an SSP with age t_{SSP} . **Left panel:** The parameter ϵ_2 normalized to a total stellar mass $M_\star = 10^6 M_\odot$ for the double power-law Kroupa IMF ($\alpha = 2.3$ for $M > 0.5 M_\odot$) as a function of t_{SSP} . **Middle panel:** Ratio in ϵ_2 between the Kroupa IMF and a top-heavy IMF ($\alpha = 1.6$ for $M > 0.5 M_\odot$) as a function of t_{SSP} . At fixed integrated mean flux \bar{F}_{cl} , ϵ_2 only depends weakly on the IMF slope for massive stars. **Right panel:** Ratio in ϵ_2 between different metallicity choices as a function of t_{SSP} . We compare $Z = 0.1 Z_\odot$ and $1 Z_\odot$ to the fiducial metallicity $Z = 0.6 Z_\odot$ found for the LyC Knot (Chisholm et al. 2019). At fixed integrated mean flux \bar{F}_{cl} and for stellar populations younger than ~ 6 Myr, ϵ_2 only depends mildly on metallicity. In all panels, we show curves for three HST wide filters, F275W, F814W and F160W, as well as one JWST NIRCAM wide filter F444W, respectively, all computed for the source redshift $z_s = 2.37$ of the Sunburst Arc. These filters are chosen as examples to cover UV, optical, and IR wavelength ranges in the observer frame. We note that an age $t_{\text{SSP}} = 3.0\text{--}3.6$ Myr was reported for the LyC Knot with or without including binary stellar evolution (Chisholm et al. 2019).

3.2 Microlens Population of G1

Having studied the intrinsic flux distribution of source stars in the LyC Knot, we now quantify flux (de-)magnification acting on individual stars, which is another ingredient in modeling the overall variability of the star cluster. We first need to characterize the stellar population of the microlens host galaxy G1.

As presented in App. B, we perform SED fitting of the rest-frame UV continuum using the archival MUSE IFU data (Program 297.A-5012; PI: N. Aghanim). Our SED fitting procedures disentangle the light of G1 from the light of the lensed Images of the LyC Knot, and reveal that G1 has a spectroscopic redshift $z = 0.458$, slightly higher than that of PSZ1-G311. As far as lensing of sources at $z_s = 2.37$ is concerned, however, it is an excellent approximation to simply set the redshift of G1 to be the same as that of the galaxy cluster.

Utilizing synthetic template SEDs from the BPASS model (Eldridge et al. 2017; Stanway & Eldridge 2018), we determine that the stellar population of G1 has an age $t_{\text{SSP}} = 10^{8.7\text{--}8.8}$ yr, a metallicity in the range $0.4 \lesssim Z/Z_\odot \lesssim 0.7$, and probably some small amount of internal dust extinction $E_{G1}(B - V) \approx 0.05$. Based on these parameters, we determine that the average surface mass density of microlenses amounts to $\kappa_\star = 0.007\text{--}0.02$, accounting for the fact that the line of sight toward Image 5 intersects the outskirts of G1.

In a preliminary lens model applicable to the region of Arc 1 as presented in Rivera-Thorsen et al. (2019), Image 5 is on the interior of the lensing critical curve, and hence its macro magnification factor is negative $\mu_{\text{macro}} < 0$. The absolute value $|\mu_{\text{macro}}|$ is not precisely known. A conservatively estimated range would be $10 < |\mu_{\text{macro}}| < 100$. Ongoing efforts to construct a detailed lens model for the Sunburst Arc (Vanzella et al. 2020a) may provide a more reliable number. The bottom line is that the expected large value of $|\mu_{\text{macro}}|$, together with the estimate $\kappa_\star = 0.007\text{--}0.02$ from G1, indicates that $\kappa_\star |\mu_{\text{macro}}| \sim \mathcal{O}(0.1\text{--}1)$. Therefore, Image 5 could be subject to optically thick statistical microlensing due to the intervening G1. Moreover, the negative sign of μ_{macro} favor-

ably enhances the microlensing effects on the flux compared to the $\mu_{\text{macro}} > 0$ case (Schechter & Wambsganss 2002).

3.3 Magnification PDF

The PDF $P(\delta)$ for the flux fluctuation of any single source star needs to be calculated for given microlens mass function and surface density κ_\star , and at given macro magnification parameters μ_t and μ_r . Previously, much efforts have been made to analytically model it in the study of microlensing induced variability of lensed quasars (Katz et al. 1986; Deguchi & Watson 1987; Wyithe & Turner 2001). A number of convenient approximations have been found to work successfully in the low optical depth regime (Peacock 1982; Vietri & Ostriker 1983; Marchandon & Nottale 1991; Kofman et al. 1997; Lee et al. 1997). For general values of κ_\star , however, accurate closed-form expressions for $P(\delta)$ that are fully satisfactory are not known. The major difficulty has to do with the nonlinear coupling between multiple microlenses at sufficiently high values of κ_\star and the formation of many micro-images (Fleury & García-Bellido 2020). With an interest in quasar microlensing, some past studies presented numerical results for moderate values of μ_{macro} (e.g. Lewis & Irwin (1995)).

A task more analytically tractable than evaluating the full PDF $P(\delta)$ is to compute only the second cumulant $\langle \delta^2 \rangle_c$, or more generally the two-point correlation function at nonzero lags. In the literature, semi-analytic methods have been developed to avoid direct numerical simulation of random microlensing realizations (Deguchi & Watson 1987; Seitz & Schneider 1994; Seitz et al. 1994; Neindorf 2003; Tuntsov et al. 2004). However, the presented numerical results in those references were for either $\kappa_\star \sim \mathcal{O}(1)$ or $|1 - \kappa_\star| \approx 0$, but are not easily quoted for the case under study in this work, i.e. a small mean convergence for the microlens $\kappa_\star = \mathcal{O}(10^{-2})$ coupled to a large macro magnification $|\mu_{\text{macro}}|$.

We therefore resort to numerical methods for evaluating $P(\delta)$. One class of methods generate randomized magnification patterns on the source plane through inverse ray-shooting (Wambs-

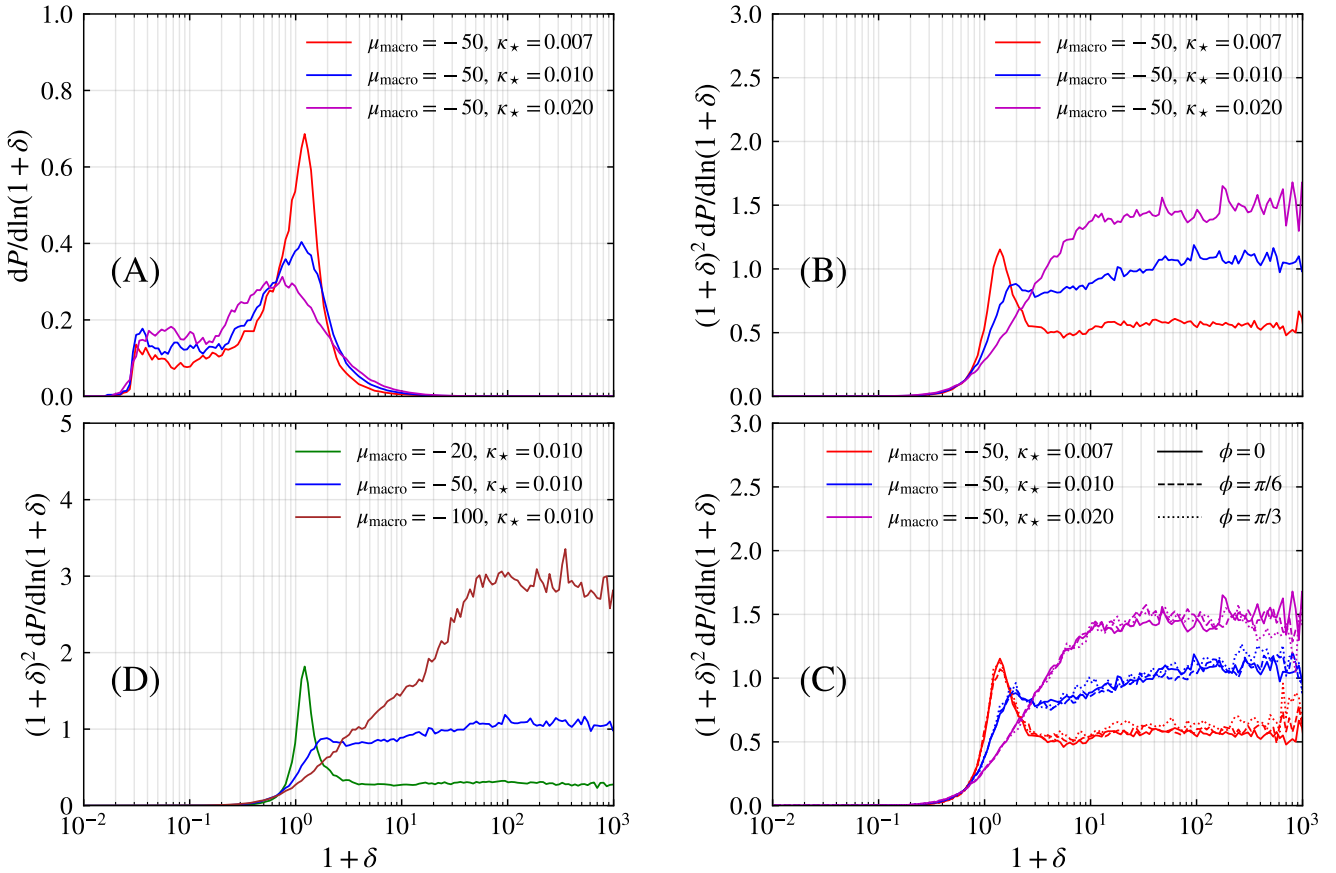


Figure 2. Probability distributions of the fractional variable flux $(1 + \delta)$ induced by stochastic microlensing, for any individual source star residing in the LyC Knot. The distributions are derived, from randomly generated microlensing realizations, for Image 5 of the LyC Knot behind the foreground galaxy G1, for $\kappa_0 = 0.6$ and different values of μ_{macro} . The microlens mass function follows from an SSP of an age $t_{\text{SSP}} = 630$ Myr and metallicity $Z = 0.4 Z_{\odot}$ with the standard IMF ($\alpha = 2.3$), as inferred for G1 (App. B). Except otherwise noted in Panel (C), the source trajectory is parallel to the macro shear $\phi = 0$. **Panel (A):** Probability density function (PDF) $dP/d \ln(1 + \delta)$ for $\mu_{\text{macro}} = -50$ and several choices for the mean convergence of microlenses $\kappa_{\star} = 0.007, 0.01, 0.02$. While having the most support around $\delta = 0$, the PDF has a power-law tail $dP/d \ln(1 + \delta) \propto (1 + \delta)^{-2}$ for $\delta \gg 1$, and is only truncated by the effect of finite source size (typically at $1 + \delta \approx 10^4 (R_s/10 R_{\odot})^{1/2}$ where R_s is the stellar radius (Venumadhav et al. 2017)). **Panel (B):** Same as (A), but with curves multiplied by $(1 + \delta)^2$ to highlight the contribution to the second cumulant $\langle \delta^2 \rangle_c$. The power-law tail at $\delta \gg 1$ can make a dominant contribution to $\langle \delta^2 \rangle_c$ if $\kappa_{\star} \geq 1/|\mu_{\text{macro}}|$, which grows logarithmically with the (physical) cutoff in δ . **Panel (C):** Same as (B), but with additional curves overplotted showing that the PDF is insensitive to the angle ϕ between the source trajectory and the direction of the macro shear. **Panel (D):** PDFs for a fixed microlens abundance $\kappa_{\star} = 0.01$, but with $\mu_{\text{macro}} = -20, -50, -100$. The jaggedness of the curves in all panels, most noticeable for $\delta \gg 1$, arise from Monte-Carlo sampling noise.

gans 1999). For example, this was recently employed to quantify the statistics of microlensing on strongly lensed supernovae (SNe) (Goldstein et al. 2018).

However, inverse ray-tracing is computationally costly for adequately resolving very large magnifications $\delta \gtrsim 10^2$ near micro-caustics, a regime crucial for the more compact stellar sources than for quasars and SNe. As we have argued before, the large- δ tail makes a significant or even dominant contribution to $\langle \delta^2 \rangle_c$.

Instead, we implement a method proposed by Lewis et al. (1993) (see also Witt (1993); Lewis & Irwin (1995)), which solves the continuous motion and the possible creation and annihilation of the micro-images of a moving point source. Based on solving ordinary differential equations, this method can find all micro-images and accurately track them to arbitrarily high magnifications (Venumadhav et al. 2017).

We first generate a random realization of microlenses in the presence of macro convergence and shear, and then solve the light curve for a point source that traverses a few Einstein lengths on the

source plane, summing over all micro-images. By computing for a large number of random realizations, we build up a statistically adequate collection of random light curves, which enable us to evaluate $P(\delta)$ and $P_{\delta}(f)$, or any other statistics. While we numerically solve for point sources, we manually truncate the PDF at $\delta = 10^4$, appropriate for the typical stellar radii of massive main sequence stars or supergiants (Venumadhav et al. 2017).

In Fig. 2, we present $P(\delta)$ numerically evaluated for the parameter ranges appropriate for Image 5, $0.007 < \kappa_{\star} < 0.02$ and $-100 < \mu_{\text{macro}} < -20$. Qualitatively, most probability comes from a central core at $\delta \approx 0$, and some significant probability is associated with the strongly de-magnified ($1 + \delta \ll 1$) regions expected for a negative μ_{macro} (Chang & Refsdal 1984; Schechter & Wambsgans 2002; Diego et al. 2018). However, in nearly all cases what matters for the variance of flux fluctuation $\langle \delta^2 \rangle_c$ is primarily the $\delta \gg 1$ tail which nevertheless has very small weight in terms of probability. When $(1 + \delta)^2 dP/d \ln(1 + \delta)$ is plotted, this clearly shows as a plateau for $\delta > 0$, which confirms the power law $P(\delta) \propto \delta^{-3}$ and

supposedly extends all the way to the cutoff magnification. As can be seen from Fig. 2, the height of the plateau being comparable to the height of the $\delta \approx 0$ core marks the transition between optically thin and optically thick microlensing. Fig. 2 also shows that for fixed μ_{macro} and κ_{\star} , the full PDF $P(\delta)$ and hence the value of $\langle \delta^2 \rangle_c$ are insensitive to the direction of source motion ϕ .

For $\mu_{\text{macro}} = -20$ and reliably extrapolating to a truncation at $\delta_{\text{max}} = 10^4$ following the $P(\delta) \propto \delta^{-3}$ law, we find $\langle \delta^2 \rangle_c = 1.6, 2.5$ and 6.4 , for $\kappa_{\star} = 0.007, 0.01$ and 0.02 , respectively. If $\mu_{\text{macro}} = -50$, those numbers become $\langle \delta^2 \rangle_c = 4.6, 8.2$ and 11.4 . In this case, the RMS of the integrated flux variability Δ (defined as $\langle \Delta^2 \rangle_c$ in Eq. (13)) for Image 5 is enhanced by at least a factor $\sqrt{\langle \delta^2 \rangle_c} = 2-3$ relative to ϵ_2 , reaching 2–3%.

3.4 Power Spectrum

To understand the temporal structure of the flux variability, in Fig. 3 we present the variability power spectrum $P_{\delta}(f)$ for an individual source star, derived from simulated light curves. Since the variability power spectrum $P_{\Delta}(f)$ of the entire star cluster is proportional to $P_{\delta}(f)$, information on the same timescales can be extracted by measuring $P_{\Delta}(f)$.

We measure the spatial frequency f in units of ξ_0^{-1} , where ξ_0 is a fiducial Einstein scale

$$\xi_0 = \left(\frac{4G M_{\odot}}{c^2} \frac{D_L D_{LS}}{D_S} \right)^{1/2}, \quad (19)$$

corresponding to a $1 M_{\odot}$ microlens. In the case of microlensing toward Image 5, $\xi_0 \approx 2600$ AU. For the majority of the microlenses, which are sub-solar main-sequence stars of masses $\sim 0.2-0.3 M_{\odot}$, the Einstein scale is roughly half of ξ_0 .

Limited by computational cost, we only solve for light curve segments of finite lengths following the source motion across about order unity times ξ_0 , resolving no more than a dozen micro-caustic crossing features per light curve for the values of μ_{macro} and κ_{\star} we consider. We thus do not have information of $P_{\delta}(f)$ for $f\xi_0 \ll 1$, but we do have reliable results for $f\xi_0 \gg 1$ thanks to the fine resolution of our adaptively sampled light curves.

As shown in Fig. 3, $P_{\delta}(f)$ has a non-trivial shape around the characteristic Einstein scale of the microlenses, which should depend on the detailed microlens mass function. On the other hand, well below the Einstein scale $f\xi_0 \gg 1$, $P_{\delta}(f)$ shows a universal scale-invariant slope $P_{\delta}(f) \propto f^{-1}$, until truncated at $fR_S \sim \mathcal{O}(1)$ for a finite source radius R_S . This scale invariant power spectrum is a direct consequence of the universal behavior in the vicinity of a fold caustic, that $|\mu| \propto s^{-1/2}$ where s is the distance to the caustic on the source plane (Blandford & Narayan 1986), and therefore its Fourier transform goes as $f^{-1/2}$. We note that for $\mu_{\text{macro}} = -100$ and $\kappa_{\star} = 0.01$, our numerically evaluated $P_{\delta}(f)$ shows departure from perfect scale-invariance for moderately large values of $f\xi_0$. This probably can be attributed to the fact that at a very high number density micro-caustics strongly interfere with one another, which leads to violation of the simple scaling $|\mu| \propto s^{-1/2}$.

The scale invariance of $P_{\delta}(f)$ for $f\xi_0 \gg 1$ leads to an interesting conclusion: since $P_{\Delta}(f)$ is proportional to $P_{\delta}(f)$, and because the flux variability of the star cluster Δ is efficiently Gaussianized from the superposition of a large number of contributing source stars, Δ follows a scale-invariant process on those timescales.

In the case of the LyC Knot of Sunburst Arc, and assuming a typical observer-frame relative transverse velocity $v_t \sim 300$ km/s between the source and the lens (see Eq. (12) of Venumadhav et al. (2017)), the source-plane Einstein scales for typical microlenses

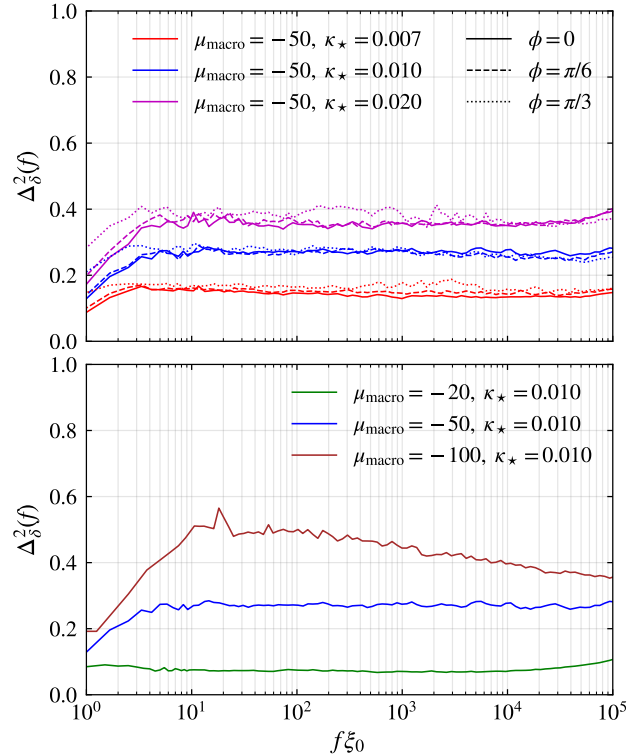


Figure 3. Characteristic power $\Delta_{\delta}^2(f) = f P_{\delta}(f)$ for the microlensing induced flux fluctuations of individual source stars, plotted as a function of the inverse spatial scale f on the source plane. Curves have been computed for the inferred microlens population in the foreground galaxy G1. For the horizontal axis, f is normalized to the inverse of the Einstein scale ξ_0 for a fiducial point microlens of $1 M_{\odot}$, which is about 2600 AU in the case of the LyC Knot, or equivalent to a timescale $\tau_0 \approx 40 \text{ yr} (v_t/300 \text{ km/s})^{-1}$ in the observer frame where v_t is the effective transverse velocity (Eq. (12) of Venumadhav et al. 2017) of the LyC Knot relative to the lens. Most stellar microlenses in G1 have masses $\sim 0.2-0.3 M_{\odot}$. **Top panel:** Cases of a fixed $\mu_{\text{macro}} = -50$ but with increasing values of the microlens surface density $\kappa_{\star} = 0.007, 0.01, 0.02$. We plot for different source motion directions across the magnification pattern on the source plane, $\phi = 0$ (solid), $\pi/6$ (dashed), and $\pi/3$ (dotted). **Bottom panel:** Cases of a fixed $\kappa_{\star} = 0.01$ but with different macro magnification factors $\mu_{\text{macro}} = -20, -50$, and -100 for Image 5. Except for the case of $\mu_{\text{macro}} = -100$ and $\kappa_{\star} = 0.01$ for which highly entangled micro-caustics are expected, $\Delta_{\delta}^2(f)$ exhibits a scale-invariant slope on scales well below the Einstein scales of the typical microlenses ($f\xi_0 > 10$). The trend that the curves start to rise near the shortest length scales (or the largest values of f) is likely due to artificial power induced by numerical interpolation of the simulated light curves near caustic crossing spikes.

translate to a rather long timescale on the order of decades. This is unappealing that accumulating adequate statistics to make inference about $P_{\delta}(f)$ around the Einstein scale would probably be prohibitive from the observational perspective. On the other hand, we suggest that it would be much more feasible to measure the scale-invariant portion of the power spectrum, e.g. at $f\xi_0 \approx 10^2-10^4$, which correspond to monitoring the LyC Knot over much more accessible timescales ranging from days to a year.

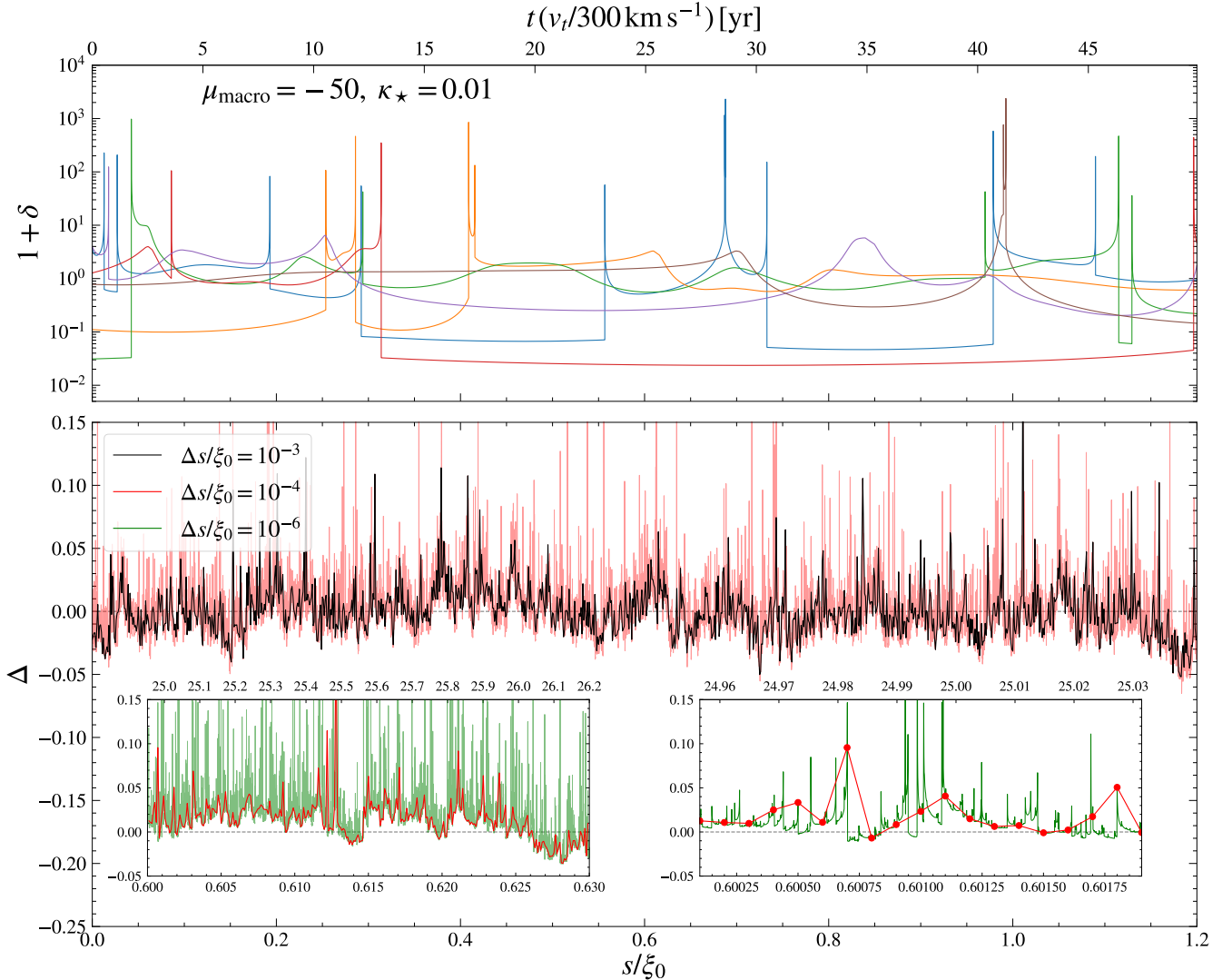


Figure 4. Example light curve of a mock star cluster that comprises the stellar population of the LyC Knot and has a total mass $M_{\star} = 10^7 M_{\odot}$, assuming $\mu_{\text{macro}} = -50$ and $\kappa_{\star} = 0.01$. Because of computational limitation, we only explicitly evaluate microlensing light curves for about 34000 brightest stars, which account for 90% of the mean integrated flux and 99.8% of the flux variance, but neglect the variabilities of the other fainter members. The bottom axis is the source-plane distance in units of the unit-solar-mass Einstein scale ξ_0 , and in the top axis this is converted into a total elapsed time of about 50 yr for a fiducial $v_t = 300$ km/s. We show the light curve in the F814W filter, but light curves in other filters from UV to near-IR are highly correlated. **Top panel:** Random light curves for several individual stars (in different colors) uniformly sampled at $\Delta s = 10^{-6} \xi_0$. Individual stars brighten by orders of magnitude during micro-caustic transits, which nevertheless occur infrequently over the timescale of decades. **Bottom panel:** Variability for the integrated flux uniformly sampled at different cadences $\Delta s/\xi_0 = 10^{-4}$ (red) and 10^{-3} (black). The **left and right insets** zoom into short-timescale details, over a one-year timescale (left inset) and a one-month timescale (right inset) respectively, which are difficult to discern in the plot of the full light curve. For the insets, two cadences $\Delta s/\xi_0 = 10^{-4}$ (red) and 10^{-6} (green) are shown. The integrated flux exhibits the usual highly non-Gaussian flaring behavior on very short timescales (green), increasingly Gaussianizes on intermediate timescales (red), and approaches a Gaussian process on long timescales (black).

3.5 Light Curve Example

To gain insight into the variable nature of the integrated flux, we show in Fig. 4 one example of a synthetic light curve for the LyC Knot. To be specific, we consider $\mu_{\text{macro}} = -50$ and $\kappa_{\star} = 0.01$, and set a cluster mass $M_{\star} = 10^7 M_{\odot}$ (Vanzella et al. 2020b). It is computationally prohibitive to explicitly simulate all member stars (Lewis et al. 2000) as there are $\sim \mathcal{O}(10^7)$ of them! For a good approximation, we only include the flux variability from a subset consisting of ~ 34000 bright stars. This subset account for 90% of the total mean flux \bar{F}_{cl} and 99.8% of the variance parameter ϵ_2 (and hence of

$\langle \Delta^2 \rangle_{\text{c}}$). According to Eq. (10), higher order cumulants of the cluster variability, $\langle \Delta^n \rangle_{\text{c}}$ for $n = 3, 4, \dots$, are increasingly dominated by the brightest member stars. We include the contribution to \bar{F}_{cl} from the numerous other fainter stars but simply neglect their variabilities. We note however that during micro-caustic crossings many of the faint stars can be magnified by larger factors than the brighter stars due to their small stellar sizes.

For an individual star, the flux evolves slowly for most of the time. Occasionally, the flux rapidly increases by several orders of magnitude around micro-caustic crossings, but (e.g. for $\mu_{\text{macro}} = -50$ and $\kappa_{\star} = 0.01$) such events only occur infrequently

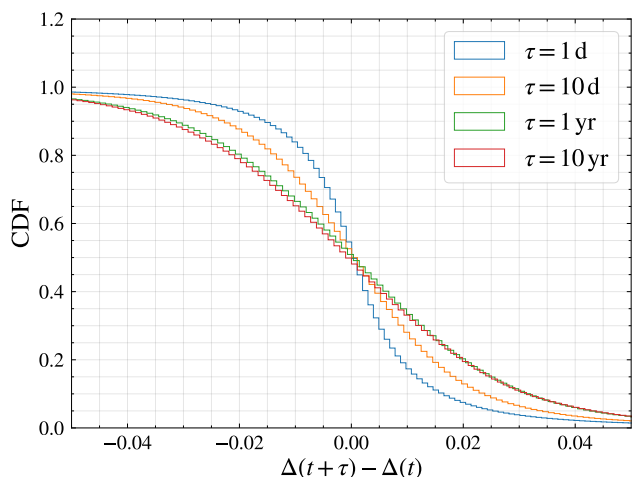


Figure 5. Cumulative distribution function of the fractional flux difference between two epochs separated by an amount of time τ , derived from the simulated light curve in Fig. 4. We show curves for several temporal separations, $\tau = 1$ d, 10 d, 1 yr and 10 yr.

over decades. For $\mu_{\text{macro}} < 0$, the flux can hide for a long period of time, but in this case the fractional flux change is bounded $\delta > -1$.

By contrast, the integrated flux reflects the collective effect of many source stars. As shown in the lower panel of Fig. 4, on long timescales $f\xi_0 \approx 1$ –100, i.e. from years to decades, the light curve resembles a scale-invariant Gaussian process, which, for the case $\mu_{\text{macro}} < 0$ we consider, is mainly driven by many source stars independently entering or leaving the extensive regions of low magnification in between caustic pairs (Chang & Refsdal 1984; Kelly et al. 2018; Diego et al. 2018). On the other hand, when sampled at high resolutions, i.e. $f\xi_0 \approx 100$ – 10^4 on timescales from days to months, the variable flux exhibits a highly non-Gaussian behavior. The most prominent feature of this behavior are flares from micro-caustic crossings that occur intermittently and frequently (referred to as the “extreme events” by Lewis & Ibata (2001); also note that the fast transients reported by Rodney et al. (2018) may be of such nature). We further note that on short timescales our simplification of including only the brightest stars must have left out a large number of lower-amplitude flaring events from micro-caustic crossings of the many more fainter stars.

Numerically, we find an overall $\sqrt{\langle \Delta^2 \rangle_c} \approx 2.5\%$ (corresponding to the green curve in Fig. 4 with the finest temporal sampling $\Delta s/\xi_0 = 10^{-6}$), in good agreement with Eq. (13) and Eq. (14). Fig. 5 shows the probability distribution for the fractional flux difference between two random epochs separated by an amount of time τ , i.e. $\Delta(t+\tau) - \Delta(t)$, computed using the simulated light curve shown in Fig. 4. Even at $\tau = 1$ d, the typical values for $|\Delta(t+\tau) - \Delta(t)|$ can be $\sim 0.5\%$, and this number grows to ~ 1 – 2% as the temporal baseline increases to one year. The distribution stays nearly the same between one-year and one-decade separations, except for extreme values of $|\Delta(t+\tau) - \Delta(t)|$. These results regarding the temporal structure indicate that the timescales of observational baselines must be taken into account when flux variability measurements at multiple epochs are compared to theory, and also have implications for the optimal observing strategy.

Another inevitable consequence is the random variation of the apparent color of the star cluster, defined as the magnitude difference between two chosen filters. This is examined in Fig. 6, which

is made based on the same realization used in Fig. 4. Measured relative to the HST F814W filter as the reference filter, we show the colors for the HST F275W filter, which uniquely probes hydrogen ionizing radiation, that for the HST F160W filter, which measures the rest-frame optical flux, and that for the JWST NIRCAM F444W filter, which measures the rest-frame near-IR flux. On timescales of years to decades, the color variabilities follow a Gaussian random behavior, with amplitudes on par with that of the flux variability. Similarly, non-Gaussian transient events of large-amplitude color variation are expected to be seen on shorter timescales, which are mainly associated with the most luminous individual stars transiting micro-caustics.

4 DISCUSSION

In this section, we comment on two issues: confusion with intrinsic variable stars, and stellar multiplicity. Furthermore, we suggest that the phenomenon of statistical microlensing can be used to constrain the abundance of compact objects that could partially constitute the DM.

4.1 Intrinsically Variable Stars

So far in our analysis we have assumed that all source stars have stationary fluxes in the absence of microlensing. This is an idealized assumption, as stars can exhibit variable fluxes, on a range of timescales (days to decades) and to various degrees, due to changes in the physical properties of the stars themselves.

When the microlensing optical depth is significant $O(0.1$ – $1)$, as we have studied before, every star is subject to order-unity flux variation due to lensing. Intrinsic variability can be a significant confusion effect in the context of this study only if the stars that dominate the observed total flux in a given band commonly have order-unity intrinsic variability. Observational data suggest that this is not the case for young stellar populations of only a few Myr old observed at short wavelengths. For example, Conroy et al. (2018) monitored the Whirlpool Galaxy (M51) using the HST for stellar variability over the timescale of days to months. Complete to $M_{I,814} < -4$ in the *i*-band, that study is sensitive to a diversity of luminous stars of a wide range of ages. The study found that all stars with $M_{I,814} < -3$ have relatively small amplitudes of intrinsic flux variation, $\sqrt{\langle \delta^2 \rangle_c} < 0.45$ (Figure 11 of Conroy et al. (2018)), except for Mira variables (Miras) and semi-regular variables (SRVs) which are typically red giant stars.

Our population model for the LyC Knot indicates that red stars with $M_{V,606} - M_{I,814} > 1$ have completely negligible contribution to F814W (rest-frame UV ~ 2000 – 2500) for $t_{\text{age}} < 50$ Myr, and only have a sub-percent contribution to the total flux in F160W (rest-frame optical ~ 5000) even at 50 Myr. Therefore, pulsating red giant stars such as Miras and SRVs are unimportant for the study of very young star clusters (a few Myr) at rest-frame UV or optical wavelength. Neglecting these red variables, all other luminous stars with $M_{I,814} < -3$ contribute an upper limit $\sqrt{\langle \delta^2 \rangle_c} < 0.45$, which, when applied to all stars, corresponds to only $\sqrt{\langle \Delta^2 \rangle_c} < 0.45 (\mu_{\text{macro}}/20)^{1/2} 0.6\% \approx (\mu_{\text{macro}}/20)^{1/2} 0.3\%$ for the variability of the integrated flux. In fact, at 3 Myr all stars with $M_{I,814} < -3$ account for 90% (87%) of the mean integrated flux $\langle F \rangle$ and 99.8% (99.9%) of the second moment $\langle F^2 \rangle$ in the F814W (F160W) filter. Additionally, this is likely to significantly overestimates the intrinsic variability, as Conroy et al. (2018) reported that

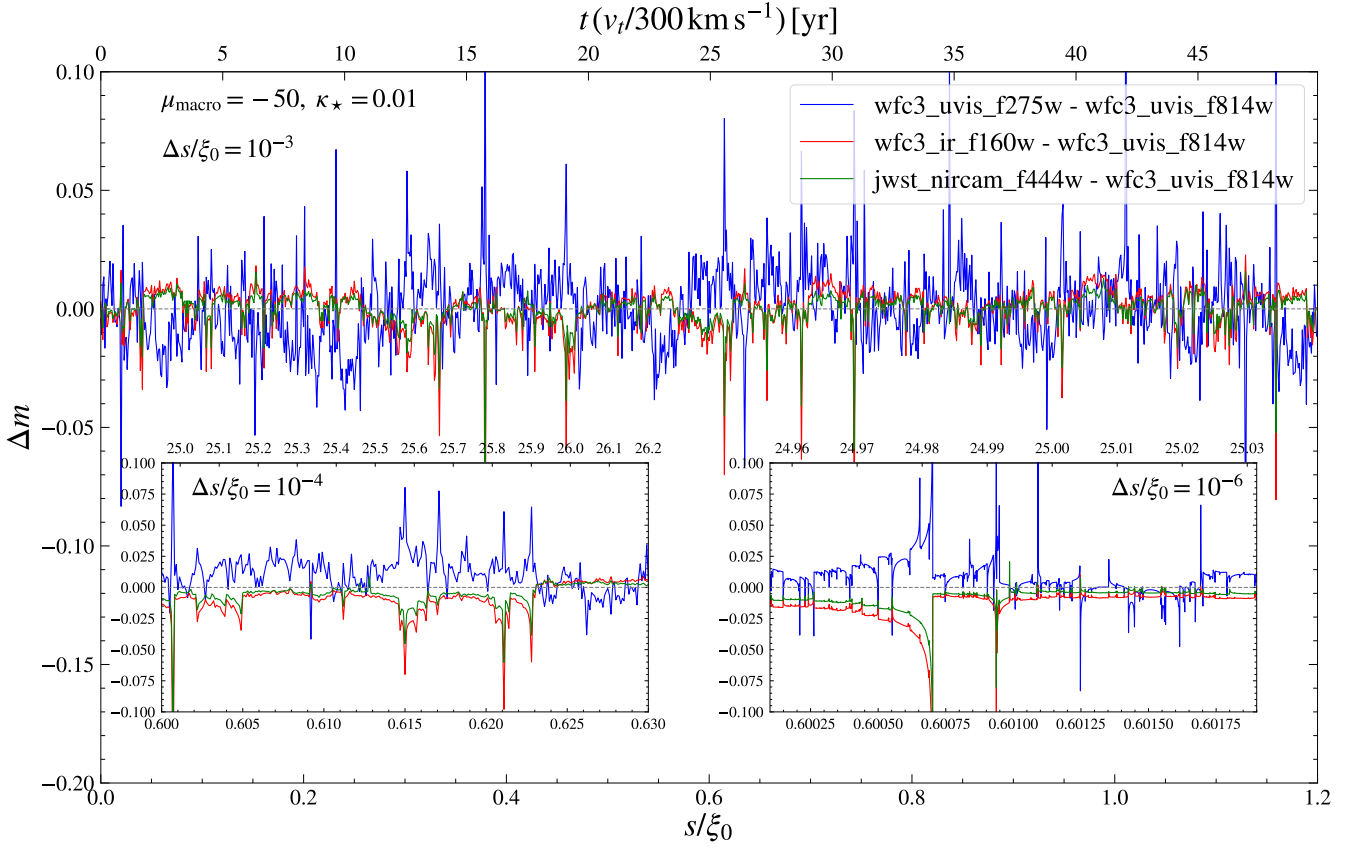


Figure 6. The same random light curve realization as in Fig. 4 but showing the color variation measured in magnitudes. Curves are uniformly sampled at a coarse resolution $\Delta s = 10^{-3} \xi_0$ in the **main panel**, and are sampled at finer resolutions at $\Delta s = 10^{-4} \xi_0$ and $10^{-6} \xi_0$ in the **left and right insets**, respectively. On long timescales, the color variability resembles a Gaussian random process. On short timescales, rapid events of large amplitude color variation are induced when individual bright stars transit micro-caustics.

the bright stars with $M_{I,814} < -6$ only have $\sqrt{\langle \delta^2 \rangle_c} < 0.1$. At 3 Myr, these stars contribute to 54% (57%) of $\langle \bar{F} \rangle$ and 90% (94%) of $\langle \bar{F}^2 \rangle$ in the F814W (F160W) filter. Also, Conroy et al. (2018) showed that the bluest stars on average have the lowest variability fraction, as those are mainly main sequence stars. Indeed, at 3 Myr we expect stars with $M_{V,606} - M_{I,814} < 0.4$ account for nearly the entirety of the integrated flux in F814W and F160W. Hence, intrinsic stellar variability in comparison to microlensing is probably negligible for a young star cluster of only a few Myr old when observed at rest-frame UV or optical wavelengths, as hot main sequence stars are photometrically very stable (Laur et al. 2017), and hot evolved (super-)giant stars (e.g. luminous blue variables (Smith 2017)) are rare in such a young cluster. For aged stellar systems at 100 Myr or older, or observation at rest-frame infrared wavelengths, pulsation or stellar outbursts of unstable evolved stars may cause confusion with microlensing induced flux variability.

Apart from time-dependent changes in the intrinsic properties of the stars, eclipsing binaries can exhibit variable fluxes of large amplitudes. What is relevant to the case of the LyC Knot is the eclipsing binary fraction for OB stars. Using HIPPARCOS data, Lefèvre et al. (2009) identified 127 eclipsing binaries out of ~ 2400 local OB stars. In particular, the fraction of eclipsing binaries with large flux variability amplitude > 0.4 mag is $\sim 1\%$. The generally small fraction of eclipsing binaries does not lead to any sizable flux variability in the cluster integrated flux \bar{F} that is comparable to microlensing effects.

More dramatic events such as supernova explosions are rarer to be found in any single star cluster, and when such an event does happen it will show unique photometric and spectroscopic behaviors to be distinguished from microlensing effects.

We emphasize that it is possible to empirically distinguish between microlensing effects and intrinsic variability by observing multiple lensed images of the same star cluster as in the case of the LyC Knot. It is expected that some images (such as Image 5) are subject to significantly more foreground microlensing than the others. The flux variance observed for the latter can be used to set the base intrinsic variability, and any excess variability observed for the former is a robust signature of microlensing.

4.2 Stellar Multiplicity

when building up a synthetic stellar population using FSPS and computing ϵ_2 , we have neglected the possibility of two or more stars forming a tightly bound system. In reality, massive stars in young star clusters are found to commonly have significant companions (Sana et al. 2012, 2013; Dunstall et al. 2015). This leads to an underestimate in ϵ_2 because members of a multiple stellar system traverse the source-plane magnification pattern in a correlated way. For example, if every star has an equal companion with perfectly correlated variability, then ϵ_2 increases by a factor $\sqrt{2} \approx 1.4$. This, in principle, opens up the interesting possibility of inferring the binary or multiple fraction of massive stars from the RMS amplitude

of the flux variation. To obtain robust results, however, one must accurately model both the microlens and the source population. State-of-the-art modeling of the source population must be carried out using all photometric and spectroscopic data, accounting for uncertainty in IMF, metallicity, dust reddening, and so on. The results should be verified for a more general star-formation history than what is assumed here. Another dynamical effect, which is interesting in its own right, is that rejuvenated or exotic stars may efficiently form from stellar collisions in a dense star cluster (Gaburov et al. 2008, 2010), which may significantly alter the stellar luminosity function and hence the RMS flux variation.

Those degeneracies may be broken by considering the imprints of binary or multiple stars in the variability power spectrum $P_{\Delta}(f)$. This is because, depending on the separation, the correlation in variability can be perfect on long timescales but degrades substantially on short timescales. Even very close companion stars still cross the same micro-caustic at slightly different times. This effect should break the scale-invariance of the power spectrum $P_{\Delta}(f)$ found on sub-Einstein scales (Fig. 3). Through a detailed measurement of the integrated flux variability, we may be able to infer not only the abundance of binary or multiple stars but also their separation distribution. More detailed investigations are left for future work.

4.3 Probing Compact Dark Matter

Extremely compact objects such as primordial black holes (PBHs) (Carr & Hawking 1974; Meszaros 1974; Chapline 1975) may constitute some significant fraction of the Dark Matter (DM) or make gravitational-wave sources detectable to LIGO/Virgo (Nakamura et al. 1997; Bird et al. 2016; Sasaki et al. 2016). Their abundance can be observationally probed based on their microlensing effects (Paczynski 1986; Griest 1991). Microlensing analyses of Galactic stars (Alcock et al. 2001; Griest et al. 2013), toward local-group galaxies (Tisserand et al. 2007; Niikura et al. 2019), and of extragalactic SNe (Zumalacárregui & Seljak 2018), stars (Diego et al. 2018; Oguri et al. 2018) and quasars (Mediavilla et al. 2017), together with other astrophysical (Brandt 2016) or cosmological tests (Ali-Haïmoud & Kamionkowski 2017), have constrained the mass fraction f_{PBH} of planetary or stellar mass PBHs to be no more than a few to tens of percent.

Flux variability induced by statistical microlensing acting on lensed star clusters would offer a novel opportunity to constrain the PBH abundance owing to a large DM column density along the line of sight. Indeed, it was proposed before that one could monitor surface brightness variations in galaxies at $z \sim 0.5$ (Lewis & Ibata 2001) or in giant arcs at $z \sim 1$ magnified by galaxy clusters (Lewis et al. 2000; Tuntsov et al. 2004) in order to probe a PBH population of $f_{\text{PBH}} = O(1)$.

Regarding the Sunburst Arc, the multiple images of the LyC Knot other than Image 5 could serve this purpose because our ICL modeling suggests $\kappa_{\star} \lesssim 0.002$ at the location of the Sunburst Arc for normal stellar microlenses. Given a PBH mass function, one would neglect intracluster stars by setting $\kappa_{\star} = f_{\text{PBH}} \kappa_0$, and repeat the same analysis as in the previous sections.

According to the discussion of Venumadhav et al. (2017), if PBHs in the mass range $\sim 10^{-6} M_{\odot} - 10^2 M_{\odot}$ contribute more than $f_{\text{PBH}} \approx 2\%$ of all DM, there will be an equivalent $\kappa_{\star} \gtrsim 0.01$, comparable to or exceed the (effectively) optically thick surface abundance of microlens stars from G1, and possibly a resultant fluctuating flux varying at the percent level. This suggests, through dedicated photometric monitoring of just one system, that statistical microlensing of magnified extragalactic star clusters has the

potential to probe $f_{\text{PBH}} \approx O(10^{-2})$ over a wide range of PBH masses. This would be highly competitive, and in particular complementary at sub-solar PBH masses, to other proposed methods based on microlensing of extragalactic radio bursts (Muñoz et al. 2016; Liao et al. 2020), GRBs (Ji et al. 2018) and chirping gravitational waves (Jung & Shin 2019). We note, however, that a higher abundance of PBHs is not necessarily easier to probe; a value for $\kappa_{\star} = f_{\text{PBH}} \kappa_0$ that is too large can instead reduce the fractional variability (Tuntsov et al. 2004). The optimal value of f_{PBH} to probe therefore depends on μ_{macro} , which will be different toward different macro images of the same star cluster or different star clusters. In light of future access to applicable photometric data, further study is warranted to derive the precise amplitude of flux variability for a broad range of f_{PBH} and PBH mass distributions.

5 CONCLUSION

We have studied the microlensing-induced flux variability of gravitationally magnified star clusters, which are typically found in lensed star-forming galaxies at high redshifts. We have made the suggestion that in galaxy cluster strong lenses the condition for optically thick microlensing is often satisfied because the usually small surface density of foreground stellar microlenses $\kappa_{\star} = O(10^{-2})$ is coupled to a large macro magnification $|\mu_{\text{macro}}| = O(10^{-10^2})$. Very young star clusters of only a few Myr old are particularly advantageous because their fluxes are dominated by a relatively small number of massive O main-sequence stars or evolved B supergiants (Gil-Merino & Lewis 2006), and hence the flux variability does not average out completely.

We have used a statistical framework to quantify the flux variability. We have shown that the variable flux exhibits rapid, frequent flares of large amplitude which arise from micro-caustic transits by individual stars, typically on timescales from days to months, while has a scale-invariant, Gaussian behavior of small amplitude on longer timescales from years to decades. We have conducted a case study of the LyC Knot discovered in the Sunburst Arc behind the galaxy cluster lens PSZ1-G311. We found that a minor foreground galaxy intervenes the line of sight toward one magnified image of the LyC Knot, providing a significant surface density of stellar microlenses. Through a population synthesis study, we predict that this lensed image should exhibit flux variability at the $\sim 1-2\%$ level, across a wide wavelength range from rest-frame UV to near-IR, and depending on the precise value of its macro magnification μ_{macro} . This prediction only mildly depends on the IMF slope and metallicity. Measuring the variability statistics therefore offers a microlensing-based method to determine the macro magnification μ_{macro} , and to probe the fraction of stellar multiplicity in massive stars.

How feasible would it be to measure this light variability? Theoretically, an SNR of 100 is needed to have 1% precision. Since the lensed images of the LyC Knot are remarkably bright (22–23 mag), 1–2% photometry should be achievable with $\sim 1-3$ ks exposures using the HST through observed optical or near-IR filters. With the forthcoming JWST, even shorter exposures will be sufficient. In fact, Rivera-Thorsen et al. (2019) detected the lensed images at $\text{SNR} > 10^3$ and reported aperture photometry precisions better than 0.1% with a total ~ 5 ks exposure in the HST F814W filter. Clearly, photon count noise will not be the major limitation from the space, in line with the earlier assessment of Lewis & Ibata (2001). Even without adaptive optics, percent-level photometry should be feasible from the ground using 4-meter class telescopes if photometry

is calibrated using reference stars. In either case, the key will be to uniform image reduction procedures and photometry schemes so as to prevent any systematic biases when comparing data acquired at different epochs, and to carefully subtract any uniform sky background that might differ between epochs. In this regard, applying the same analysis to lensed images unaffected by microlensing can serve as a control test. Ideally, measurements made at a large number of epochs will be required to enable a statistically meaningful comparison with the theory. It would be certainly an economic strategy to combine such a program with others that monitor the field of view of interesting galaxy cluster lenses in search of SNe and caustic transients.

Finally, we have pointed out that magnified images without significant intervention from foreground stars are clean targets to probe planetary to stellar mass compact objects in the galaxy cluster's DM halo. In the case of the Sunburst Arc, for example, we suggest that monitoring flux variability for the other lensed images of the LyC Knot will allow to probe as little as just a few percent of all DM in this form.

ACKNOWLEDGEMENTS

The author would like to thank Matthew Bayliss, Geraint Lewis, Jordi Miralda-Escudé, Emil Rivera-Thorsen, Artem Tuntsov, and Keiichi Umetsu for useful comments. The author also thanks the anonymous referee for a constructive review of the manuscript. The author acknowledges the support of John Bahcall Fellowships at the Institute for Advanced Study. This work is based on observations made with the NASA/ESA Hubble Space Telescope, obtained from the data archive at the Space Telescope Science Institute (STScI). STScI is operated by the Association of Universities for Research in Astronomy, Inc. under NASA contract NAS 5-26555. For our analysis, we use archival HST images available from the Barbara A. Mikulski Archive for Space Telescopes¹, taken with the Wide Field Camera 3 (WFC3) from three programs: (1) Proposal # 15101 (PI: H. Dahle) using F555W, F814W, F105W and F140W; (2) Proposal # 15377 (PI: M. Bayliss) using F606W, F125W and F160W; (3) Proposal # 15418 (PI: H. Dahle) using F275W. This work is also based on spatially resolved spectroscopic data collected at the European Southern Observatory (ESO) under ESO programme 297.A-5012 (PI: N. Aghanim). We accessed the public VLT/MUSE data from the ESO Archive².

DATA AVAILABILITY STATEMENT

The original data underlying this work are all publicly available from the Barbara A. Mikulski Archive for Space Telescopes and from the European Southern Observatory Archive. The derived data generated in this research will be shared upon reasonable request to the corresponding author.

¹ <https://mast.stsci.edu/portal/Mashup/Clients/Mast/Portal.html>

² <http://archive.eso.org/scienceportal/home>

APPENDIX A: DARK MATTER HALO OF PSZ1 G311.65-18.48 AND INTRACLUSTER LIGHT

The foreground lens, massive galaxy cluster PSZ1 G311.65-18.48 at $z_l = 0.44$, was first catalogued by the all-sky *Planck* galaxy cluster survey (Planck Collaboration et al. 2014) through the Sunyaev-Zeldovich (SZ) effect (Sunyaev & Zeldovich 1972). At the source redshift $z_s = 2.37$ of the Sunburst Arc, it has a spectacular angular Einstein radius $\theta_E \approx 30''$ (Dahle et al. 2016).

The multiple sections of the Sunburst Arc trace an ideal Einstein ring, which suggests that the cluster lens is highly spherical and dynamically relaxed. Neglecting a possibly small ellipticity, the observed Einstein radius θ_E translates into a total projected mass $M_E = 1.7 \times 10^{14} M_\odot$ enclosed within the Einstein radius.

Assuming that the stellar and gaseous contributions to M_E are subdominant (Rasia et al. 2004), the observed Einstein scale can be reproduced by a spherical Navarro-Frenk-White (NFW) profile (Navarro et al. 1996, 1997) with a concentration parameter $C_{200} = 8$ (throughout overdensity is defined relative to the cosmic critical density), a dynamic mass $M_{200} = 1.0 \times 10^{15} M_\odot$, a scale radius $R_s = 224$ kpc, and a lensing convergence at the Einstein radius $\kappa_E = 0.6$. This model gives $M_{500} = 7.7 \times 10^{14} M_\odot$, in agreement with the SZ-based dynamic mass estimate $M_{500SZ} = 6.6^{+0.9}_{-1.0} \times 10^{14} M_\odot$ (Dahle et al. 2016).

The chosen concentration parameter $C_{200} = 8$ is reasonable for high-mass “superlens” clusters (Oguri & Blandford 2009; Umetsu 2020) at $z \lesssim 0.5$ characterized by large Einstein radii $\theta_E \gtrsim 30''$ (for $z_s \sim 2$) and $M_{200} \sim 2 \times 10^{15} M_\odot$. If instead a higher value $C_{200} = 12$ is adopted, the profile more resembles an isothermal one. The dynamic mass is then reduced to $M_{200} = 7 \times 10^{14} M_\odot$, with a decreased scale radius $R_s = 133$ kpc and a smaller Einstein convergence $\kappa_E = 0.5$. Even in this case, $M_{500} = 5.6 \times 10^{14} M_\odot$ is still fairly consistent with the SZ-based mass measurement. For both choices, $2(1 - \kappa_E) = 0.8-1.0$ is close to unity, which is consistent with the narrow appearance of the arc's width. We use the value $C_{200} = 8$ for our fiducial cluster halo profile, according to which we use a fiducial value $\kappa_0 = 0.6$ throughout this work.

To determine the intracluster stellar population of PSZ1 G311, we perform SED fitting using the archival MUSE WFM spatially resolved spectroscopy data (Program 297.A-5012; PI: N. Aghanim). The data were acquired during May–August 2016 with seeing 0.5–0.8'' and a total exposure time of 1.2 hr, and have been reduced using the standard MUSE data reduction pipeline. The SNR is insufficient for a direct detection of the ICL in the MUSE data out to the Einstein radius, so we instead perform SED fitting within an aperture nearer to the brightest cluster galaxy (BCG), using BPASS synthetic SED templates. Assuming an SSP with a standard IMF and subject to Milky-Way dust reddening, we find an age $t_{SSP} \approx 5$ Gyr and a sub-solar metallicity $Z = 0.4 Z_\odot$. We then apply this solution to the ICL surface brightness in HST wide filters at the Einstein radius, which we find is fainter than 25 mag/arcsec² in the F160W filter, and hence derive $\kappa_\star \lesssim 0.002$.

APPENDIX B: STELLAR POPULATION OF THE FOREGROUND GALAXY G1

To infer its microlens population, we would like to model the stellar population of the foreground galaxy G1. To that end, we again choose to fit the stellar SED continuum using the MUSE IFU data, which offer richer information than from the photometry of wide HST filters. However, the seeing limited spatial resolution prevents

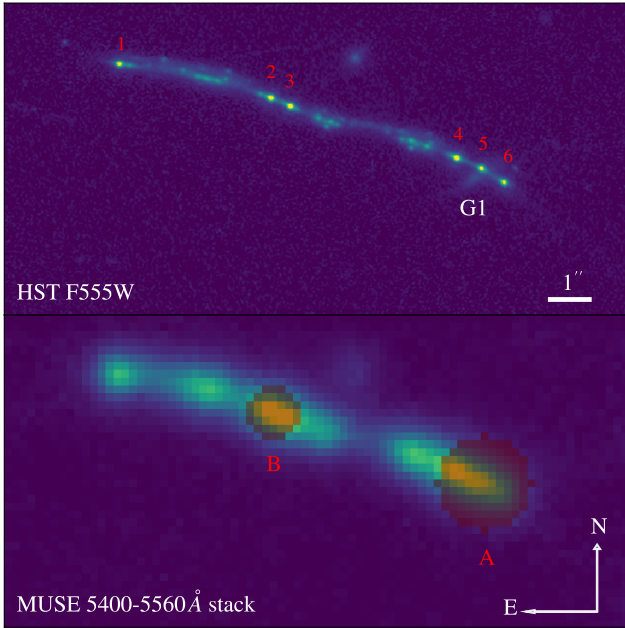


Figure B1. The same field of view through the HST F555W filter (top panel) and in the MUSE IFU image stacked across the wavelength range 5400–5560Å (bottom panel). The field of view shows Arc 1 of the Sunburst Arc, along which multiple lensed images of the LyC Knot denoted as Images 1 through 6 are clearly seen ((we follow the denotation of [Rivera-Thorsen et al. 2019](#))). The faint foreground galaxy G1 appears as a low-surface-brightness, elongated ellipsoid whose outskirts overlaps with the line of sight toward Image 5. In the MUSE IFU image, we show Apertures A and B which we use to extract the continuum SED of the galaxy G1.

us from directly separating the light of G1 from that of the nearby lensed images (Image 4, 5 and 6) of the LyC Knot, unlike in the case of wide filter HST images (Fig. B1).

To overcome this problem, we define two circular apertures and compute the integrated spectra enclosed within them. Aperture A has a radius of $1.2''$ and has its integrated flux dominated by Images 4, 5, and 6 of the LyC Knot, as well as the intervening G1. Aperture B has a radius of $0.38''$ and has its flux dominated by Images 2 and 3 of the LyC Knot. These aperture radii have been chosen to minimize contamination from nearby surface brightness features along the arc. We measure and subtract any uniform background, for Apertures A and B, respectively, by positioning the same aperture at a nearby empty place within the field of view. Under the assumption that fluxes from the various lensed images of the LyC Knot have identical spectral shapes, we subtract a suitable rescaling of the Aperture B flux from Aperture A flux such that the residuals reflect the contribution from G1.

When 1.6 times of the Aperture B flux is subtracted from the Aperture A flux, as shown in the top panel of Fig. B2, we completely remove the major spectral features of the LyC Knot, which include C IV $\lambda\lambda 1548, 1550 \text{ \AA}$, He II $\lambda 1640 \text{ \AA}$, O III] $\lambda\lambda 1661, 1666 \text{ \AA}$, and C III] $\lambda\lambda 1907, 1909 \text{ \AA}$ ([Vanzella et al. 2020b](#)). The remaining flux, which supposedly comes from G1, exhibit Balmer absorption features H ϵ , H δ and H γ , from which we infer a redshift $z_{G1} = 0.458$. This is close to, but slightly larger, than the spectroscopic redshift derived for the cluster lens PSZ1-G311 $z_l = 0.443$ ([Dahle et al. 2016](#)), suggesting that G1 is not a cluster member galaxy. In any case, for the purpose of quantifying the intervening microlensing effects

toward Image 5 under the influence of macro lensing by the galaxy cluster PSZ1-G311, it is a good approximation to assume that G1 has the same redshift as PSZ1-G311.

For a control test, we now center Aperture A at Image 1 and repeat the same procedure. This time, we remove the major spectral features by subtracting from its flux 0.47 times the flux of Aperture B. The residuals are more than a factor of 5 smaller than the magenta curve in Fig. B2 for G1. The small residuals are biased toward positive values probably because even away from the images of the LyC Knot the Sunburst Arc itself has a finite surface brightness, which is not expected to be removed by the subtraction method. This verifies that the continuum SED we have extracted for G1 is robust and is not severely contaminated by other sources.

Due to possible extra dust reddening by G1, Image 5 may differ from the other lensed images of the LyC Knot in the SED shape, which is unaccounted for in the above subtraction procedure. However, dust reddening due to G1 is likely mild $E_{G1}(B-V) \simeq 0.05$ for $R_V = 3.1$, according to the SED fitting of G1 as we will show next, and also because the line of sight toward Image 5 intersects only the outskirts of G1.

The subtraction residuals are dereddened assuming a Milky Way dust attenuation law $R_V = 3.1$ and $E_{MW}(B-V) = 0.08$ along the line of sight according to the model of [Schlafly & Finkbeiner \(2011\)](#), redshifted to the source frame, and are then fitted to BPASS ([Eldridge et al. 2017](#); [Stanway & Eldridge 2018](#)) synthetic SEDs for a simple stellar population (SSP). The BPASS synthetic SEDs are for a double power-law initial mass function (IMF) ([Kroupa 2001b](#)) which has a slope $\alpha = -d \log N / d \log m = 1.30$ for main sequence masses $0.1 M_\odot < m < 0.5 M_\odot$, and a steepened slope $\alpha = 2.35$ for $0.5 M_\odot < m < 100 M_\odot$.

The SED continuum is best fit by an SSP of $t_{SSP} = 630 \text{ Myr}$ old at sub-solar metallicity $Z = 0.4 Z_\odot$, with some additional amount of local dust reddening $R_V = 3.1$ and $E_{G1}(B-V) = 0.05$ (bottom panel of Fig. B2). When the stellar population is allowed to comprise multiple components of different ages $10 \text{ Myr} < t_{SSP} < 10 \text{ Gyr}$, and of different metallicities $0.05 < Z/Z_\odot < 2$, the fitting results remain stable and point toward the same SSP. When the source-frame dust reddening is varied $0.0 < E_{G1}(B-V) < 0.2$, the best-fit metallicity alters, but the stellar mass and age only change by small amounts. The best-fit stellar population also reproduce the photometric magnitudes of G1 in the wide HST filters.

Due to blending, it is difficult to directly measure the surface brightness of G1 at the centroid of Image 5. Instead, we assume that the isophotes of G1 have ellipsoidal symmetry, and find the line of sight whose location is symmetric about the G1 center (peak of surface brightness) with the one toward Image 5. We assume that the surface brightness along that line of sight equals that along the line of sight toward Image 5.

As listed in Tab. B1, several SSP solutions that provide a good fit all predict an average contribution to the lensing convergence from G1 member stars toward Image 5, $\kappa_\star = \Sigma_\star / \Sigma_{\text{crit}} \approx 1-3\%$. Here Σ_\star is the stellar surface mass density, and $\Sigma_{\text{crit}} = (c^2/4\pi G)(D_S/D_L D_{LS}) = 2.0 \times 10^9 M_\odot/\text{kpc}^2$ is the critical surface density, with D_L , D_S and D_{LS} being the angular diameter distances to the lens plane at $z_l = 0.44$, to the source plane at $z_s = 2.37$, and from the lens plane to the source plane. Variation in metallicity, stellar age and local dust reddening in these SSP solutions only leads to moderate changes in κ_\star , while the most noticeable uncertainty comes from the choice of HST filter used to estimate the surface brightness. This may be due to the filter dependence of the point spread function (PSF).

The numbers quoted in Tab. B1 are not corrected for stel-

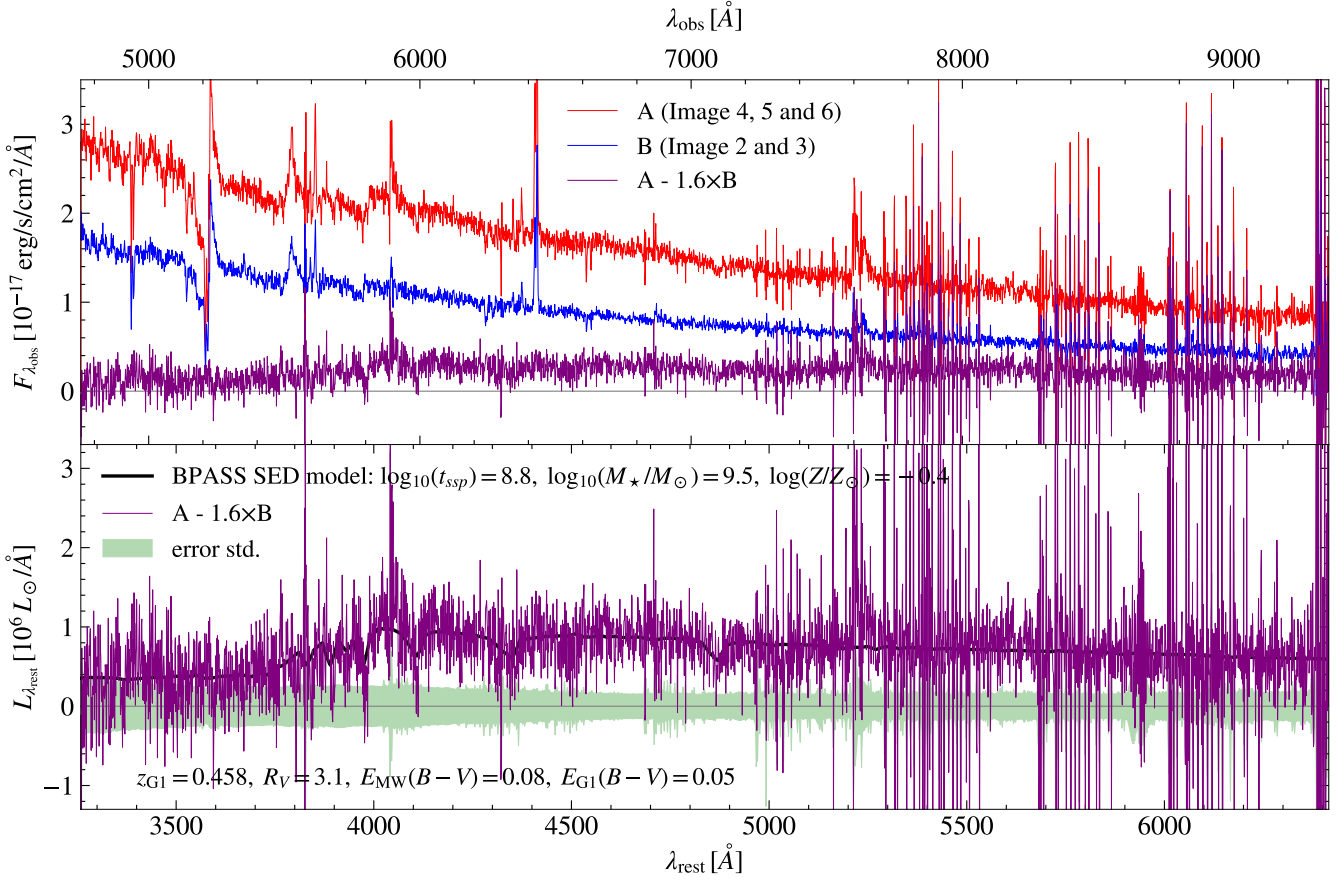


Figure B2. Modeling of the stellar population of the foreground galaxy G1 at $z_{G1} = 0.458$ using MUSE IFU data. *Top:* Observed-frame flux density integrated over Aperture A (red; containing Images 4, 5 and 6, and G1), Aperture B (blue; containing Images 2 and 3), and the subtraction of latter from the former (magenta). *Bottom:* Fitting of G1 rest-frame luminosity density (magenta; same as in top panel) using BPASS synthetic SEDs. We show the best-fit SSP (black) at an age $t_{SSP} \approx 600$ Myr, a metallicity $Z = 0.4 Z_{\odot}$, and a total stellar mass $M_{\star} = 3 \times 10^9 M_{\odot}$. Curves are dust reddened from within the Milky Way with $E_{MW}(B - V) = 0.08$ (Schlafly & Finkbeiner 2011), and from within G1 with $E_{G1}(B - V) = 0.05$. The green shaded band indicates the size of the statistical uncertainty including sky background contamination. Many strong and sharp spectral features are due to instrumental artifacts or sky lines unrelated to the astrophysical sources.

lar mass loss. According to stellar population synthesis codes (e.g. Flexible Stellar Population Synthesis (FSPS) (Conroy et al. 2009; Conroy & Gunn 2010)), at an age $\log_{10}(t_{SSP}/\text{yr}) = 8.7-8.8$ about 70% of the total initial stellar mass (including stellar remnants) survives either mass loss or stellar explosion. Applying this correction, we estimate that $\kappa_{\star} = 0.7-2^{\sigma}$.

REFERENCES

- Alcock C., et al., 1999, *ApJ*, 521, 602
 Alcock C., et al., 2001, *ApJ*, 550, L169
 Ali-Haïmoud Y., Kamionkowski M., 2017, *Phys. Rev. D*, 95, 043534
 Bird S., Cholis I., Muñoz J. B., Ali-Haïmoud Y., Kamionkowski M., Kovetz E. D., Raccanelli A., Riess A. G., 2016, *Phys. Rev. Lett.*, 116, 201301
 Blandford R., Narayan R., 1986, *ApJ*, 310, 568
 Brandt T. D., 2016, *ApJ*, 824, L31
 Carr B. J., Hawking S. W., 1974, *MNRAS*, 168, 399
 Chang K., Refsdal S., 1984, *A&A*, 132, 168
 Chapline G. F., 1975, *Nature*, 253, 251
 Chen W., et al., 2019, *ApJ*, 881, 8
 Chisholm J., Rigby J. R., Bayliss M., Berg D. A., Dahle H., Gladders M., Sharon K., 2019, *ApJ*, 882, 182
 Conroy C., Gunn J. E., 2010, *ApJ*, 712, 833
 Conroy C., Gunn J. E., White M., 2009, *ApJ*, 699, 486
 Conroy C., et al., 2018, *ApJ*, 864, 111
 Crotts A. P. S., 1992, *ApJ*, 399, L43
 Crotts A. P. S., Tomaney A. B., 1996, *ApJ*, 473, L87
 Crowther P. A., et al., 2016, *MNRAS*, 458, 624
 Dahle H., et al., 2016, *A&A*, 590, L4
 Dai L., Venumadhav T., Kaurov A. A., Miralda-Escud J., 2018, *ApJ*, 867, 24
 Dai L., et al., 2020, *MNRAS*, 495, 3192
 Deguchi S., Watson W. D., 1987, *Phys. Rev. Lett.*, 59, 2814
 Diego J. M., 2019, *A&A*, 625, A84
 Diego J. M., et al., 2018, *ApJ*, 857, 25
 Dunstall P. R., et al., 2015, *A&A*, 580, A93
 Eldridge J. J., Stanway E. R., Xiao L., McClelland L. A. S., Taylor G., Ng M., Greis S. M. L., Bray J. C., 2017, *Publ. Astron. Soc. Australia*, 34, e058
 Fleury P., García-Bellido J., 2020, *Physics of the Dark Universe*, 29, 100567
 Gaburov E., Gualandris A., Portegies Zwart S., 2008, *MNRAS*, 384, 376
 Gaburov E., Lombardi James C. J., Portegies Zwart S., 2010, *MNRAS*, 402, 105
 Gil-Merino R., Lewis G. F., 2006, *ApJ*, 643, 260
 Goldstein D. A., Nugent P. E., Kasen D. N., Collett T. E., 2018, *ApJ*, 855,

Table B1. Several comparably good fits to the stellar population of G1 using the BPASS models. We list the corresponding predictions for the stellar contribution to the lensing convergence κ_{\star} along the line of sight to Image 5 as inferred from the HST images in the F555W, F606W and F814W filters, respectively. Multiple visits of the same filter are averaged over. A Milky Way like dust extinction law $R_V = 3.1$ is assumed when applying source-frame dust reddening set by $E_{G1}(B - V)$. Mass loss effects from stellar evolution are not accounted for in the table, which we estimate cause a further 70% reduction in κ_{\star} .

$\log_{10}(Z/Z_{\odot})$	$\log_{10}(t_{\text{SSP}}/\text{yr})$	$E_{G1}(B - V)$	κ_{\star} (from F555W)	κ_{\star} (from F606W)	κ_{\star} (from F814W)
-0.30	8.8	0.00	0.012	0.025	0.016
-0.30	8.8	0.02	0.013	0.027	0.017
-0.40	8.8	0.05	0.013	0.029	0.018
-0.15	8.7	0.08	0.014	0.029	0.018
-0.15	8.7	0.10	0.015	0.032	0.019

- 22
- Gott J. R. I., 1981, *ApJ*, **243**, 140
- Griest K., 1991, *ApJ*, **366**, 412
- Griest K., Cieplak A. M., Lehner M. J., 2013, *Phys. Rev. Lett.*, **111**, 181302
- Jeřábková T., Kroupa P., Dabringhausen J., Hilker M., Bekki K., 2017, *A&A*, **608**, A53
- Ji L., Kovetz E. D., Kamionkowski M., 2018, *Phys. Rev. D*, **98**, 123523
- Jung S., Shin C. S., 2019, *Phys. Rev. Lett.*, **122**, 041103
- Katz N., Balbus S., Paczynski B., 1986, *ApJ*, **306**, 2
- Kaurov A. A., Dai L., Venumadhav T., Miralda-Escudé J., Frye B., 2019, *ApJ*, **880**, 58
- Kelly P. L., et al., 2018, *Nature Astronomy*, **2**, 334
- Kelly P., et al., 2019, Flashlights: Many Extremely Magnified Individual Stars as Probes of Dark Matter and Stellar Populations to Redshift z , 2, HST Proposal
- Kofman L., Kaiser N., Lee M. H., Babul A., 1997, *ApJ*, **489**, 508
- Kroupa P., 2001a, *MNRAS*, **322**, 231
- Kroupa P., 2001b, *MNRAS*, **322**, 231
- Laur J., Kolka I., Eenmäe T., Tuvikene T., Leedjärv L., 2017, *A&A*, **598**, A108
- Lee M. H., Babul A., Kofman L., Kaiser N., 1997, *ApJ*, **489**, 522
- Lefèvre L., Marchenko S. V., Moffat A. F. J., Acker A., 2009, *A&A*, **507**, 1141
- Lewis G. F., Ibata R. A., 2001, *ApJ*, **549**, 46
- Lewis G. F., Irwin M. J., 1995, *MNRAS*, **276**, 103
- Lewis G. F., Irwin M. J., 1996, *MNRAS*, **283**, 225
- Lewis G. F., Miralda-Escudé J., Richardson D. C., Wambsganss J., 1993, *MNRAS*, **261**, 647
- Lewis G. F., Ibata R. A., Wyithe J. S. B., 2000, *ApJ*, **542**, L9
- Liao K., Zhang S. B., Li Z., Gao H., 2020, *ApJ*, **896**, L11
- Lin Y.-T., Mohr J. J., 2004, *ApJ*, **617**, 879
- Ma X., Hopkins P. F., Kasen D., Quataert E., Faucher-Giguère C.-A., Kereš D., Murray N., Strom A., 2016, *MNRAS*, **459**, 3614
- Marchandon S., Nottale L., 1991, *A&A*, **251**, 393
- Mediavilla E., Jiménez-Vicente J., Muñoz J. A., Vives-Arias H., Calderón-Infante J., 2017, *ApJ*, **836**, L18
- Meszáros P., 1974, *A&A*, **37**, 225
- Miralda-Escudé J., 1991, *ApJ*, **379**, 94
- Muñoz J. B., Kovetz E. D., Dai L., Kamionkowski M., 2016, *Phys. Rev. Lett.*, **117**, 091301
- Nakamura T., Sasaki M., Tanaka T., Thorne K. S., 1997, *ApJ*, **487**, L139
- Navarro J. F., Frenk C. S., White S. D. M., 1996, *ApJ*, **462**, 563
- Navarro J. F., Frenk C. S., White S. D. M., 1997, *ApJ*, **490**, 493
- Neendorf B., 2003, *A&A*, **404**, 83
- Niikura H., et al., 2019, *Nature Astronomy*, **3**, 524
- Oguri M., Blandford R. D., 2009, *MNRAS*, **392**, 390
- Oguri M., Diego J. M., Kaiser N., Kelly P. L., Broadhurst T., 2018, *Phys. Rev. D*, **97**, 023518
- Paczynski B., 1986, *ApJ*, **304**, 1
- Peacock J. A., 1982, *MNRAS*, **199**, 987
- Planck Collaboration et al., 2014, *A&A*, **571**, A29
- Priewe J., Williams L. L. R., Liesenborgs J., Coe D., Rodney S. A., 2017, *MNRAS*, **465**, 1030
- Rasia E., Tormen G., Moscardini L., 2004, *MNRAS*, **351**, 237
- Ricotti M., 2002, *MNRAS*, **336**, L33
- Rivera-Thorsen T. E., et al., 2017, *A&A*, **608**, L4
- Rivera-Thorsen T. E., et al., 2019, *Science*, **366**, 738
- Rodney S. A., et al., 2018, *Nature Astronomy*, **2**, 324
- Rosdahl J., et al., 2018, *MNRAS*, **479**, 994
- Sana H., 2017, in Eldridge J. J., Bray J. C., McClelland L. A. S., Xiao L., eds, IAU Symposium Vol. 329, The Lives and Death-Throes of Massive Stars. pp 110–117 ([arXiv:1703.01608](https://arxiv.org/abs/1703.01608)), [doi:10.1017/S1743921317003209](https://doi.org/10.1017/S1743921317003209)
- Sana H., Evans C. J., 2011, in Neiner C., Wade G., Meynet G., Peters G., eds, IAU Symposium Vol. 272, Active OB Stars: Structure, Evolution, Mass Loss, and Critical Limits. pp 474–485 ([arXiv:1009.4197](https://arxiv.org/abs/1009.4197)), [doi:10.1017/S1743921311011124](https://doi.org/10.1017/S1743921311011124)
- Sana H., et al., 2012, *Science*, **337**, 444
- Sana H., et al., 2013, *A&A*, **550**, A107
- Sasaki M., Suyama T., Tanaka T., Yokoyama S., 2016, *Phys. Rev. Lett.*, **117**, 061101
- Schechter P. L., Wambsganss J., 2002, *ApJ*, **580**, 685
- Schlafly E. F., Finkbeiner D. P., 2011, *ApJ*, **737**, 103
- Schneider P., 1987, *ApJ*, **319**, 9
- Schneider P., Weiss A., 1988, *ApJ*, **327**, 526
- Seitz C., Schneider P., 1994, *A&A*, **288**, 1
- Seitz C., Wambsganss J., Schneider P., 1994, *A&A*, **288**, 19
- Smith N., 2017, *Philosophical Transactions of the Royal Society of London Series A*, **375**, 20160268
- Stanway E. R., Eldridge J. J., 2018, *MNRAS*, **479**, 75
- Stanway E. R., Eldridge J. J., Becker G. D., 2016, *MNRAS*, **456**, 485
- Sunyaev R. A., Zeldovich Y. B., 1972, Comments on Astrophysics and Space Physics, **4**, 173
- Tisserand P., et al., 2007, *A&A*, **469**, 387
- Tuntsov A. V., Lewis G. F., Ibata R. A., Kneib J. P., 2004, *MNRAS*, **353**, 853
- Umetsu K., 2020, arXiv e-prints, [p. arXiv:2007.00506](https://arxiv.org/abs/2007.00506)
- Vanzella E., et al., 2017, *ApJ*, **842**, 47
- Vanzella E., et al., 2020a, *MNRAS*,
- Vanzella E., et al., 2020b, *MNRAS*, **491**, 1093
- Venumadhav T., Dai L., Miralda-Escudé J., 2017, *ApJ*, **850**, 49
- Vietri M., Ostriker J. P., 1983, *ApJ*, **267**, 488
- Wambsganss J., 1999, Journal of Computational and Applied Mathematics, **109**, 353
- Wambsganss J., 2001, in Menzies J. W., Sackett P. D., eds, Astronomical Society of the Pacific Conference Series Vol. 239, Microlensing 2000: A New Era of Microlensing Astrophysics. p. 351 ([arXiv:astro-ph/0008419](https://arxiv.org/abs/astro-ph/0008419))
- Witt H. J., 1993, *ApJ*, **403**, 530
- Wyithe J. S. B., Turner E. L., 2001, *MNRAS*, **320**, 21
- Wyithe J. S. B., Turner E. L., 2002, *ApJ*, **575**, 650
- Young P., 1981, *ApJ*, **244**, 756
- Zibetti S., White S. D. M., Schneider D. P., Brinkmann J., 2005, *MNRAS*, **358**, 949
- Zumalacáregui M., Seljak U., 2018, *Phys. Rev. Lett.*, **121**, 141101
- Zwicky F., 1951, *PASP*, **63**, 61