OGLE-2016-BLG-1195 AO: Lens, Companion to Lens or Source, or None of the Above?

Andrew Gould[1,2], Yossi Shvartzvald[3] , Jiyuan Zhang[4], Jennifer C. Yee[5] , Sebastiano Calchi Novati[6] ,
Weicheng Zang[4,5] , and Eran O. Ofek[3]

[1] Max-Planck-Institute for Astronomy, Königstuhl 17, D-69117 Heidelberg, Germany; gould.34@osu.edu
[2] Department of Astronomy, Ohio State University, 140 W. 18th Ave., Columbus, OH 43210, USA
[3] Department of Particle Physics and Astrophysics, Weizmann Institute of Science, Rehovot 76100, Israel
[4] Department of Astronomy, Tsinghua University, Beijing 100084, People's Republic of China
[5] Center for Astrophysics | Harvard & Smithsonian, 60 Garden St., Cambridge, MA 02138, USA
[6] IPAC, Mail Code 100-22, Caltech, 1200 E. California Blvd., Pasadena, CA 91125, USA

Abstract

We systematically investigate Vandorou et al.'s claim to have detected the host star of the low-mass-ratio ($q < 10^{-4}$) microlensing planet OGLE-2016-BLG-1195Lb, via Keck adaptive optics (AO) measurements $\Delta t = 4.12$ yr after the event's peak (t_0). If correct, this measurement would contradict the microlens-parallax measurement derived from Spitzer observations taken near t_0. We show that this host identification would be in 4σ conflict with the original ground-based relative lens–source proper-motion measurements. By contrast, Gould estimated a probability $p = 10\%$ that the "other star" resolved by single-epoch late-time AO would be a companion to the host or the microlensed source, which is much more probable than a 4σ statistical fluctuation. Independent of this proper-motion discrepancy, the kinematics of this host identification are substantially less probable than those of the Spitzer solution. Hence, this identification should not be accepted, pending additional observations that would either confirm or contradict it, which could be taken in 2023. Motivated by this tension, we present two additional investigations. We explore the possibility that Vandorou et al. identified the wrong "star" for their analysis. Astrometry of KMT and Keck images favors a star (or asterism) lying about 175 mas northwest of Vandorou et al.'s star. We also present event parameters from a combined fit to all survey data, which yields a more precise mass ratio, $q = (4.6 \pm 0.4) \times 10^{-5}$. Finally, we discuss the broader implications of minimizing such false positives for the first measurement of the planet mass function, which will become possible when AO on next-generation telescopes are applied to microlensing planets.

Unified Astronomy Thesaurus concepts: High-resolution microlensing event imaging (2138); Satellite microlensing parallax (2148); Gravitational microlensing exoplanet detection (2147)

Supporting material: tar.gz file

1. Introduction

The most systematically applicable method of measuring host masses for microlensing planets is late-time imaging of the system when the host and source have separated sufficiently to resolve them. The only fundamental requirements of the method are that the host be luminous and that the planet-to-host mass ratio q and Einstein timescale t_E had been adequately measured from the original event. If the host is luminous, the measurement yields the heliocentric lens–source relative proper motion $\mu_{\rm rel,hel}$ and the lens flux (in, e.g., the K band) K_L. By adjusting $\mu_{\rm rel,hel}$ to the geocentric value $\mu_{\rm rel}$, and combining this with t_E, one obtains the angular Einstein radius $\theta_E = \mu_{\rm rel} t_E$. Then,

$$\theta_E \equiv \sqrt{\kappa M \pi_{\rm rel}}, \qquad \kappa \equiv \frac{4G}{c^2 {\rm au}} \simeq 8.14 \, \frac{\rm mas}{M_\odot}, \qquad (1)$$

and,

$$K_L = M_K(M_{\rm host}) + 5\log\frac{D_L}{10 \, \rm pc} + A_K(D_L), \qquad (2)$$

providing two relations between the host mass $M_{\rm host}$ and distance D_L, which can then be solved for both quantities.

Then, the mass of the planet $m_p = qM_{\rm host}$ can be found from the known value of q. Here, $\pi_{\rm rel} \equiv {\rm au}(D_L^{-1} - D_S^{-1})$ is the lens–source relative parallax, $M_K(M_{\rm host})$ is the absolute magnitude of the lens, and $A_K(D_L)$ is its extinction.

Gould (2022) examined a wide range of issues associated with this method, including degeneracies in converting from $\mu_{\rm rel,hel}$ to $\mu_{\rm rel}$ and in combining Equations (1) and (2), uncertainties in various parameters such as the mass–luminosity relation $M_K(M_{\rm host})$, the source parallax π_S, the extinction $A_K(D_L)$, and the light-curve measurement of t_E.

The most vexing issue identified by Gould (2022) was that the "other object" (i.e., other than the source) might not be the lens. Rather it could be either a companion to the lens, a companion to the source, or in rare cases, an ambient star. Gould (2022) showed that such misidentifications do not pose an issue of principle: they can almost always be resolved by additional late-time observations. In all nonpathological cases, if the object is not the lens, the relative proper-motion vector derived from the two late-time observations will not be consistent with uniform motion and lens–source coincidence at the time of the event. From the discrepancy, one can usually distinguish among these cases, and in particular if the object is a companion to the source, then it will hardly move between epochs, thereby permitting the lens to appear subsequently.

The main issue is that, in general, one may not know when to take a second late-time observation. Adaptive optics (AO) observations on large, or extremely large telescopes (ELTs) are

expensive, while Gould (2022) estimated that of order 150 mass measurements could be made at AO first light on ELTs from the 2016 to 2022 planet detections by the Korean Microlensing Telescope Network (KMTNet; Kim et al. 2016) alone. Hence, he considered the issue of the false-positive rate to be crucial.

Motivated by these considerations, Gould (2022) distinguished between two cases: those with and without measurements of $\mu_{\rm rel}$ from the event itself. Such measurements require that the normalized source radius $\rho = \theta_*/\theta_{\rm E}$ be measured, which in turn requires that the source passes over a caustic or close to a cusp (and that these are covered by the observations). In this case, it is possible to determine $\theta_{\rm E} = \theta_*/\rho$ and so $\mu_{\rm rel} = \theta_{\rm E}/t_{\rm E}$ from the angular source radius θ_*, which can be determined by standard techniques (Yoo et al. 2004).

Gould (2022) estimated that for about 1/4 of microlensing planets, $\mu_{\rm rel}$ cannot be measured from the light curve (i.e., finite-source effects are not detected). For these, he estimated that the false-positive rate (mainly from companions to the lens), would be about 10%. Nominally, the false-positive rate would be exactly the same for events with $\mu_{\rm rel}$ measurements. However, in most cases, one would receive a warning from the fact that this light-curve-based $\mu_{\rm rel}$ was inconsistent with the one derived from the late-time imaging. Gould (2022) estimated that if these cases were excluded (because they could be scheduled for additional late-time observations) then the false-positive rate could be reduced to about 3%. With this precaution, the overall false-positive rate would be about $(1/4) \times 10\% + (3/4) \times 3\% \sim 5\%$. If this is considered to be an acceptable level, then only about $(3/4) \times 7\% \sim 5\%$ of events would have to be subjected to additional late-time observations. Otherwise, a more aggressive approach would be needed.

Recently, Vandorou et al. (2023) have adopted the opposite approach. They imaged the planetary microlensing event OGLE-2016-BLG-1195 and reported a strong, 3σ, disagreement with the $\mu_{\rm rel}$ measurements from the discovery papers (Bond et al. 2017; Shvartzvald et al. 2017). None of the six previous late-time imaging efforts, OGLE-2005-BLG-071 (Bennett et al. 2020), OGLE-2005-BLG-169 (Batista et al. 2015; Bennett et al. 2015), MOA-2007-BLG-400 (Bhattacharya et al. 2021), MOA-2009-BLG-319 (Terry et al. 2021), OGLE-2012-BLG-0950 (Bhattacharya et al. 2018), and MOA-2013-BLG-220 (Vandorou et al. 2020), had yielded such a strong disagreement. Nevertheless, rather than regarding this disagreement as a warning sign that this was one of the $\sim 10\%$ expected rate of false positives, they assumed that the "other star" was indeed the lens, that the two original light-curve-based $\mu_{\rm rel}$ measurements were each in error by 3σ, and that these were superseded by their own measurement.

Here, we investigate these issues further. We note that the two original $\mu_{\rm rel}$ measurements were based on completely independent data sets and were consistent with each other at the 1σ level. We therefore combine these two measurements in two ways. First, we use the standard method of combining independent measurements, and, second, we combine the two data sets at the light-curve level. On this basis, we further refine the conflict between the light-curve and late-time-imaging determinations. We then propose a test to determine whether the Vandorou et al. (2023) identification is indeed a false positive.

2. $\mu_{\rm rel}$ Tension I: Comparison to Two Independent Light-curve Analyses

In their Table 2, Vandorou et al. (2023) reported the offset between the source and another star (which they identified as the lens) to be $\Delta\theta = 54.49 \pm 2.70$ mas at $t_0 + \Delta t$, where $\Delta t = 4.12$ yr and t_0 is the peak of the event, when the lens and source were separated by $\sim 15\,\mu$as, i.e., far too small to be of interest here. If the "other star" is indeed the lens, then (ignoring for the moment the lens–source relative parallactic motion—see the next paragraph), this would correspond to a heliocentric relative proper motion,

$$\mu_{\rm rel,hel} = \frac{\Delta\theta}{\Delta t} = 13.22 \pm 0.66 \text{ mas yr}^{-1}. \qquad (3)$$

Before continuing, we note that Vandorou et al. (2023) incorrectly report an error of $\sigma = 0.89$ mas yr^{-1} in their Table 2, evidently by adding in quadrature the errors of the two components of $\mu_{\rm rel,hel}$. During the 4.12 (equally, 0.12) yr of elapsed time, Earth moved approximately east by about 0.7 au, causing the lens to appear to move west by $\delta\theta = 0.7\,\pi_{\rm rel} \sim 90\,\mu$as$(\pi_{\rm rel}/130\,\mu$as$)$, where we have normalized to the Shvartzvald et al. (2017) value, which Vandorou et al. (2023) argue, based on their measurement, is way too large. Even if the Shvartzvald et al. (2017) value is correct, and taking account of the direction of the offset, $-32°$, north through east, this would affect the proper-motion determination by only about 0.01 mas yr^{-1}, substantially more than an order of magnitude below the measurement errors. Hence, this effect can safely be ignored.

We now argue that this measurement is in 4σ disagreement with the $\mu_{\rm rel}$ measurements made by Bond et al. (2017) and Shvartzvald et al. (2017), via the relation,

$$\mu_{\rm rel} = \frac{\theta_*}{t_*} = \frac{\theta_*}{\rho t_{\rm E}}, \qquad (4)$$

where θ_* is the angular radius of the source, $\rho = \theta_*/\theta_{\rm E}$, and t_* is the source self-crossing time. Substantially different methods are used to measure θ_* and t_*. Hence, we treat them separately. We note that these two measurements are virtually uncorrelated.

2.1. t_* Measurement

Bond et al. (2017) and Shvartzvald et al. (2017) analyzed disjoint photometric data sets, and each fit these to planetary models, which automatically yielded estimates of t_*. In principle, one might simply take these two measurements and combine them in the standard way to obtain the best overall estimate. However, for several reasons, we adopt a more comprehensive approach.

Our main concern is to check that the overall fits presented in these two papers are consistent. If they were not, it would be evidence that one or the other of these two measurements were dominated by systematics, in which case it would not be appropriate to combine them in the naive way. Second, each group reported multiple models, i.e., two and eight models, respectively. Both groups reported two classes of models, which both labeled "wide" and "close," but which are more accurately called "inner" and "outer" (Gaudi & Gould 1997; Yee et al. 2021). That is, in the inner model, the source passes inside the planetary caustic (over a ridge between the planetary and central caustics), while in the outer model, it passes outside the planetary wing of a central caustic. In both cases, these

Table 1
Combination of Two Fits Comparison

Parameter	B+	σ(B+)	S+	σ(S+)	Δ/σ	Comb.	σ(C)
$t_0 - 2{,}457{,}560$	8.7716	0.0020	8.7693	0.0013	0.9432	8.7700	0.0011
t_{eff} (day)	0.5225	0.0069	0.5296	0.0046	−0.8605	0.5274	0.0038
t_{E} (day)	10.1850	0.2550	9.9600	0.1100	0.8102	9.9953	0.1010
t_* (day)	0.0337	0.0022	0.0286	0.0037	1.1730	0.0324	0.0019
t_q (hr)	0.0103	0.0016	0.0133	0.0018	−1.2249	0.0117	0.0012
α (deg)	55.2850	0.2600	55.4988	0.1300	−0.7353	55.4560	0.1163
s_{inner}	1.0698	0.0078	1.0858	0.0078	−1.4505	1.0778	0.0055
s_{outer}	0.9957	0.0073	0.9839	0.0072	1.1508	0.9897	0.0051

Note. B+ = Bond et al. (2017); S+ = Shvartzvald et al. (2017); C = Combined.

models (or groups of models) are indistinguishable at the 1σ level. Therefore, we must determine whether these models make essentially identical predictions for t_*. Third, Shvartzvald et al. (2017) simultaneously analyzed Spitzer data and so reported measurements of the microlensing parallax vector π_{E}. Such measurements are subject to a well-known four-fold degeneracy (Refsdal 1966; Gould 1994), which is the reason that Shvartzvald et al. (2017) reported four times more models. Although we are not directly concerned with the π_{E} measurements in this section, we must determine whether including π_{E} in the fit substantially impacts the parameters that we are interested in.

To conduct these investigations, we first reparameterize the fits reported by the two papers in terms of "invariants." That is, one usually expresses the solutions to planetary microlensing events in terms of the seven parameters, $(t_0, u_0, t_{\mathrm{E}}, \rho, q, \alpha, s)$, where u_0 is the impact parameter (normalized to θ_{E}), α is the angle of the source trajectory relative to the planet–host axis, and s is the planet–host separation (also normalized to θ_{E}). However, Yee et al. (2012) showed that for planets detected at high magnification the four quantities $t_{\mathrm{eff}} \equiv u_0 t_{\mathrm{E}}$, $t_* \equiv \rho t_{\mathrm{E}}$, $t_q \equiv q t_{\mathrm{E}}$, and $f_S s t_{\mathrm{E}}$, are usually "invariants," i.e., the fractional errors in these quantities are smaller than those naively inferred from the fractional errors of the two factors because these are anticorrelated. Here, f_S is the source flux, although we will not be making use of the $f_S t_{\mathrm{E}}$ invariant until Section 2.2.

Thus, our seven parameters are $(t_0, t_{\mathrm{eff}}, t_{\mathrm{E}}, t_*, t_q, \alpha, s)$. We infer best estimates of the invariant parameters by taking the products of the two factors from the published tables. For the error estimates, we proceed as follows. First, for the cases that asymmetric errors are reported (all from Shvartzvald et al. 2017), we symmetrize them. Second, for an invariant parameter ηt_{E} (i.e., the product of two parameters from the fit, η and t_{E}), we adopt $\sigma(\eta t_{\mathrm{E}}) = \eta t_{\mathrm{E}}\sqrt{[\sigma(\eta)/\eta]^2 - [\sigma(t_{\mathrm{E}})/t_{\mathrm{E}}]^2}$. Note that Bond et al. (2017) already reported t_*, while Shvartzvald et al. (2017) reported ρ and t_{E} separately.

Our first step was to check whether the two models (or eight models) were consistent among themselves, i.e., they had essentially the same values and errors (except for s). Half of the eight Shvartzvald et al. (2017) models had negative t_{eff} and α, but we just considered the absolute values of these quantities for this purpose. We found that the error bars were all the same to within a few percent. Furthermore, we found that the largest difference between models was generally much smaller than the error bars. For Bond et al. (2017), the largest difference was 30% of the error bar (for t_0), while most were of order 20%. For Shvartzvald et al. (2017), the largest difference was 45% of the

error bar (for t_{E} for one solution), while the others were of order 10%–20%. Therefore, for each parameter, we report, in Table 1, the simple average of all solutions (two or eight) from each paper, except that we report the values of s from the two topologies separately. In column 6, we report the difference between the two papers divided by the quadrature sum of their errors. Under the assumption that the results reported from the two papers are not dominated by systematics, we expect these differences to be unit-variance Gaussian distributed. The first six parameters (i.e., excluding s_{inner} and s_{outer}) are essentially uncorrelated, implying that χ^2 is just the sum of the squares of the values in this column, i.e., $\chi^2 = 5.7$ for six degrees of freedom (dof). The last two rows are essentially uncorrelated with the others, but highly correlated with each other because $\sqrt{s_{\mathrm{inner}} s_{\mathrm{outer}}} \simeq s_+^{\dagger} \equiv (\sqrt{4 + u_{\mathrm{anom}}^2} + u_{\mathrm{anom}})/2$, where $u_{\mathrm{anom}} = t_{\mathrm{eff}} \csc \alpha / t_{\mathrm{E}} = 0.64\,\mathrm{day}/t_{\mathrm{E}}$ (Gould et al. 2022; Hwang et al. 2022; Ryu et al. 2022). If we add this seventh dof and evaluate it at the mean of their absolute values, i.e., 1.35, then $\chi^2 = 7.5$ for seven dof. In either case, this test constitutes strong evidence against systematics in either analysis.

Hence, it is justified to combine them, which is done in columns 7 and 8. In particular, we find,

$$t_* = 0.0324 \pm 0.0019\,\mathrm{day}, \tag{5}$$

i.e., a 5.9% error.

2.2. θ_* Measurement

As for the t_* measurement, the two papers relied on completely independent data sets for the measurement of θ_*, which provides a powerful consistency check. On the other hand, both papers employed the same overall method, which is subject to the same systematic errors. Therefore, the two θ_* measurements cannot simply be averaged together as was done for t_*, in Table 1.

The basic method (Yoo et al. 2004) is: first, measure the offset of the source relative to the red clump, $\Delta[(V-I), I] = [(V-I), I]_S - [(V-I), I]_{\mathrm{cl}}$, on a color–magnitude diagram (CMD); second, make use of the "known" dereddened position of the red clump, $[(V-I), I]_{\mathrm{cl},0}$ (Bensby et al. 2013; Nataf et al. 2013), to calculate the dereddened source values $[(V-I), I]_{\mathrm{s},0} = [(V-I), I]_{\mathrm{cl},0} + \Delta[(V-I), I]$; and third, use a color/surface-brightness relation to derive θ_* from $[(V-I), I]_{\mathrm{s},0}$. While the first step is a straightforward measurement with (usually) equally straightforward error estimation, the other two steps are subject to systematic errors that are more difficult to quantify.

Bond et al. (2017) and Shvartzvald et al. (2017) found $\Delta[(V - I), \quad I]_{B+2017} = (-0.355 \pm 0.021, \quad 3.369 \pm 0.018)$ and $\Delta[(V - I), \quad I]_{S+2017} = (-0.37 \pm 0.03, \ 3.40 \pm 0.04)$, respectively. Before continuing, we note that, of course, each group used its own measurement of I_S from its fit to the data. Because $f_S t_E$ is an invariant, and because the two fits differed by $t_{E,B+2017}/t_{E,S+2017} = 1.0226$, these I_S values would differ (after calibrating to the same CMD) by 0.024 mag. We put these on the same system by adopting the combined t_E from Table 1, which yields adjusted values $\Delta[(V - I), \quad I]_{B+2017} = (-0.355 \pm 0.021, \ 3.349 \pm 0.018)$ and $\Delta[(V - I), \quad I]_{S+2017} = (-0.37 \pm 0.03, 3.404 \pm 0.04)$.

In our view, the Bond et al. (2017) error bars are underestimated. First, they report the error in I_S itself as "0.001," although their reported error in t_E, combined with $f_S t_E$ invariance, implies that it is 0.027, which is substantially larger than their reported total error. In addition, based on our extensive experience, including making all of the CMDs used for the Bensby et al. (2013) calibrations, we do not believe that the clump can be centroided to the precision given by Bond et al. (2017) who quote $I_{clump} = 16.212 \pm 0.018$ and $(V - I)_{clump} = 2.468 \pm 0.007$. We note that Bond et al. (2017) also report a separate measurement based on the Optical Gravitational Lensing Experiment (OGLE)-IV data, $\Delta(V - I) = -0.39 \pm 0.03$. We then adopt an average of the two values for ΔI and of the three values for $\Delta(V - I)$, and we use the Shvartzvald et al. (2017) error bars, which basically reflect the difficulty of centroiding the clump,

$$\Delta[(V - I), I]_{adopted} = (-0.37 \pm 0.03, 3.38 \pm 0.04). \quad (6)$$

As mentioned above, the next two steps require estimates of the systematic errors. For this line of sight, the "known" position of the clump is $[(V - I), I]_{cl,0} = (1.06, 14.44)$. Based on our experience carrying out the Bensby et al. (2013) calibration, we estimate the clump color error as 1.06 ± 0.03. The Nataf et al. (2013) clump magnitude measurement, when combined with various stellar physics arguments, led to an estimate for a Galactocentric distance of $R_0 = 8.1$ kpc, in remarkable agreement with subsequent direct observations of SgrA*. Therefore, we estimate the systematic error in the clump magnitude as 14.44 ± 0.02, and so find,

$$[(V - I), I]_{S,0} = (0.69 \pm 0.04, 17.82 \pm 0.05). \quad (7)$$

As described above, the Yoo et al. (2004) method derives θ_* from the dereddened color and magnitude using a color/surface-brightness relation. Usually, one employs such a relation that is calibrated from angular diameter measurements. However, as the source has almost exactly the color of the Sun, we use its color and absolute magnitude, $[(V - I), M_I]_\odot = (0.71, 4.10)$, its radius $R_\odot = 695,700$ km, as well as the differential color relation $d \ln \theta_*/d(V - I) = 0.966$ quoted by Bond et al. (2017) to obtain $\theta_* = 0.822 \pm 0.037 \, \mu$as. Finally, we must account for the fact that the source star has an unknown composition and so may have a somewhat different surface brightness from the Sun at the same color. To account for this, we add 2% in quadrature to the error and finally obtain,

$$\theta_* = 0.822 \pm 0.041 \, \mu\text{as}. \quad (8)$$

For comparison, Bond et al. (2017) derived $\theta_* = 0.856 \pm 0.019 \, \mu$as, while Shvartzvald et al. (2017) derived $\theta_* = 0.82 \pm 0.07 \, \mu$as.

Combining Equations (5) and (8) yields,

$$\mu_{rel} = \frac{\theta_*}{t_*} = 9.27 \pm 0.72 \text{ mas yr}^{-1}. \quad (9)$$

Ignoring for the moment the difference between heliocentric and geocentric proper motions,[7] Equations (3) and (9) differ by 3.95 ± 0.98 mas yr^{-1}, i.e., a 4.0σ discrepancy.

3. Spitzer π_E

Gould (2022) also mentions contradictions between light-curve-based measurements of the microlens parallax π_E and the lens parameters derived from the AO imaging. The microlens parallax is defined by,

$$\pi_E \equiv \frac{\pi_{rel}}{\theta_E} \frac{\boldsymbol{\mu}_{rel}}{\mu_{rel}}. \quad (10)$$

In principle, it can be measured from light-curve distortions generated by Earth's annual motion (Gould 1992), but because for most events (and OGLE-2016-BLG-1195, in particular), $t_E \ll 1$ yr, this is often impossible, and it is difficult in most other cases.

Nevertheless, OGLE-2016-BLG-1195 was observed by Spitzer from solar orbit (Yee et al. 2015), and such observations can in principle measure π_E even for very short events (Refsdal 1966; Gould 1994). However, such satellite-parallax measurements are only straightforward if the satellite observations cover both the rising and falling sides of the light curve. Because of constraints on Spitzer operations, observations could not be immediately triggered. In particular, observations of OGLE-2016-BLG-1195 did not begin until 1.6 days after the ground-based peak. Moreover, the total flux variation was only about 2.5 flux units, whereas in several other cases, it has been shown that Spitzer light curves show systematics at the level of 0.5–1 flux unit. Hence, in our view, results derived from cases with few-flux-unit variations must be treated cautiously, but can still provide valuable information. For example, Spitzer observations of Kojima-1 (Zang et al. 2020) covered only the extreme falling wing of the light curve, with a flux variation of only 5 units, yet it delivered precise parallax information that has been independently confirmed by other techniques (Dong et al. 2019; Fukui et al. 2019).

However, rather than seeing the contradiction with previous π_E measurements as a reason for caution and deeper investigation of their own results, Vandorou et al. (2023) took this contradiction as "proof" that Spitzer π_E measurements derived from low flux variations are unreliable. By contrast, in our view, such a conclusion would only be appropriate if the AO results were independently confirmed in some other aspect, such as agreeing with the very precise, and multiple confirmed measurements of μ_{rel}. In fact, the radical disagreement of the AO-based and light-curve-based proper-motion measurements implies that the AO observations cannot be used to cast doubt on the Spitzer π_E. Thus, in the following section, we will retain an open mind regarding the Spitzer π_E measurement.

---

[7] Under the assumption that the AO measurement correctly gives the host–source relative proper motion, which is appropriate for this test, $\pi_{rel} = (\mu_{rel} t_E)^2/\kappa M \to 27 \, \mu$as. Then, according to Equation (11), below, the correction to the vector proper motion is $\Delta\boldsymbol{\mu}_{rel} = 0.17$ mas yr$^{-1} \boldsymbol{v}_{\oplus,\perp}/v_{\oplus,\perp}$, so the correction to the scalar proper motion is $\Delta\boldsymbol{\mu}_{rel} \cdot \boldsymbol{\mu}_{rel,hel}/\mu_{rel,hel} = -0.095$ mas yr^{-1}. Hence, it is justified to ignore this effect, but if we included it, it would increase the tension by about 0.1σ.

4. Other Scenarios

In general, when a planet/host mass measurement is made based on a single late-time AO observation, one should always consider the possibility that the "other star" is not the lens, but rather is a companion to either the lens or source or is an ambient star. Moreover, as pointed out by Gould (2022), these possibilities must be taken even more seriously when the AO measurement appears to contradict previously known facts about the event. The contradiction that is expected to be most frequent is between the AO-based $\mu_{\rm rel,hel}$ and the light-curve-based $\mu_{\rm rel}$. In the present case, this contradiction is quite severe. Moreover, as just discussed in Section 3, it is augmented by a conflict between the AO measurement and the Spitzer $\pi_{\rm E}$. See Figure 3 of Vandorou et al. (2023).

4.1. Companion to the Lens (Host)

The "other star" could be a companion to the lens (host). As we will soon show, such a companion must be separated from the lens by at least tens of $\theta_{\rm E}$ (so tens of astronomical units), and therefore it will be moving with nearly identical proper motion to the lens. Thus, it can be robustly predicted that a second epoch, which could be taken "now" (in 2023) will show a vector displacement from its 2020 position of about $(2023 - 2020)\mu_{\rm rel} \simeq 28$ mas. Such an observation would rule out a source companion (which would hardly move) and would render the explanation of an ambient star extremely unlikely. Hence, it would confirm the "lens-companion" hypothesis. Furthermore, by extrapolating the companion proper motion back to t_0 (in 2016) one would find the separation between the "lens" (planet host) and its companion in arcseconds.

However, such a measurement would not, in itself, tell us the mass or distance of the host or the planet. That is, the lens could be anywhere along the line of sight, and at each possible distance, the lens would have a mass that is consistent with the measured $\theta_{\rm E} = \sqrt{\kappa M \pi_{\rm rel}}$, and the companion would have a mass consistent with its measured flux. Nevertheless, such a measurement would tell us the future positions as functions of time of the lens relative to both the source and the companion, which would permit an informed decision on when the lens could be imaged (possibly with more advanced instruments). Even if the lens were dark, one could still determine the distance of the companion (and so the lens system) by multicolor, or possibly spectral observations of the companion. Combining this distance with the $\theta_{\rm E}$ measurement would then yield the host (and planet) mass. Thus, the first step is simply to obtain another AO epoch, which could be done immediately.

Next, is the lens-companion hypothesis consistent with the Spitzer $\pi_{\rm E}$ measurement being correct? Recall first that there were eight such $\pi_{\rm E}$ solutions. However, these come in four (inner/outer) pairs, whose $\pi_{\rm E}$ are nearly identical. Moreover, the scalar amplitudes are very similar among these four solutions: $\pi_{\rm E} = (0.437, 0.473, 0.482, 0.430)$, leading to very similar $\pi_{\rm rel} = (0.111, 0.120, 0.122, 0.109)$ mas. Hence, all solutions are at similar distances $D_L = 4.2$ kpc, and they also have similar lens masses $M_{\rm host} \sim 0.07\,M_\odot$. At this distance, the measured flux of the companion $K = 19.96$, together with the extinction $A_K = 0.24$ (Gonzalez et al. 2012) and the mass–luminosity relation of Benedict et al. (2016), yields a companion mass $M_{\rm comp} = 0.4\,M_\odot$, hence, a mass ratio[8] $Q = M_{\rm comp}/M_{\rm host} \simeq 6$.

---

[8] We use an upper case "Q" to distinguish this putative stellar companion from the planet, whose mass ratio is designated "q."

In principle, such a massive companion could be ruled out because it might predict a light-curve distortion near the peak of this relatively high-magnification event that is not seen. To determine whether this is the case, we predict the position of the lens relative to the source, using the measured $\mu_{\rm rel}$ and the four values of $\pi_{\rm rel}$, to obtain the corresponding $\mu_{\rm rel} = \mu_{\rm rel}\pi_{\rm E}/\pi_{\rm E}$, and finally convert to heliocentric,

$$\boldsymbol{\mu}_{\rm rel,hel} = \boldsymbol{\mu}_{\rm rel} + \frac{\pi_{\rm rel}}{\rm au}\boldsymbol{v}_{\oplus,\perp}, \qquad (11)$$

where $\boldsymbol{v}_{\oplus,\perp}(N, E) = (-0.76, +28.94)$ km s^{-1} is the velocity of Earth, projected on the sky at t_0.

These models are illustrated in Figure 1. The green arrows represent the four heliocentric proper motions, propagated over the 4.12 yr between t_0 and the Keck observations. The blue point is the measured position of the "other star" in 2020 relative to the source, which is at the origin. In these models the "other star" is assumed to be a companion to the host. The black points then represent the companion position relative to the source (and so the host) at t_0. The red points represent the position of the host relative to the source in 2020.

Note that for two of the solutions, the companion was separated from the source at t_0 by 65 or 70 mas, so if there had been an AO observation at that time, such models could have been confirmed or ruled out. However, to the best of our knowledge, no such observations were taken. There were AO observations taken in 2018 by Vandorou et al. (2023), when the companion would have been separated from the source by about 57 or 60 mas for these two models. These observations were of lower quality than the 2020 observations, so it is not clear whether they could have detected the companion at the position predicted by these two models (i.e., models 3 and 4 in Figure 1). In any case, there are no constraints on models 1 and 2.

Note that models (1, 2, 3, 4) predict separations between the lens companion and the source of about (56, 60, 76, 78) mas in mid 2023. Hence, all would be detectable under good conditions, while models 3 and 4 (the two solutions that could not have been probed by the 2018 observations) would be detectable even under moderately good conditions.

For all four solutions, the companion is separated from the lens by at least $\Delta\theta_{\rm comp} > 27$ mas, which corresponds to $s_{\rm comp} = \Delta\theta_{\rm comp}/\theta_{\rm E} > 106$. This would induce a Chang & Refsdal (1979, 1984) caustic of radius $w = 2Q/s_{\rm comp}^2 \sim 10^{-3}$, which is about 50 times smaller than the closest passage of the source, i.e., $u_0 \sim 0.05$. Hence, the companion would not have induced any noticeable effect on the light curve near peak.

4.2. Companion to the Source

The "other star" could be a companion to the source. If so, at $D_S \sim 8$ kpc and $A_K = 0.24$, it would have $M_K \simeq 5.2$ and so $M = 0.6\,M_\odot$, with projected separation $a_{\perp,S} \simeq 435$ au. According to Figure 7 from Duquennoy & Mayor (1991), the corresponding period, $\log(P/{\rm day}) \sim 6.9$, is within the broad peak of the distribution for companions to solar-mass stars. Similarly, companions of mass ratio ~ 0.6 are also relatively common according to their Table 7. Thus, there is no reason to discount this possibility. AO observations in 2023 could confirm this hypothesis, provided that they were taken under similarly good conditions as the 2020 observation because the "other star" would remain at $\Delta\theta = 54$ mas. Note that two of the four lens-companion models (under the assumption that the Spitzer $\pi_{\rm E}$ measurement is correct) make similar predictions for
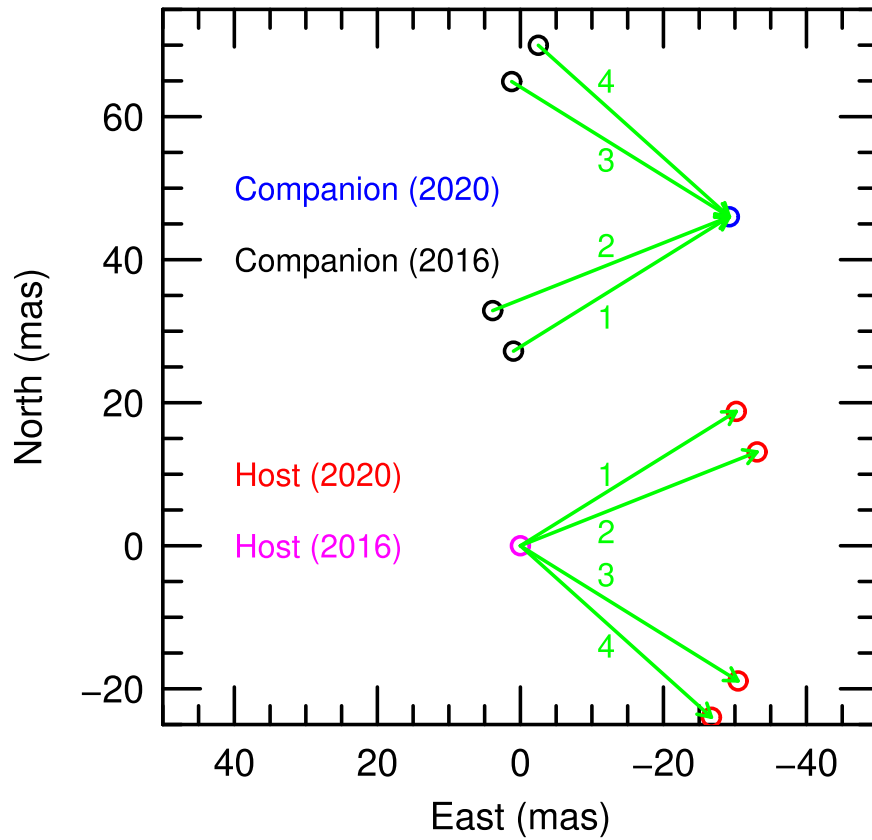
Figure 1. Illustration of four models (corresponding to four degenerate solutions for the microlens parallax π_E), in which the "other star" is assumed to be a companion to the host. The green arrows show the vector motion over 4.12 yr relative to the source (at the origin) for the companion (upper part of diagram) and host (lower part of diagram) for the four solutions, which are labeled sequentially by decreasing $\pi_{E,N}$. In all cases, the host is superposed with the source in 2016, while the companion is at the location measured by Vandorou et al. (2023) in 2020.

$\Delta\theta$ (see Section 4.1), but these have very different position angles (see Figure 1). Note also that even if the conditions are less than ideal, these observations would easily detect the "other star" at $\Delta\theta = 92$ mas under the Vandorou et al. (2023) hypothesis that the "other star" is the host. Note finally that if the "other star" is a companion to the source, it would lie well outside the Einstein ring and so would not be magnified at all during the event.

4.3. Ambient Star

The only other logical possibility (apart from host, companion to the host, and companion to the source) is that the "other star" is an ambient star that is unrelated to the event. In this case, the star would have $I_0 \sim 21.2$. The surface density of stars within a magnitude of this value toward Baade's Window is about 1000 arcmin^{-2} (Holtzman et al. 1998), while the surface density toward OGLE-2016-BLG-1195 is twice that of Baade's Window (Nataf et al. 2013). Therefore, the expected number of such stars within 54 mas is $p = 5 \times 10^{-3}$. While small, it is still much larger than the Gaussian probability that the light-curve-based proper-motion measurement is in error by 4σ ($p < 10^{-4}$). Hence, it must be considered.

An additional AO epoch of the "other star" could confirm the ambient-star hypothesis provided that the resulting vector-proper-motion measurement meets the following conditions: (1) it conflicts with the vector proper motion derived by Vandorou et al. (2023) under the assumption that this star is the host; (2) it also conflicts with the scalar proper motion derived

from the light curve (which should apply to either the host or a companion to the host); and (3) it is inconsistent with zero (as would be expected for a companion to the source).

5. Two Other Issues

There are two additional issues that impact the plausibility of the identification of the "other star" as the lens. While neither appears to be as severe as the 4σ discrepancy in the proper-motion measurement, both do need to be considered.

5.1. Galactic Kinematics

One reason for concern about the original Shvartzvald et al. (2017) Spitzer parallax measurement is that it appeared to imply that the kinematics of the lens are strongly at variance with what is expected for stars at ~ 4 kpc within the Galactic disk. That is, using our refined measurements of μ_{rel} and t_E (which are only slightly different from theirs) and their measurements of π_E, we find for the least retrograde solution ("solution 1" in Figure 1), that $\boldsymbol{\mu}_{rel,hel}(N, E) = (+4.56, -7.33)$ mas yr^{-1}, which corresponds to $\boldsymbol{\mu}_{rel,hel}(l, b) = (+0.25, +8.63)$ mas yr^{-1}.

For comparison, if the source had the mean motion of bulge stars as measured by Gaia (Gaia Collaboration et al. 2016, 2018) toward this direction, $\boldsymbol{\mu}_{S,hel}(l, b) = (-5.74, -0.13)$ mas yr^{-1}, this would imply lens motion, $\boldsymbol{\mu}_{L,hel} = (-5.49, +8.50)$ mas yr^{-1}. By contrast, taking account of the motion of the Sun relative to the LSR, and adopting an asymmetric drift at 4.2 kpc of -25 km s^{-1}, the mean motion of

a disk lens at this distance would be $\mu_{\rm hel,meanDisk}(l, b) = (-1.86, -0.35)$ mas yr^{-1}. Of course, neither the source nor the lens can be expected to be moving exactly at the mean motion of their respective populations, but the difference, $\Delta\mu_{L,\rm hel}(l, b) = \mu_{L,\rm hel} - \mu_{\rm hel,meanDisk}(l, b) = (-3.63, +8.85)$ mas yr^{-1}, i.e., a total of $|\Delta\mu_{L,\rm hel}| = 9.6$ mas yr^{-1}, would require peculiar motions of one or both of these stars that are relatively rare. That is, the dispersions of bulge sources are only $(3.0, 2.7)$ mas yr^{-1} in the (l, b) directions, while the dispersions of the disk lens at this distance are about $(3.1, 2.1)$ mas yr^{-1}. Hence, the offset comes to $(1.0, 2.6)\sigma$ in the two directions and thus a probability $p = 2\%$. The implausibility of this scenario was undoubtedly one of the motivations that led Vandorou et al. (2023) to begin observing OGLE-2016-BLG-1195 in 2018, when the lens and source were expected to be separated by only $\sim$18 mas, meaning that it was impossible to resolve them using Keck.

However, the identification of the "other star" as the lens does not resolve this motivating issue. Repeating the same steps as above, but assuming an asymmetric drift of -40 km s^{-1} (at $D_L \sim 7$ kpc), we obtain $\mu_{\rm rel,hel}(N, E) = (+11.16, -7.09)$ mas yr^{-1}, $\mu_{\rm rel,hel}(l, b) = (+6.16, +11.70)$ mas yr^{-1}, $\mu_{L,\rm hel} = (+0.42, +11.57)$ mas yr^{-1}, $\mu_{\rm hel,meanDisk}(l, b) = (-1.59, -0.21)$ mas yr^{-1}, $\Delta\mu_{L,\rm hel}(l, b) = (+2.01, +11.78)$ mas yr^{-1}, and $|\Delta\mu_{L,\rm hel}| = 11.95$ mas yr^{-1}. At this distance, the disk proper-motion dispersions are similar, so that the kinematics of this "solution" are intrinsically less probable, $p = 0.24\%$, than those of the "bad Spitzer measurement" that it was intended to solve.

5.2. Apparent Source Star Is Too Bright

Another issue is the flux from the apparent source star. That is, after Vandorou et al. (2023) subtracted the flux of the "other star," the remaining flux that they attribute to the source star ($K = 16.98 \pm 0.05$) is substantially higher than one would predict based on models of the microlensing event. Our best estimate of the intrinsic source color and magnitude (Equation (7)) combined with the color–color relations of Bessell & Brett (1988), would imply $K_{S,0} = 17.02 \pm 0.08$. Adopting $A_K = 0.24$ (Gonzalez et al. 2012), this implies $K_S = 17.26 \pm 0.08$. Hence, the apparent source is $\Delta K = 0.28 \pm 0.10$ mag brighter than expected.

Vandorou et al. (2023) do not comment on this discrepancy, but there are a number of possible explanations for it, some inconsequential but others that would substantially impact the interpretation of the lens system.

In the absence of any other information, by far the most likely explanation would be that the excess light is due to the lens. Indeed, this was the tentative conclusion of the Keck team when they first detected the excess light during the analysis of their 2018 observations using the NIRC2 camera on Keck (J.-P. Beaulieu 2019, private communication). These observations, made when the lens–source separation was only $\sim$20 mas, which could not possibly have resolved the lens from the source, regardless of its brightness. Hence, the main value of these observations would be to detect excess light, or at the least, serve as a first AO epoch that could help clarify later AO observations. Nevertheless, the Keck team ultimately adopted the cautious approach of waiting for confirmation by subsequent AO observations.

In fact, it is exactly their 2020 Keck OSIRIS observations that rule out the lens interpretation of this excess light: if the source/lens flux ratio really were 3:1, and if they were

separated by $\sim$38 mas (as predicted by the microlensing model, Equation (9)), then the lens would have been resolved by exactly the same technique as was used by Vandorou et al. (2023) to detect the much fainter object at 54 mas. For example, Bhattacharya et al. (2018) securely resolved the lens of OGLE-2012-BLG-0950 at $\Delta\theta \sim 34$ mas, with a flux ration 1.46:1. See their Figure 3.

Two other possibilities are that this excess light is due to a companion to the lens or to the source. Broadly, these scenarios are similar to those discussed in Sections 4.1 and 4.2, so we do not discuss them in detail here. The main difficulty is that, if one takes the 4σ proper-motion discrepancy at face value, then both this excess light and the "other star" reported by Vandorou et al. (2023) are due to companions, without the lens yet being detected. Nevertheless, this is certainly possible in principle.

Another possibility is that "3σ errors happen." That is, there is in reality no excess light: the problem is an incorrect measurement of the star's flux or a misestimate of the source color and magnitude from the analysis of the microlensing event. However, this 3σ error is occurring on the back of another, independent, 4σ error, which begins to strain credulity.

Finally, it is also possible that Vandorou et al. (2023) have identified the wrong star as the (previously) microlensed source, and the actual source is the neighbor that lies $\sim$175 mas to the northwest, at a position angle $\sim -31°$. This would be a very unsatisfying "solution" because it creates as many problems as it solves, but it does require investigation.

Under this hypothesis, the "other star" would almost certainly be a companion to its brighter neighbor at 54 mas, just as was considered in Section 4.2. Then the microlensed source would be one component of an asterism that is several tenths brighter than the star identified by Vandorou et al. (2023), i.e., with excess light almost equal to the microlensed source (instead of a 1:3 ratio). Thus, this scenario trades the problem of explaining a moderate amount of excess light for a larger amount of excess light. Again, in the context of this hypothesis, the excess light could be the lens or a companion to the lens or source. The one difference from the original scenario is that the lens explanation cannot (yet) be ruled out because no one has applied the techniques of Bhattacharya et al. (2018) and Vandorou et al. (2023) to this star.

Have Vandorou et al. (2023) misidentified the (formerly) microlensed source? We cannot say with certainty. They do not discuss how they made their identification, except to note that they had already done so based on their 2018 NIRC2 observations. And they certainly did so using image-level microlensing data that are not available to us. However, we have carried out our own determination using KMT image-level data, which we report in Appendix A. In brief, we find that in 2018 the "northwest star" (lying about 175 mas from the "southeast star" identified by Vandorou et al. 2023) lies roughly 30 ± 15 mas west of the microlensed source (as determined in 2016), while in 2020, it lies roughly 45 ± 15 mas west of the microlensed source. According to this analysis, the "northwest star" is a substantially better candidate for the source position.

6. $\mu_{\rm rel}$ Tension II: Reanalysis of the Original Light Curve

As discussed in Section 2, the first response to an apparent conflict between the light-curve-based measurement of $\mu_{\rm rel}$ and the AO-based measurement of $\mu_{\rm rel,hel}$ should be a review of the published literature to determine how secure this conflict

Table 2
Static 2L1S Models for OGLE-2016-BLG-1195 (OGLE+MOA+KMT-new)

Parameters	Central		Resonant	
	Central Inner	Central Outer	Resonant Inner	Resonant Outer
χ^2/dof	29,461.4/30,470	29,459.6/30,470	29,487.3/30,470	29,486.6/30,470
$t_0 - 7560$	8.7707 ± 0.0010	8.7706 ± 0.0009	8.7707 ± 0.0009	8.7707 ± 0.0009
u_0	0.0528 ± 0.0007	0.0528 ± 0.0007	0.0530 ± 0.0007	0.0531 ± 0.0007
t_E (days)	9.91 ± 0.11	9.91 ± 0.11	9.89 ± 0.11	9.88 ± 0.10
ρ (10^{-3})	3.20 ± 0.18	3.20 ± 0.17	3.52 ± 0.09	3.56 ± 0.10
α (degree)	55.25 ± 0.11	55.23 ± 0.11	55.28 ± 0.11	55.31 ± 0.11
s	1.0763 ± 0.0049	0.9911 ± 0.0044	1.0402 ± 0.0006	1.0254 ± 0.0007
q (10^{-5})	4.62 ± 0.44	4.60 ± 0.43	2.75 ± 0.13	2.77 ± 0.15
$\log q$	-4.337 ± 0.042	-4.339 ± 0.041	-4.561 ± 0.021	-4.559 ± 0.023
$t_{\rm eff}$ (days)	0.5237 ± 0.0015	0.5235 ± 0.0015	0.5241 ± 0.0016	0.5246 ± 0.0016
t_* (days)	0.0317 ± 0.0017	0.0317 ± 0.0016	0.0348 ± 0.0009	0.0352 ± 0.0009
t_q (hr)	0.0110 ± 0.0010	0.0110 ± 0.0010	0.0065 ± 0.0003	0.0066 ± 0.0003
$f_{S, \rm OGLE}$	0.2304 ± 0.0030	0.2304 ± 0.0030	0.2311 ± 0.0031	0.2315 ± 0.0029

actually is. We carried such a review and found the conflict to be at the 4.0σ level. The strength of this conflict provided the context for investigating other possible explanations for the detection of the "other star" in Sections 4 and 5.

However, another possibility is that the original light-curve analysis was incorrect and that a corrected value of $\mu_{\rm rel}$ might be more consistent with the AO-based $\mu_{\rm rel,hel}$. In Section 2, we argued that the original analyses were likely to be robust because there had been two such analyses that used independent data sets and independent codes, and we measured the difference between these as having χ^2/dof = 7.5/7.

Nevertheless, there are several reasons for pursuing this course. The most important is that the analysis of even seemingly "simple" planetary events like OGLE-2016-BLG-1195 has progressed substantially over the intervening 7 yr and continues to do so. In particular, new degeneracies have been discovered and new methods for finding degenerate solutions are being developed. As we will review below, one of these degeneracies can, in principle, lead to dramatic changes in t_* and therefore in $\mu_{\rm rel} = \theta_*/t_*$. Second, while seemingly unlikely, it remains possible that mistakes were made in the original analyses. Third, the disagreement between the two measurements of t_* was at the 1.2σ level, with the weighted average strongly dominated by the slower-$\mu_{\rm rel}$ result of Bond et al. (2017). Hence, if there were an error in that analysis, this could, by itself, significantly reduce the tension. Finally, there have recently been important improvements to the KMT data reductions, and these could in principle also change the result.

Moreover, Vandorou et al. (2020) charted a path of such light-curve reanalyses when they found that their AO observations conflicted with the upper limit on lens light derived from the original light-curve analysis by Yee et al. (2014) of MOA-2013-BLG-220.[9]

While we mainly report in this section on the results using rereduced KMT data, we note that we first checked on the

result of combining all the light curves (i.e., OGLE, MOA, and KMT) from the published literature (Bond et al. 2017; Shvartzvald et al. 2017). We found that the best-fit values of all eight parameters listed Table 1 were well within the 1σ ranges shown in the final two columns of that table. By far, the largest "discrepancy" was for t_0, which differed by $\sim 0.5\sigma$. In addition, we found that while the central value of $t_{\rm eff}$ was almost identical to that of Table 1, its error bar was only half as big.

As stated above, to conduct our reanalysis, we first rereduced the KMT data using an improved pipeline (H. Yang et al. 2023, in preparation) that is now routinely applied to essentially all new published KMT events. Inspection of the images showed that the KMTC01 data are affected by a "spike" from a neighboring bright star. Hence, we eliminated these from the analysis. Because KMTC does not cover the anomaly and because (by favorable chance) the event lies in two other high-cadence fields (BLG41 and BLG42), the elimination of these data is not expected to have a major effect.

The results of our analysis are given in Table 2, where we show both the standard parameters and the three invariant parameters ($t_{\rm eff}$, t_*, t_q) in order to enable comparison with Table 1.

There is one major and one minor feature of note about the new analysis. The major feature is that there are four solutions, including the two solutions found in the original papers (Bond et al. 2017; Shvartzvald et al. 2017), plus two additional solutions that were identified in our new grid search (Figure 2). In the original solutions, the bump peaking at 7569.13 in Figure 3 is a major-image anomaly generated by the source crossing a ridge that extends from the central (or resonant) caustic in the direction of the planet. In the additional solutions, the source crosses two closely spaced caustics in the planetary wing of a resonant caustic. See Figure 4. This degeneracy was discovered by Ryu et al. (2022) and was dubbed the "central-resonant" degeneracy by Yang et al. (2022).

Of particular relevance in the current context is that such a difference in morphologies can significantly impact t_*. This occurred for both of the cases in Ryu et al. (2022; i.e., KMT-2021-BLG-1391 and KMT-2021-BLG-0253), as well as for one of the two cases in Yang et al. (2022; i.e., KMT-2021-BLG-1689 but not KMT-2021-BLG-0171). As Ryu et al. (2022) recount, this degeneracy was discovered by a systematic
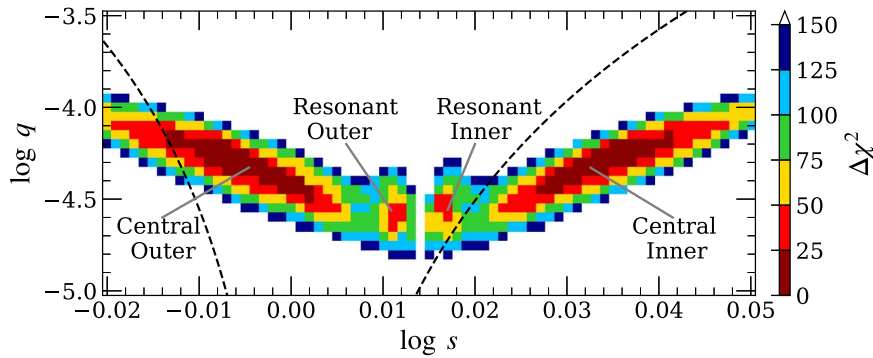
---

[9] In their initial arXiv preprint, their revised fit showed a reduction of the Einstein timescale by $\Delta \ln t_E = -0.147$, and corresponding increase in source flux by $\Delta \ln f_S = +0.144$. This would have substantially increased the tension by leaving even less of the baseline flux available for lens light. However, in the final published version of the paper, they corrected this fit and essentially reproduced the Yee et al. (2014) fit parameters. They attributed the modest remaining discrepancy to the effects of faint, unmodeled background stars. In any case, this appears to be the first effort to track down such discrepancies systematically by checking the original light-curve models.

Figure 2. Results of a "grid search" for solutions in which (s, q) are held fixed on a grid of values, while the remaining five parameters are allowed to vary. The two "central-caustic" solutions that were previously discovered (Bond et al. 2017; Shvartzvald et al. 2017) were recovered, but two additional "resonant-caustic" solutions were also discovered. Full analysis of the combined data set (Table 2) rules out these solutions. The boundaries between the central-caustic and resonant-caustic regimes are indicated by dashed lines.
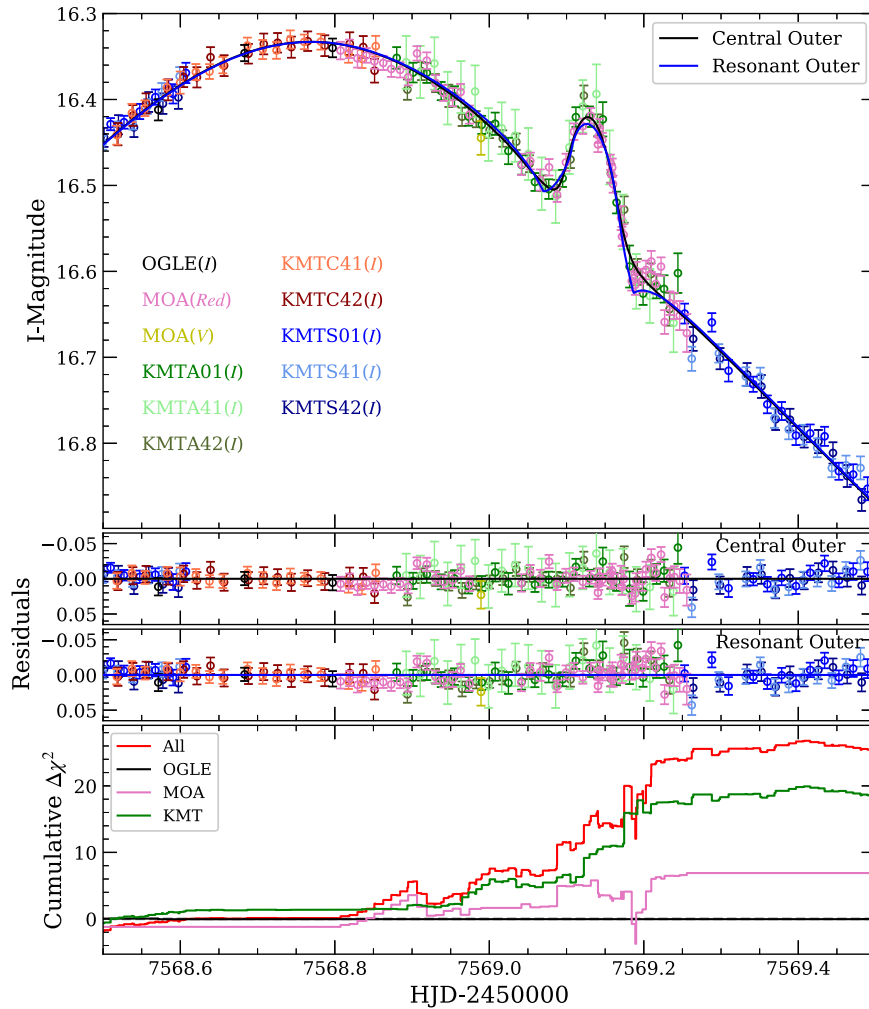
Figure 3. Data (colored points) and two representative models of OGLE-2016-BLG-1195, a central-caustic model (black) that was originally found in the discovery papers (Bond et al. 2017; Shvartzvald et al. 2017), and a new resonant-caustic model. The cumulative $\Delta\chi^2$ function shows that the new resonant model is strongly disfavored, primarily due to its failure to match the anomalous region, 7569.08 – 7569.22. However, rejection of this model would have been less decisive based on the partial data sets in each discovery paper. See also Figure 4.

effort to identify symmetries among multiple solutions. They hypothesized that because the degeneracy was found in two out of four randomly chosen events, it could be common and therefore should be checked for in archival events with bump-type anomalies at relatively high magnification. OGLE-2016-BLG-1195 is indeed a typical example of such an anomaly.

Table 2 shows that the additional solutions (labeled "resonant") indeed have significantly different values of t_* from the original solutions. Specifically, the values are $t_* = 0.0317 \pm 0.0017$ days (central) versus $t_* = 0.0350 \pm 0.0009$ days (resonant). This difference is far smaller than the factor ~ 2 difference found by Ryu et al. (2022) for their two cases. Moreover, in the present context,
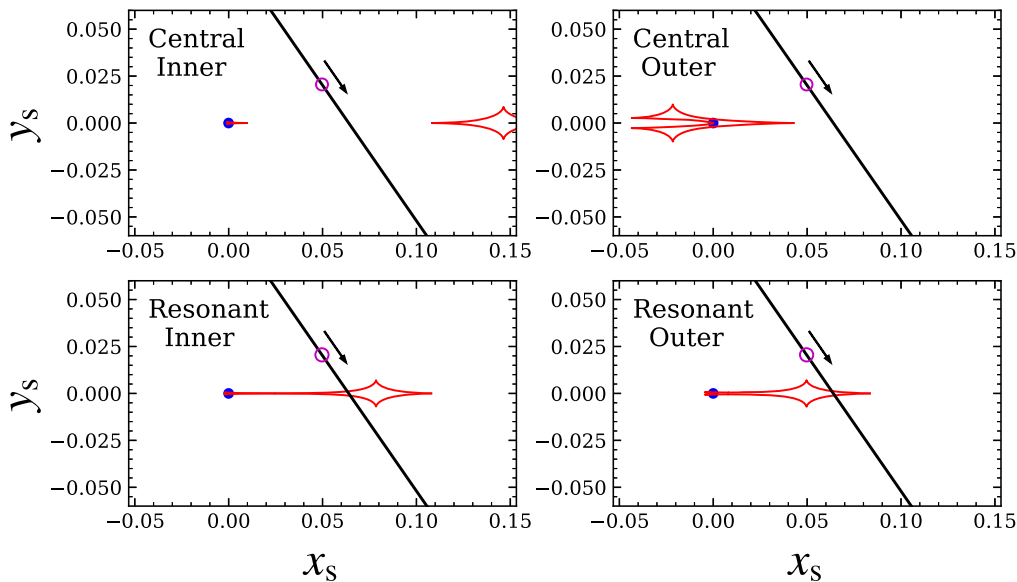
Figure 4. Caustic geometries for the four solutions shown in Table 2 and illustrated in Figure 3. In the "central-caustic" models, the "bump" in the light curve centered at 7569.13 is caused by the source passing over a ridge that extends from the central (or resonant) caustic, whereas in the "resonant-caustic" models, it is caused by the source passing over two closely spaced caustics of the planetary wing of a resonant caustic.

the new solutions have larger t_* (so smaller μ_{rel}), which would only strengthen the tension, not relieve it. Finally, the resonant solutions are disfavored in the combined analysis by $\Delta\chi^2 = 27$, which would normally lead to their being reported but discounted in the final assessment of the planet's characteristics. Nevertheless, because the χ^2 difference is divided between the two independent data sets (see Figure 3), it would have been substantially smaller in the original analyses had these additional solutions been recognized, and it could not have been so decisively rejected.[10] In particular, if the resonant solution had been recognized, it could not have been ruled out based solely on the $\Delta\chi^2 \sim 7$ difference from an analysis restricted to OGLE and MOA data. Nevertheless, this modest $\Delta\chi^2$ combined with the phase-space arguments of Yang et al. (2022), would (in a more modern context) probably be enough to reject this solution.

The minor feature is that the best-fit value of t_* is slightly smaller than the one in Table 1 and also has a smaller error bar: $t_* = 0.0317 \pm 0.0017$ days versus the previous value of $t_* = 0.0324 \pm 0.0019$ days. While this change is small compared to the errors, it still must be taken into account to make the best estimate of the discrepancy between the light-curve-based and AO-based proper-motion measurements.

The net result of this investigation is that the light-curve-based proper motion increased from $\mu_{rel} = 9.27 \pm 0.69$ mas yr^{-1} (Equation (9)) to $\mu_{rel} = 9.52 \pm 0.69$ mas yr^{-1}. Comparing to Equation (3) (and ignoring the correction from heliocentric to geocentric), this implies a conflict of 3.9σ rather than 4.0σ. Thus our light-curve reanalysis confirms the strong tension between the light-curve-based and AO-based proper-motion measurements.

7. Discussion

High-resolution imaging of the hosts of microlensing planets in which the host star is resolved from the source, which was pioneered by Batista et al. (2015) and Bennett et al. (2015) for Keck-AO and Hubble Space Telescope (HST) imaging, respectively, has immense scientific prospects. As demonstrated by Gould (2022), this approach can yield planetary and host masses for a statistically complete sample of well over 100 planets, beginning at first AO light on ELTs, probably about 2030. There is no other technique that can do this, in particular covering such a wide range of host and planet masses, projected separations, and Galactic environments.

However, as also shown by Gould (2022), only very fragmentary results are possible prior to ELT observations because only a small subset of the full sample will be accessible to present-day instruments. Hence, the main scientific return of high-resolution observations today is technical in nature, in effect, establishing and refining the viability of the technique. A good example is provided by Vandorou et al. (2023), who were able to detect and measure the 16:1 flux ratio of a neighbor at 54 mas, i.e., inside the Keck-AO FWHM. This, together with earlier related achievements by the same group, greatly increases confidence in the method, a confidence that will be sorely needed to gain the necessary observing time on ELTs in a highly competitive environment.

Nevertheless, as also systematically analyzed by Gould (2022), the technical issues arising from this method are not restricted to imaging technology: false positives and false negatives must be driven down to an acceptable level. Hence, the techniques for identifying these must not only be cataloged in theory, they must also be tested in practice.

In this paper, we have shown that the host identification reported by Vandorou et al. (2023) raises three different red flags. The most striking of these, and also the one that Gould (2022) argued would be the most common indication of a false

---

[10] Note that there is a plotting error in Figure 2 of Bond et al. (2017), which makes it appear as though the MOA data would favor the resonant-caustic models, whereas actually they disfavor these models. That is, there appears to be one prominent point that is below the displayed central-caustic model at 7569.08 and several more near 7569.22, i.e., exactly the places that the resonant-caustic model predicts dips relative to the central-caustic model. The most likely explanation is that, in making these plots, Bond et al. (2017) first calculated the flux residuals at time t_i and observatory j as $\Delta f_{i,j} = f_j(t_i) - (f_{S,j}A(t_i) + f_{B,j})$ and then plotted their data points at $A_j(t_i) = A(t_i) - \Delta f_{i,j}/f_{S,j}$. However, these should have been plotted at $A_j(t_i) = A(t_i) + \Delta f_{i,j}/f_{S,j}$. See the residuals panel in Figure 3 of the present paper for the correct plotting.

Table 3
20 Brightest KMTC01 Stars

Star	ΔE (″)	ΔN (″)	X (pixel)	Y (pixel)	$\sigma(X)$	$\sigma(Y)$	I	$V - I$	K_{pseudo}
1	3.85	−8.90	143.4649	128.0831	0.0094	0.0064	13.85	2.68	11.16
2	2.87	−10.66	145.9149	123.6762	0.0074	0.0123	14.64	3.26	11.38
3	1.44	5.08	149.4823	163.0342	0.0077	0.0124	15.55	2.82	12.73
4	−2.58	−8.52	159.5384	129.0413	0.0039	0.0009	15.61	3.10	12.51
5	−11.20	−4.03	181.0989	140.2636	0.0586	0.0264	15.89	2.46	13.43
6	9.34	−3.38	129.7343	141.8794	0.0020	0.0032	15.96	2.60	13.36
7	−11.34	4.35	181.4430	161.2108	0.0042	0.0074	15.98	2.60	13.38
8	2.63	4.38	146.5206	161.2719	0.0073	0.0170	16.04	1.81	14.23
9	7.01	−1.43	135.5583	146.7614	0.0044	0.0024	16.29	2.41	13.88
10	−5.31	1.37	166.3558	153.7593	0.0060	0.0065	16.34	2.45	13.89
11	3.69	6.95	143.8606	167.7049	0.0086	0.0165	16.59	2.27	14.32
12	0.17	1.74	152.6748	154.6741	0.0027	0.0155	16.66	2.35	14.30
13	4.75	6.03	141.2199	165.4033	0.0083	0.0267	16.72	2.52	14.21
14	−4.12	−4.27	163.3926	139.6519	0.0033	0.0066	17.04	2.85	14.19
15	−1.57	−1.13	157.0192	147.5050	0.0121	0.0160	17.21	2.45	14.77
16	−9.61	2.57	177.1237	156.7633	0.0388	0.0146	17.26	2.47	14.79
17	−4.49	7.99	164.3092	170.3038	0.0108	0.0095	17.31	1.85	15.46
18	9.71	−4.78	128.8218	138.3825	0.0485	0.0667	17.61	2.36	15.25
19	−8.00	−8.57	173.0839	128.8976	0.0022	0.0273	17.64	1.97	15.66
20	−2.77	6.73	160.0079	167.1636	0.0254	0.0040	17.75	2.64	15.11

Note. $(X, Y)_{source} = (153.0913, 150.3299) \pm (0.0116, 0.0141)$.

positive, is the strong conflict between the AO-based proper-motion measurement (Equation (3)) and the one derived from the published light-curve analyses (Equation (9)). The second red flag (argued by Gould 2022 to be less common) is the conflict with a previous microlens-parallax measurement. In addition to these two known red flags, we also found that the Vandorou et al. (2023) host identification leads to a solution that is kinematically disfavored relative to the one derived from the Spitzer microlens-parallax measurement.

None of these red flags proves that this host identification is incorrect. Together, however, they do imply that the result must be more deeply investigated, and if these investigations do not lead to clear rejection, then the identification must be confirmed by directly measuring the proper motion of the putative host relative to the source by additional late-time observations.

Prompted by these three red flags, we have conducted such an investigation based on existing data. We noted that the source appears to be "too bright" relative to the flux values implied by the light-curve analysis, something that was also previously noted by the Vandorou et al. (2023) authors but not reported in their published paper. We considered various explanations for this discrepancy, including one in which Vandorou et al. (2023) may have identified the wrong "star" (or stellar asterism) on which to conduct their analysis. We then found preliminary evidence in favor of this hypothesis.[11]

An important conclusion of Vandorou et al. (2023) was that by "disproving" the Spitzer parallax solution, they helped to demonstrate the general unreliability of Spitzer parallax measurements. While we do not agree that the Spitzer measurement has in fact been "disproved," we do agree that testing of Spitzer parallax measurements by high-resolution imaging is extremely important. Gould (2022) cataloged 12 planetary events with Spitzer parallaxes (his Table 3), of which

five have giant sources and lens–source relative proper-motion measurements (his Figure 4). Because giant sources can require much longer wait times, these potential targets may have to wait until well after ELT AO first light. In such cases, the Spitzer-based mass measurements will be essential to including these planets in statistical studies. Hence, testing the reliability of Spitzer-based parallaxes on events with fainter sources, e.g., OGLE-2016-BLG-1195, is crucial for establishing the conditions under which these giant-source planetary events can be included. We note that new Spitzer planets are still being discovered (so not yet cataloged by Gould 2022), including OGLE-2019-BLG-0679 (Jung et al. 2023) and OGLE-2017-BLG-1275 (Ryu et al. 2023), with the former having a giant source. Thus, one of the technical goals of current high-resolution studies should definitely be to test Spitzer parallaxes when feasible, in particular focusing on those that, like OGLE-2016-BLG-1195, do not have giant-star sources.

Appendix A
Position of the Source Star

We measure the position of the source relative to the frame of field stars in its neighborhood. Such a measurement can

---

[11] After this paper was submitted to AJ and posted to arXiv, Andrzej Udalski (2023, private communication) carried out an independent analysis of the source location using OGLE-IV data. This analysis confirmed that the microlensed star is indeed very close to the "northwest star" in Figures 6 and 7.
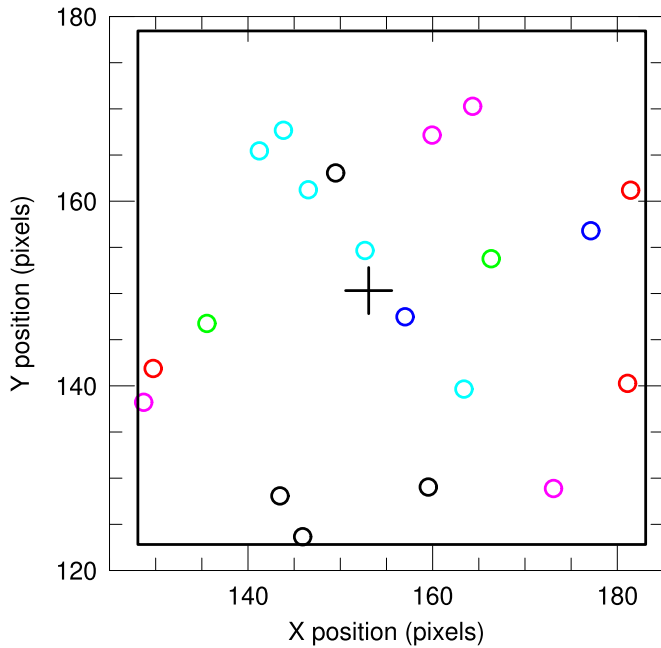
Figure 5. Positions of the 20 brightest stars (in I band) from the KMTC01 pyDIA reduction that overlap the Keck OSIRIS image (solid rectangle). The star positions are in the original reference-frame pixel coordinates. The pixels are approximately $0.4''$. For ease of comparison to the Keck images, the stars are color coded according to $K_{\mathrm{pseudo}} \equiv I - (V - I)$ by $K_{\mathrm{pseudo}} < (13, 13.5, 14, 14.5, 15, 15.5) \rightarrow$ (black, red, green, cyan, blue, magenta). The source position $(X, Y)_{\mathrm{source}} = (153.074, 150.317)$ is shown by a cross.

potentially have two different applications. First, it allows us to identify the star (or asterism) on the Keck-AO images that contains the source and therefore should be the target of detailed analysis. Second, it could also permit a measurement of the source's proper motion relative to this frame. Then, in turn, the frame proper motion could be put on an absolute scale using the Gaia sources within it. Combining such a measurement of $\mu_{S,\mathrm{hel}}$ with a measurement of $\mu_{\mathrm{rel,hel}}$ from the imaging would yield $\mu_{L,\mathrm{hel}} = \mu_{\mathrm{rel,hel}} + \mu_{S,\mathrm{hel}}$, which could help clarify the nature of the host.

These two potential applications have very different accuracy requirements. The surface density of stars (or asterisms) that are bright enough to contain the $K \sim 17.25$ source is relatively low. Hence, barring a pathological pileup of such stars, an accuracy of a few tens of milliarcseconds should be adequate. In fact, there are only two "stars" in the broad neighborhood of the source that are sufficiently bright to contain the source, and these are separated by ~ 175 mas. See Figure 1 of Vandorou et al. (2023). Hence, few-tens-of-milliarcseconds accuracy is indeed enough. On the other hand, in order to constrain μ_L significantly, the accuracy of μ_S should be $\lesssim \sigma_\mu \sim 3$ mas yr^{-1}, i.e., the known dispersion of bulge stars. Hence, the position measurement should have an accuracy $\lesssim \sigma_\mu \Delta t \sim 12$ mas yr^{-1}, which is more demanding.

The steps toward making this measurement are basically standard but still require some description. The first step is to measure the source position on the reference frame of the field stars in the pixel-coordinate frame of the camera. For this purpose, we start by making pyDIA (Albrow 2017) reductions of the KMTC01 images. The reference image is formed by stacking many good images. It is aligned to and then subtracted from a series of images in which the source is magnified. The resulting difference images then basically consist of an isolated

point-spread function (PSF), whose position is easy to measure.[12]

Next we repeat this procedure on three additional data sets: KMTC41, KMTS01, and KMTA01. We align the images by cross-matching relatively bright, $I < 17.5$ stars, making a two-dimensional linear (six parameter) transformation, and iteratively rejecting outliers. Then, for each cross-matched star (and also for the microlensed source position), we have four separate measurements, all in the KMTC01 reference system. From these we find the mean position, scatter (standard deviation), and standard error of the mean (s.e.m.). In particular, for the microlensed source, we find transformed positions (153.0735, 150.3169), (153.0633, 150.2900), (153.1118, 150.3588), and (153.1165, 150.3539) for KMTC01, KMTC41, KMTS01, and KMTA01, respectively. This yields a mean and s.e.m. of,

$$(X, Y)_{S,\mathrm{KMTC01}} = (153.0913, 150.3299)$$
$$\pm (0.0134, 0.0168) \text{ pixels}, \qquad (A1)$$

corresponding to errors of (5.4, 6.7) mas.[13] In Table 3, we list the pixel positions and instrumental magnitudes and colors of the 20 brightest stars lying in the region of the Keck OSIRIS image. We also list the pseudo-K-band magnitude, $K_{\mathrm{pseudo}} \equiv I - (V - I)$ as a rough indicator of the relative brightness expected in the K band.[14] These indicators are used in Figure 5 to color the points that represent these 20 stars, which should enable easy comparison to the Keck K-band images.

Using these, together with the microlensed source position in Equation (A1), it will be possible for any reader to make his/her own estimate of the source position on the Keck image. In addition, we provide the original KMTNet pyDIA CMD and star list files to enable complete reproduction of the analysis, if desired.

Next we proceed to make our own such measurements for the NIRC2 (2018) and OSIRIS (2020) epochs.

First, we retrieved the Keck images via the Keck Observatory Archive (KOA[15]). For the NIRC2 2018 epoch we use 13 good images. For the OSIRIS 2020 epoch we use 23 good images. For each image (in each epoch) we construct an astrometric catalog using PSF fitting for the centroids of all sources detected in the vicinity of the event, using the software package described in Ofek (2014) and E. O. Ofek et al. (2023, in preparation) and specifically the astrometric tools used in Ofek (2019). The (23 or 13) catalogs are then aligned using

---

[12] We note that under certain conditions, this method can fail, even catastrophically. However, in Appendix B, we derive a general formula that describes such potential failures and show that they do not apply to the present case.

[13] Note that we include KMTC01 in the pyDIA astrometric analysis even though it was excluded from the photometric analysis in Section 6. The astrometric analysis depends on a relative handful of images, and we check by eye that the contamination from the spike is low. Finally we check that in the above list of four measurements, KMTC01 contributes $\chi^2_{\mathrm{KMTC01}} = 0.8$, compared to an average value $\chi^2_{\mathrm{ave}} \equiv 1.5$, which is enforced by setting $\chi^2_{\mathrm{total}} \equiv \mathrm{dof} = 6$ for eight measurements.

[14] This is because, first, $(I - K)_0 \sim (V - I)_0$ over a broad range of stellar types (Bessell & Brett 1988). If it were also the case that $E(I - K) \sim E(V - I)$, then K_{pseudo} would be a very good proxy for K. In fact, for this line of sight, it is roughly true that $E(V - I) \sim E(I - K)$. Even for other lines of sight, the great majority of the stars in these images are in the bulge, and so suffer very similar extinction. To the extent that this is the case, the relative brightness in K is accurately predicted by K_{pseudo}, making it a valuable, even if imperfect, indicator of K.

[15] https://www2.keck.hawaii.edu/koa/public/koa.php

Table 4
13 KMT–Gaia Matches

Star	μ_α	μ_δ	RUWE	ΔE(OS)	ΔN(OS)	ΔE(N2)	ΔN(N2)
1	−1.36	−0.88	1.02				
	0.05	0.03					
2	−4.93	−5.97	0.81			2.5054	−10.6576
	0.11	0.07				0.0393	0.0336
3	−8.82	−3.92	1.19	1.4265	4.9511	1.4354	4.9147
	0.16	0.10		0.0005	0.0005	0.0328	0.0202
4	−3.45	−5.70	1.06	−2.7065	−8.2932	−2.7759	−8.3684
	0.12	0.07		0.0003	0.0008	0.0165	0.0222
6	2.33	−1.25	1.09	9.2114	−3.4319	9.2014	−3.4270
	0.15	0.11		0.0012	0.0005	0.0078	0.0043
7	−4.66	−10.95	1.01			−11.4346	4.4242
	0.15	0.10				0.0030	0.0020
9	−2.16	−7.47	1.20	6.9454	−1.5280	6.9400	−1.5243
	0.19	0.13		0.0004	0.0003	0.0014	0.0024
10	−5.20	0.82	1.19	−5.3049	1.4491	−5.3170	1.4445
	0.22	0.13		0.0004	0.0002	0.0016	0.0020
11	−4.76	−3.20	1.56	3.7042	6.7277	3.7029	6.7341
	0.28	0.18		0.0004	0.0003	0.0009	0.0022
12	−3.00	−5.05	1.18	0.2190	1.7090	0.2019	1.7090
	0.27	0.17		0.0004	0.0003	0.0034	0.0048
14	−8.23	−9.56	1.28	−4.1584	−4.1288	−4.1599	−4.1144
	0.36	0.22		0.0005	0.0002	0.0011	0.0013
15	0.42	−5.05	1.28	−1.5333	−1.1035	−1.5521	−1.0972
	0.36	0.23		0.0003	0.0002	0.0012	0.0015
17	2.08	−6.53	0.94	−4.4552	7.8875	−4.4669	7.8898
	0.26	0.15		0.0004	0.0004	0.0012	0.0015

Note. units: proper motions (mas yr^{-1}); offsets ($''$).

third-order polynomials. The mean position and standard deviation is then measured for each entry. We then set the zero-point of this catalog as the measured position of the "northwest star," which we will find below is the closest to the microlensed source. Note that the standard errors of the mean of this star are (0.10, 0.06) mas and (0.44, 2.23) mas, for OSIRIS and NIRC2, respectively, which are small compared to the other errors in the problem. We express these offsets in arcseconds. Note that in the KMT system, the first coordinate (in pixels) increases to the west, while in the Keck system, the first coordinate (in arcseconds) increases to the east.

Next, we identify a restricted subset of the KMT stars that have Gaia counterparts with proper-motion measurements. We exclude from consideration any star with KMT position errors (s.e.m.) greater than 10 mas. In fact, all but one of these excluded stars have Gaia entries but without proper-motion measurements, likely due to crowding and/or unresolved sources. For each such star, we find (if present) the OSIRIS and NIRC2 counterparts. Table 4 lists the resulting 13 stars. Column 1 is a cross reference to Table 3. Columns 2 and 3 give the Gaia proper motions as well as their errors (in the second row). Column 4 gives the Gaia RUWE indicator. Columns (5, 6) and (7, 8) give the OSIRIS and NIRC2 positions, respectively, with their standard deviations shown in the second row. Note that for four of the $2 \times 13 = 26$ cases, there is

no measurement (hence, no entry). Thus, there are potentially 10 cross-matches for OSIRIS and 12 for NIRC2. However, we find that the brightest three NIRC2 stars are seriously saturated, with bagel-morphology images. Corresponding to this, their standard deviations are dramatically larger than those of the fainter NIRC2 stars. Hence, we exclude these, leaving nine NIRC2 stars.

Next we propagate the KMT positions forward using the Gaia proper motions by 2.10 and 4.12 yr for NIRC2 and OSIRIS, respectively. For this purpose we subtract an estimate of the mean proper motion of the bulge, $\mu_{\rm bulge}(E, N) = (-3.00, -5.20)$ mas yr^{-1}, from each Gaia proper motion. That is, ultimately we will be measuring the offset between the positions of Keck stars in the 2018/2020 bulge frame from the position of the microlensed source in the 2016 bulge frame.

We multiply the reported Gaia errors by 1.5 to take account of the general difficulty of proper-motion measurements in the bulge and then add these in quadrature to the KMT position errors. We also include the s.e.m.s of the Keck measurements, but these are generally too small to matter.

We then carry out a second-order (quadratic) transformation from the KMT frame to the Keck frames and thus derive the positional offsets of the "northwest star" from the nominal position of the microlensed source. We find offsets ("northwest star" minus microlensed source) of $\Delta(E, N)_{2018} = (-26.8,$
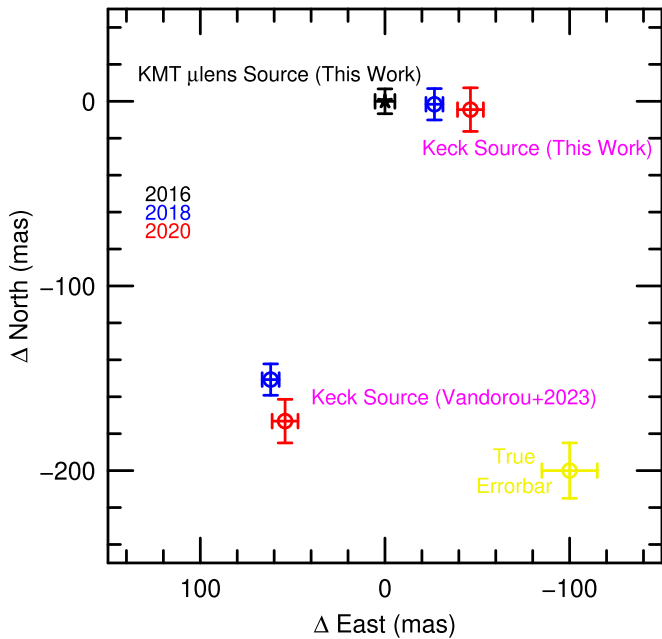
Figure 6. Offsets in the positions of two nearby stars relative to that of the microlensed source (black) as determined from a difference imaging analysis of KMT data from the microlensing event in 2016. The two stars lie in the northwest and southeast quadrants of the zoom of Figure 1 from Vandorou et al. (2023). The analysis of the present work yields the offsets of these two stars at two epochs, i.e., 2018 (blue, NIRC2) and 2020 (red, OSIRIS). The northwest star is much closer to the microlensed source, suggesting that the two are associated. By contrast, Vandorou et al. (2023) assumed that the southeast star was associated with the microlensing event. Note that the black error bars reflect the precision of centroiding the source within the KMT difference images, whereas the blue and red error bars reflect the precision of transforming from the Keck images to the KMT reference image. As discussed in the text, we estimate the true centroiding errors to be ~15 mas; for reference, we show ±15 mas error bars in yellow to the lower right of the plot.

$-1.6) \pm (1.2, \quad 2.2)$ mas with $\chi^2/\mathrm{dof} = 90/6$ and $\Delta(E, N)_{2020} = (-46.3, -4.5) \pm (1.9, 3.2)$ mas with $\chi^2/\mathrm{dof} = 108/8$, for NIRC2 and OSIRIS, respectively.

Clearly, our formalism has not captured all sources of error. One potential cause of additional errors is that the KMT and

Keck images have substantially different resolutions and bandpasses, so that each can be affected in different ways by stars that are uncataloged, either because they are too faint or are buried within the PSFs of cataloged stars. There could be others. We adopt the simple expedient of renormalizing the final error bars by $\sqrt{\chi^2/\mathrm{dof}}$ to obtain,

$$\Delta(E, N)_{2018} = (-26.8, -1.6) \pm (4.6, 8.5) \text{ mas},$$
$$\Delta(E, N)_{2020} = (-46.3, -4.5) \pm (7.0, 11.8) \text{ mas}. \quad (A2)$$

To get a further handle on the errors, we repeat these calculations by eliminating either the largest or two-largest outliers. In seven of the eight cases, (two epochs) × (two eliminations) × (two components), we find that the changes are <10 mas, while in one case (NIRC2, two eliminations, decl.), the change is 24 mas. Hence, we estimate that the true errors are of order 15 mas. Nevertheless, for purposes of display in Figure 6, we show the formal errors of Equation (A2). This figure also shows the microlensed source position and the position of the "southeast star," i.e., the one identified by Vandorou et al. (2023) as the microlensed source. Recall from the discussion above that the Keck star positions are in the bulge frame at their respective epochs, while the microlensed source position is in the bulge frame from 2016. Hence, one possible reason that the microlensed source is displaced from the northwest star is that the latter moved toward the west in the intervening years. However, the two Keck positions are actually consistent within the errors. Hence it is also possible that the northwest star has moved very little in the bulge frame, while the microlensed source has remained within the Keck PSF of this bright star.

Figure 7 shows the Keck images from the two epochs (NIRC2 2018 and OSIRIS 2020) with the (2016 bulge frame) position of the microlensed source superposed as a blue circle.

Currently, the balance of evidence favors that the microlensed source is associated with the northwest star, rather than the southeast star that was identified by Vandorou et al. (2023). However, in our view, it would be premature to claim this as a fact. First, the evidence that we have presented must be weighed against the evidence (based on completely independent microlensing survey images) that led Vandorou et al. (2023) to the conclusion that their southeast-star identification was correct.
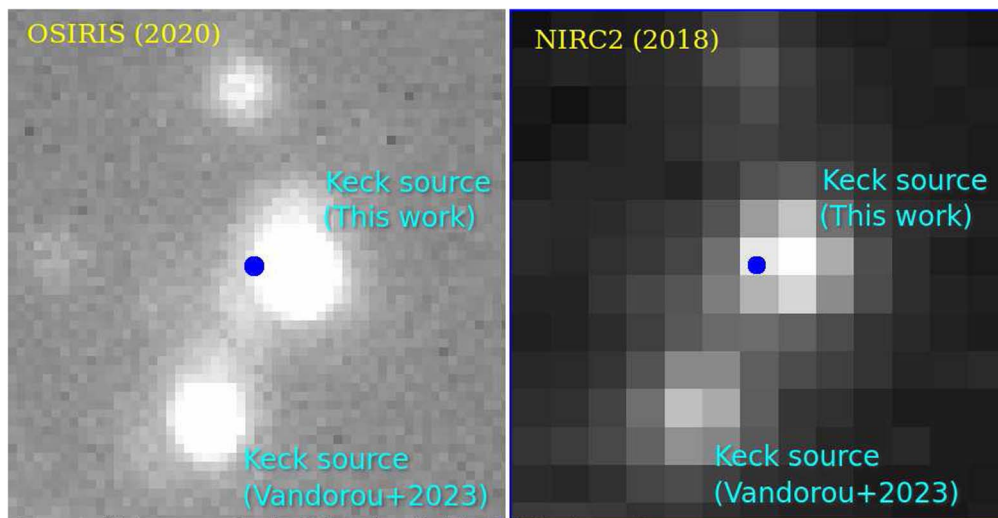

Figure 7. Keck images of the region surrounding OGLE-2016-BLG-1195 from OSIRIS 2020 (left) and NIRC2 2018 (right). In each case the position of the microlensed source (in the 2016 bulge frame) is superposed as a blue circle. The pixel sizes are, respectively, 10 and 40 mas. North is up and east is left. Vandorou et al. (2023) identified the southeast star (or asterism) as the location of the microlensing event, whereas our analysis suggests that this is more likely to be the northwest star.

Second, as we have discussed, the nonlinear transformation between the KMT and Keck frames are limited by the number of stars. In fact, we lacked enough stars even to make the standard third-order (cubic) transformation. One way to resolve this would be to take an HST image in the I band. This would require only a linear transformation. Moreover, it would allow direct comparison with the microlensing photometry (also in the I band), and so permit a more precise estimate of the excess flux.

However, perhaps the simplest approach would be to wait several more years and then take a Keck or HST image. For example, in 2025, the source and lens will be separated by ∼90 mas. If the Spitzer parallax is basically correct, then the host (near the star/BD boundary) will probably still not be visible. But if the host has a more typical mass, as argued by Vandorou et al. (2023), then it will be separately resolved in these images, probably near the northwest star, but likely visible in either case.

For the present, the main result of this appendix is to reinforce the need for a cautious approach to the identification of the host of OGLE-2016-BLG-1195.

Appendix B
Failure Mode of Difference-image Astrometry

The goal of difference-image astrometry is to locate the source position on the seeing-limited reference image, with the ultimate aim of comparing this reference-plus-source image to late-time high-resolution images. In the simplified presentation above, we implicitly assumed that nothing had moved between the epochs of the images entering the reference image and those of the magnified images. However, for a variety of reasons, including just convenience, the two may in principle be separated by some interval Δt, which could be either positive or negative. For completeness, we note that this could, in principle, lead to difficulties in aligning the reference and magnified images due to random motions relative to the bulk motion of the frame, but in practice this is almost never the case. Hence, we ignore this issue here. We further note that in most cases, the astrometric error of the source position is dominated by scintillation noise, rather than photon noise. However, this scintillation noise is automatically accounted for in our approach (which is often adopted) of estimating the astrometric errors from the scatter of multiple measurements.

After the magnified images (for simplicity, just called "image" I) are photometrically and astrometrically aligned to the reference image R, and then R is convolved to match the PSF of I, a difference image $D = I - R$ is formed by subtracting the second from the first. In the approximation that nothing has moved between the construction of R and I, $D = f_S(A - 1)P_S$, where f_S is the source flux, A is the magnification, and P_S is a normalized PSF whose centroid is at the location of the source θ_S. Then, in this approximation,

$$\theta_S = c(P_S) = \frac{c(D_S)}{(A - 1)f_S}, \quad (B1)$$

where c is the centroid operator that, in effect, just sums over the pixel values in P_S (of D).[16]

However, if either the microlensed source or the blended light, f_B, lying within the PSF (or both) have moved, then the subtraction operation will yield $P_{\text{diff}} = (I - R)/[f_S(A - 1)]$, whose centroid is displaced from θ_S by,

$$\Delta\theta = c(P_{\text{diff}}) - \theta_S = \frac{c(I - R)}{(A - 1)f_S} - \theta_S. \quad (B2)$$

In general, the blended light can take complex forms. However, as long as both the source and the components of the blended light remain well within the FWHM of the PSF, one can ignore these complexities. This is the case we consider here. It is not difficult, in principle, to consider blends that, e.g., straddle the PSF. However, such a more general treatment would take us too far afield. Then, Equation (B2) can be evaluated,

$$\Delta\theta = \frac{(f_S A\theta_{S,1} + f_B \theta_{B,1}) - (f_S \theta_{S,0} + f_B \theta_{B,0})}{(A - 1)f_S} - \theta_{S,0}, \quad (B3)$$

where the subscripts "0" and "1" refer to the reference epoch and the magnified epoch, respectively. Taking note that $\theta_{S,1} = \theta_{S,0} + \mu_S\Delta t$ and $\theta_{B,1} = \theta_{B,0} + \mu_B\Delta t$, Equation (B3) simplifies to,

$$\Delta\theta = \left[\frac{A}{A - 1}\mu_S + \frac{f_B}{(A - 1)f_S}\mu_B\right]\Delta t. \quad (B4)$$

The first point to note about Equation (B4) is that if Δt is small, then this effect is negligible. This is the case for our measurement because the reference image is constructed from epochs of the same year as the event and on either side of the peak. Specifically, the mean epoch of the reference image is displaced from t_0 by $\Delta t = -0.03$ yr. However, high-quality reference images are often used to carry out photometry for many years. This is feasible because the photometric error induced (for an isolated source and in the approximation of a Gaussian PSF) is just $\delta I = \log(32)(\delta\theta/\text{FWHM})^2$, which is tiny provided that the source has moved $\delta\theta = \mu_S\Delta t \ll \text{FWHM}$.

However, the astrometric effects can be much larger. For example, for very-low-magnification events, which includes some of the important class of giant-source free-floating planet (FFP) candidates (e.g., OGLE-2012-BLG-1323; Mróz et al. 2019), it is possible for $A/(A - 1) \gtrsim 10$, so that even with typical $\mu_S \sim 4$ mas yr^{-1} and a modest time offset from the reference image, $\Delta t \sim 5$ yr, the error in the source position could be $\Delta\theta_S \gtrsim 200$ mas, which could dramatically impact the science interpretation. See, e.g., Gould (2014). Similarly, such typical motions of a giant-star blend that was a factor 10 brighter than the magnified source could create similar artificial offsets. For example, such a blend could lie at 200 mas, so well inside the seeing-limited PSF, but easily resolved in AO images of next-generation ELTs.

In most cases, these problems can be avoided simply by constructing reference images from the same year as the peak magnification of the event. However, this may be difficult in some cases, while in others, the specific application might require even greater care to assure that Δt is as close to zero as possible.

ORCID iDs

Yossi Shvartzvald ⓘ https://orcid.org/0000-0003-1525-5041
Jennifer C. Yee ⓘ https://orcid.org/0000-0001-9481-7123
Sebastiano Calchi Novati ⓘ https://orcid.org/0000-0002-7669-1069

---

[16] In practice, the centroid is often found by maximizing the likelihood of a fit to D_S of models composed of a PSF at various positions plus a background term. Provided that the pixel counts are dominated by sources well within the PSF, this procedure has the same expectation but smaller errors. However, the process cannot be represented mathematically simply as the c operator in Equation (B1).

Weicheng Zang ⓘ https://orcid.org/0000-0001-6000-3463
Eran O. Ofek ⓘ https://orcid.org/0000-0002-6786-8774

References

Albrow, M. D. 2017, Michaeldalbrow/Pydia: InitialRelease On Github, v1.0.0, Zenodo, doi:10.5281/zenodo.268049
Batista, V., Beaulieu, J.-P., Bennett, D. P., et al. 2015, ApJ, 808, 170
Benedict, G. F., Henry, T. J., Franz, O. G., et al. 2016, AJ, 152, 141
Bennett, D. P., Bhattacharya, A., Anderson, J., et al. 2015, ApJ, 808, 169
Bennett, D. P., Bhattacharya, A., Beaulieu, J. P., et al. 2020, AJ, 159, 68
Bensby, T., Yee, J. C., Feltzing, S., et al. 2013, A&A, 549, A147
Bessell, M. S., & Brett, J. M. 1988, PASP, 100, 1134
Bhattacharya, A., Beaulieu, J.-P., Bennett, D. P., et al. 2018, AJ, 156, 289
Bhattacharya, A., Bennett, D. P., Beaulieu, J.-P., et al. 2021, AJ, 162, 60
Bond, I. A., Bennett, D. P., Sumi, T., et al. 2017, MNRAS, 469, 2434
Chang, K., & Refsdal, S. 1979, Natur, 282, 561
Chang, K., & Refsdal, S. 1984, A&A, 130, 157
Dong, S., Mérand, A., Delplancke-Strobale, F., et al. 2019, ApJ, 871, 70
Duquennoy, A., & Mayor, M. 1991, A&A, 248, 485
Fukui, A., Suzuki, D., Koshimoto, N., et al. 2019, AJ, 158, 206
Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2018, A&A, 616, A1
Gaia Collaboration, Prusti, T., de Bruijne, J. H. J., et al. 2016, A&A, 595, A1
Gaudi, B. S., & Gould, A. 1997, ApJ, 486, 85
Gonzalez, O. A., Rejkuba, M., Zoccali, M., et al. 2012, A&A, 543, A13
Gould, A. 1992, ApJ, 392, 442
Gould, A. 1994, ApJL, 421, L75
Gould, A. 2014, JKAS, 47, 279
Gould, A. 2022, arXiv:2209.12501
Gould, A., Han, C., & Zang, W. 2022, A&A, 664A, 13
Holtzman, J. A., Watson, A. M., Baum, W. A., et al. 1998, AJ, 115, 1946
Hwang, K.-H., Zang, W., Gould, A., et al. 2022, AJ, 163, 43
Jung, Y.K., Zang, W., Wang, H., et al. 2023, AJ, 165, 226
Kim, S.-L., Lee, C.-U., Park, B.-G., et al. 2016, JKAS, 49, 37
Mróz, P., Udalski, A., Bennett, D. P., et al. 2019, A&A, 622, A201
Nataf, D. M., Gould, A., Fouqué, P., et al. 2013, ApJ, 769, 88
Ofek, E. O. 2014, MAAT: MATLAB Astronomy and Astrophysics Toolbox, Astrophysics Source Code Library, ascl:1407.005
Ofek, E. O. 2019, PASP, 131, 054504
Refsdal, S. 1966, MNRAS, 134, 315
Ryu, Y.-H., Jung, Y. K., Yang, H., et al. 2022, AJ, 164, 180
Ryu, Y.-H., Udalski, A., & Yee, J.C. 2023, arXiv:2307.13359
Shvartzvald, Y., Yee, J. C., Calchi Novati, S., et al. 2017, ApJL, 840, L3
Terry, S. K., Bhattacharya, A., Bennett, D. P., et al. 2021, AJ, 161, 54
Vandorou, A., Bennett, D. P., Beaulieu, J.-P., et al. 2020, AJ, 160, 121
Vandorou, A., Dang, L., Bennett, D. P., et al. 2023, arXiv:2302.01168
Yang, H., Zang, W., Gould, A., et al. 2022, MNRAS, 516, 1894
Yee, J. C., Gould, A., Beichman, C., et al. 2015, ApJ, 810, 155
Yee, J. C., Han, C., Gould, A., et al. 2014, ApJ, 790, 14
Yee, J. C., Shvartzvald, Y., Gal-Yam, A., et al. 2012, ApJ, 755, 102
Yee, J. C., Zang, W., Udalski, A., et al. 2021, AJ, 162, 180
Yoo, J., DePoy, D. L., Gal-Yam, A., et al. 2004, ApJ, 603, 139
Zang, W., Dong, S., Gould, A., et al. 2020, ApJ, 897, 180