Microlensing Predictions: Impact of Galactic Disc Dynamical Models

Hongjing Yang, Shude Mao, Weicheng Zang, and Xiangyu Zhang

1 Department of Astronomy, Tsinghua University, Beijing 100084, China
2 National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100101, China

Accepted XXX. Received YYY; in original form ZZZ.

ABSTRACT
Galactic model plays an important role in the microlensing field, not only for analyses of individual events but also for statistics of the ensemble of events. However, the Galactic models used in the field varies, and some are unrealistically simplified. Here we tested three Galactic disc dynamic models, the first is a simple standard model that was widely used in this field, whereas the other two are more realistic taking into account the radial dependence of the velocity dispersion, and in the last model, the asymmetric drift. We found that for a typical lens mass $M_L = 0.5M_\odot$, the more accurate dynamical models predict $\sim 41\%$ or $\sim 27\%$ less long-timescale events ($t_\ell > 300$ days) and $\sim 1\%$ and $\sim 12\%$ more short-timescale events ($t_\ell < 3$ days) than the standard model. Moreover, large Einstein radius ($\theta_E > 1$ mas) and large microlensing parallax ($\pi_E > 0.3$) events are expected to be more common. The more advanced models also have an impact on the total microlensing event rate. This result will also to some degree affect the Bayesian analysis of individual events, so we recommend that modelers should be more careful when choosing the Galactic model. Additionally, we find the asymptotic power-law behaviors in both $\theta_E$ and $\pi_E$ distributions, and we provide a simple model to understand them.

Key words: gravitational lensing: micro − Galaxy: kinematics and dynamics − stars: black holes − planets and satellites: general

1 INTRODUCTION
Since the first detection of Gravitational microlensing event in 1993 (Alcock et al. 1993; Udalski et al. 1993), microlensing has become a powerful tool for probing various objects in the Milky Way, such as extrasolar planets beyond the snow line (Mao & Paczynski 1991; Gould & Loeb 1992), binary stars (e.g., An et al. 2002), isolated black holes (Gould 2000a; Mao et al. 2002; Bennett et al. 2002; Wyrzykowski et al. 2016).

However, the mass measurement of a microlens lens system is challenging. It requires two observables, i.e., any two quantities out of the angular Einstein radius $\theta_E$, the microlens parallax $\pi_E$ and the apparent brightness of the lens system (for more details, see the Introduction section in Zang et al. 2020). If $\theta_E$ and $\pi_E$ are measured, the mass ($M_L$) of the lens object can be uniquely derived from (Gould 2000b)

$$M_L = \frac{\theta_E}{\kappa \pi_E},$$

where $\kappa = 4G/(c^2 AU) \approx 8.144$ mas/M$_\odot$ is a constant, AU represents the astronomical unit. Unfortunately, $\theta_E$ and $\pi_E$ are both high-order parameters and usually unobservable, and measurements of the lens light require a $\geq 10$ years wait to resolve the lens and source by high-resolution images. This is, however, infeasible for dark lenses such as free-floating planets and black holes.

In most microlensing events, the only available parameter that can be retrieved directly from the observation is what is called the microlensing timescale,

$$t_E = \sqrt{\frac{4GM_L}{c^2} \frac{D_L(D_S - D_L)}{D_S} \frac{1}{V_T}},$$

where $V_T$ is the relative tangential velocity of the lens and source. Obviously, the physical parameters $M_L$, $D_L$, $D_S$ and $V_T$ are all degenerate. In this case, the only way to obtain these parameters for an individual event is Bayesian analysis (e.g., Beaulieu et al. 2006) and a Galactic model (GM) as the prior is essential for the analysis.

Since microlensing observations are independent of the lens light, it can detect dark objects like free-floating planets (FFP) and isolated neutron stars (NS) and black holes (BH). For free-floating planet candidates found by Sumi et al. (2011), $t_E$ is the only observable. Some other FFP candidates like OGLE-2016-BLG-1540 (Mróz et al. 2018), OGLE-2012-BLG-1323, and OGLE-2017-BLG-0560 (Mróz et al. 2019), as a result of the strong finite-source effect, their $\theta_E$ are also detectable, while $\pi_E$ is still unknown. These events are believed to be free-floating planets just because of their short timescales ($t_\ell < 2$ days) and/or small $\theta_E$ ($< 40$ mas), which lead to high Bayesian posterior probabilities that the lenses are low mass objects. For neutron star and black hole candidates (Wyrzykowski et al. 2016), $\pi_E$ can be obtained from the light curve thanks to their long timescale ($t_\ell \geq 100$ days) and the Earth motion, but $\theta_E$ remains unknown. The reason that the authors argue they are NS or BH candidates is still their Bayesian posterior. In these identification procedures, a Galactic model is essential to be utilized as prior knowledge.

In addition to individual events, Galactic model also plays a crucial role in microlensing predictions and statistics. For prediction or simulation works, Kiraga & Paczynski (1994) concludes that the microlensing event rate is $F \approx 12$ per year per $10^9$ stars monitored;
Gould (2000a) says about 20% microlensing events are caused by stellar remnants (white dwarfs, neutron stars and black holes); Han (2008) find that if the limiting magnitude is \( V_{\text{lim}} < 18 \), there will be \( \Gamma = 23 \) near-by (all-sky, without the Galactic bulge) microlensing events per year. For statistical works, Suzuki et al. (2018) shows that microlensing results challenge the core accretion planet formation theory. All these works are based or partly based on a Galactic model (whatever it is).

Therefore, the Galactic model is strongly connected with not only analysis of individual events but also prediction and statistics of ensemble events. However, up to date, the models used in this field are diverse, some are simple while some are complex. For example, in Zhu et al. (2017) (Zhu17), the Galactic model only contains two components, a disc and a bulge (bar). The dynamics is also simple, in the bulge the velocity of stars follows a three-dimensional Gaussian distribution, and in the disc, stars have a constant velocity dispersion from a constant rotation speed. In Han & Gould (1995) and Han & Gould (2003), a bulge and two disc components are considered, but the velocity distribution is similar to that in Zhu17 and is also unrelated to the position. In Bennett et al. (2014), they considered four Galactic components, a bar, a spheroidal stellar halo, a thin and a thick disc (Robin et al. 2003). Even though they considered seven stellar populations, the velocity distribution of each population is still Gaussian and similar to Zhu17 except they set a cutoff escape velocity. In Lam et al. (2020), they made use of Galaxia (Sharma et al. 2011) as their Galactic model. It is more complex and includes not only three disc components but also a spheroidal halo and a bulge. Their stellar dynamics contains many details in the real Milky Way, like radial dependence and asymmetric drift of the disc. For more information, see Table 1-3 in Sharma et al. (2011).

Despite the diverse Galactic model selections, there is little current work that directly compares different models and discusses how they affect the results. Koshimoto & Bennett (2019) and Shan et al. (2019) have done some comparisons, but the models they used are still somehow a combination of the previous simple models. Therefore it is still unknown to what degree simple models are effective and how complex models change or improve the results. In this paper, we introduce two simple but more realistic models, one is a modified Zhu17 model that considers the radial dependence of the velocity dispersions, and the other is based on the Shu distribution function (Shu 1969) which employs the asymmetric drift in the disc velocity dispersion. We compare these two models with the standard zhu model that considers the radial dependence of the velocity dispersion in the disc, the bulge the velocity of stars follows a three-dimensional Gaussian distribution, and the other is based on the Shu distribution function which employs the asymmetric drift in the disc velocity dispersion.

The structure of the paper is as follows: The models and some algorithm details are described in the next section. The simulation and Bayesian analysis results are presented in Section 3, and the discussion and conclusion are given in Section 4.

### 2 METHODS

#### 2.1 Background

For a microlensing event, the angular Einstein radius is (e.g., Gould 2000b; Mao 2012)

\[ \theta_E = \sqrt{\frac{\mu_L \pi_{\text{rel}}}{\kappa}}, \quad \pi_{\text{rel}} = A_U \left( \frac{1}{D_L} - \frac{1}{D_S} \right). \]

The microlensing timescale \( t_E \) is needed to cross an Einstein radius due to the relative motion between the lens and the source,

\[ t_E = \frac{\theta_E}{\mu_{\text{rel}}}, \]

where \( \mu_{\text{rel}} \) is the relative proper motion. The microlensing parallax is

\[ \pi_E = \frac{\pi_{\text{rel}}}{\theta_E}, \quad \pi_E = |\pi_E| = \frac{\theta_{E}}{\theta_E}, \]

where \( \pi_E \) is the direction of the relative proper motion between the lens and the source.

In a specific time interval and for a specific direction, the number of microlensing events (event rate \( \Gamma \)) is proportional to the angular area \( \pi \theta_E^2 \) and is inversely proportional to the time scale \( t_E \).

\[ \Gamma \propto \frac{\theta_E^2}{t_E} D_L^2 D_S^2 \gamma(D_S) = \theta_{E} \mu_{\text{rel}} D_L^2 D_S^2 \gamma(D_S). \]

Here \( D_L^2 \) and \( D_S^2 \) are factors that account for the increase of the number of source and lens stars with the distance, and \( \gamma(D_S) \) is a function that describes how the number of detectable sources varies from \( D_S \) (Kiraga & Paczynski 1994). All simulated events should be weighted by the event rate, for event \( i \) the weight is

\[ w_i = \Gamma_i \propto \theta_{E,i} \mu_{\text{rel},i} D_L^2 D_S^2 \gamma(D_{S,i}). \]

#### 2.2 Models

To generate a microlensing event, the information of the source \( (D_S, \mu_S) \) and the lens \( (D_L, \mu_L, \mu_f) \) is needed, where \( \mu_L \) and \( \mu_S \) are the proper motion of the lens and source, respectively. This information can be retrieved through a given Galactic model. A Galactic model can be described in three parts, the density profile, the velocity distribution, and the mass function (MF). The Galactic density model we adopt basically follows Zhu et al. (2017). The density profile includes two components, a bar (bulge) and a disc. The bar’s stellar number density follows (Kent et al. 1991; Dwek et al. 1995)

\[ n_{\text{B}} = n_{\text{B},0} e^{-\frac{1}{2} \left( \left( \frac{x'}{x_{\text{B}}} \right)^2 + \left( \frac{y'}{y_{\text{B}}} \right)^2 + \left( \frac{z'}{z_{\text{B}}} \right)^2 \right)}, \]

where \( x', y' \) and \( z' \) are the positions along the three major axes of the bar, and \( x_{\text{B}}, y_{\text{B}}, z_{\text{B}} \) are the scale length along each axis. The bar angle is taken to be 30° (Cao et al. 2013; Wegg & Gerhard 2013). All the parameters and their values are shown in Table 1. The disc’s stellar number density follows

\[ n_{\text{D}} = n_{\text{D},0} e^{-\frac{R - R_0}{R_d} - \frac{|z|}{z_d}}. \]

Here \( R \) and \( z \) are the galactic cylindrical coordinate components. \( R_d \) and \( z_d \) are the scale length along each axis. \( R_0 = 8.3 \text{ kpc} \) is the position of solar system (Gillessen et al. 2009). The density profile is axisymmetric since it does not depend on the azimuthal angle \( \phi \). The coefficient \( n_{\text{D},0} \) is determined by the local stellar density 0.14 stars/pc^3.

The velocity distributions are different for the two components. In the bulge, the velocity follows a three-dimensional Gaussian distribution,

\[ V_B \sim \exp \left( -\frac{V_x^2}{2\sigma_{B,x}^2} - \frac{V_y^2}{2\sigma_{B,y}^2} - \frac{V_z^2}{2\sigma_{B,z}^2} \right), \]

where the velocity dispersion \( (\sigma_{B,x}, \sigma_{B,y}, \sigma_{B,z}) = (120, 120, 120) \text{ km/s} \) (Zhu et al. 2017).

For the velocity distribution of the disc, we introduce three models listed below.

1) The standard model: This model is the same as the original Zhu et al. (2017) model except for some updates in parameter values.

MNras 000, 1–12 (2020)
velocity distribution of disc stars at any distance follows a Gaussian distribution,

\[ V_\phi \sim N(\overline{V}_\phi, \sigma_{D, \phi}^2), \quad V_z \sim N(\overline{V}_z, \sigma_{D, z}^2), \]

where the velocity dispersions \( \sigma_{D, \phi} \) and \( \sigma_{D, z} \) are constant at any location, and the values are taken to be solar neighborhood values, \((\overline{V}_\phi, \overline{V}_z) = (220, 0) \) km/s and \((\sigma_{D, \phi}, \sigma_{D, z}) = (33, 18) \) km/s.

2) Model B: The mean velocity is still \( \overline{V}_\phi = 220 \) km/s and \( \overline{V}_z = 0 \) km/s anywhere, but the velocity dispersion is a function of the distance from Galactic center \( R \),

\[ \sigma_{D, \phi}(R), \sigma_{D, z}(R) \propto \sqrt{n_D(R)}, \]

or

\[ \sigma_{D, \phi}(R)(\phi, z)(R) = \sigma_{D, \phi}(R_0)e^{-\frac{R-R_0}{\mp\sigma_{D, \phi}}}, \]

where \( R_0 = 8.3 \) kpc is the position of Sun, and the solar neighborhood values are \( \sigma_{D, \phi}(R_0) = 33 \) km/s and \( \sigma_{D, z}(R_0) = 18 \) km/s.

3) Model C: We use the dynamic model of Shu (1969). The disc model is more realistic with both radial dependence and asymmetric drift. We use software gaipy \(^1\) (Bovy 2015) to generate the \( V_\phi \) distribution function. The \( V_\phi \) distribution as a function of \( R \) is shown in Fig. 1. The velocity distribution along the \( z \) axis is the same as that in Model B. Note that we only use the \( V_\phi \) velocity distribution of Shu’s model.

The standard model is widely used in the microlensing field since its good simplicity, however, it is unrealistic and dynamically self-inconsistent. Model B and C introduced radial dependence and asymmetric drift to the Galactic model. Even though the two advanced models are still simple, they are sufficient for us to study the impact of different model selections.

For the mass function, here we first fix \( M_L = 0.5M_\odot \) for all models. From Eqs. 3 and 4 we know that a specific \( M_L \) value is just a coefficient of \( \log t_E \) or a translation term of \( \log t_E \left( M_L \rightarrow M_L' \right) \),

\[ t_E' \rightarrow \sqrt{\frac{M'_L}{M_L}}t_E, \text{ or } \log t_E' \rightarrow \log t_E + \frac{1}{2} \log \left( \frac{M'_L}{M_L} \right). \]

A continuous \( M_L \) distribution will broaden the \( t_E \) distribution, but the effect is the same for all models, and will not change the asymptotic behavior at \( t_E \gg 10 \) d and \( t_E \ll 200 \) d. The nature of the velocity distribution difference would be buried if we use a continuous MF, therefore in order to control the variables and focus on the influence of model difference, we choose a \( \delta \) MF in the first part (Section 3.1). After that, when doing the Bayesian analysis (Section 3.3.1), \( M_L \) is also a parameter that needs to be obtained, and the mass function must be specified as a prior. We will describe our choice of the mass function at that time.

2.3 Algorithm

For specific values, we assume that the observation is toward the center of Baade’s Window (R.A., Dec.):2000 = \((18^h0^m32.14^s, -30^\circ0'26.96''\)) or \((l, b) = (1.02^\circ, -3.92^\circ)\), and we adopt \( \gamma(D_S) = D_S^{-2} \) here (Kiraga & Paczynski 1994). Baade’s Window is near the Galactic center, so the line of sight is almost on the disc plane, therefore in Galactic coordinates, we take

\[ V_l \approx V_\phi, \quad V_b \approx V_z. \]  

We can easily covert the generated velocity to observed relative angular velocities by

\[ \mu_L = \frac{V_L - V_\odot}{D_L}, \quad \mu_S = \frac{V_S - V_\odot}{D_S}, \]

and

\[ \mu_{rel} = \mu_L - \mu_S, \quad \mu_{rel} = |\mu_{rel}|. \]

Here \( V_L = V_L^i e^i + V_L^b e^b \) and \( V_S = V_S^i e^i + V_S^b e^b \) are the lens and source velocities, respectively, and \( V_\odot \) is the velocity of the sun.

In practice, we first randomly generate a pair of lens and source position \( D_L \) and \( D_S \) from the stellar density profile. To make the computation more efficient, we adjust the stellar density profile to

\[ n'_S \propto (n_S + n_D) D_S^2 \gamma(D_S), \quad n'_L \propto (n_B + n_D) D_L^2 \gamma(D_L), \]

where \( n'_S \) and \( n'_L \) represent the adjusted source and the lens “number density”. Using this new density profile to generate \( D_L \) and \( D_S \) can avoid a large number of very low weight (e.g., small \( D_L \) events) and improve the computing efficiency.

Second, we generate their velocities according to a given Galactic model’s velocity distribution. For the standard model, the generation of velocity is straightforward. For Model B, we calculate the velocity dispersion \( \sigma_{\phi}(R) \) and \( \sigma_{\phi}(R) \) every 0.1 kpc, and use the nearest value of each lens or source for velocity generation. For Model C, the random number generator in gaipy is time consuming, so we calculated a discrete sampled \( v_\phi \) distribution PDF at specific \( R \) every 0.1 kpc, and then linearly interpolated the results. The sample interval is \( V_\phi/200 \) or 1.1 km/s, it is precise enough to show the difference between these models.

Then, we calculate the timescale \( t_E \) and the weight through Eqs. 3 and 4. Note that we have already taken \( D_L^2 \gamma(D_L) \) term when generating their distance, so the weight of an event \( i \) is now simply

\[ w_i = \theta_{\mu_e, i} \mu_{rel,i}. \]

Finally, we repeat the process (which can easily be parallelized) to simulate a large number of events to obtain the statistical properties that we are interested in.

3 RESULTS

3.1 Timescale Distribution

We generated \( \sim 4 \times 10^9 \) events for each model. Then the results are binned and normalized in \( \log t_E \) space by their weight to obtain the \( t_E \) distribution.

The first row of Fig. 2 shows the normalized \( t_E \) distribution of each model, the different colours represent different components. Obviously, each Galactic component contributes differently to the distribution. For the standard model, the disc-source disc-lens (D-D) distribution shows a significant double bump feature, and the reason is that the rotation direction of the disc changes at the two sides of the Galactic center. For Models B and C, the rotation direction changes as well, but the second bump is smoothed by their large velocity dispersions near the Galactic center.

The three models also share some common features. Short and long \( t_E \) are dominated by bulge-source bulge-lens (B-B) and bulge-source disc-lens (B-D) events, respectively. Short events are dominated by those where lens and source are close, whereas long events are dominated by those who have slow relative proper motions. The former mostly exists in the high-density bulge region, while the latter mostly exists in the dynamically colder Galactic disc. Since the event rate in a given time interval is proportional to \( t_E^{-1} \), short \( t_E \) bulge-source

\(^1\)http://github.com/jobovy/galpy
events contribute more than 80% of all microlensing events in all of the three models. This is consistent with the estimation of Kiraga & Paczynski (1994).

The comparisons of different models are shown in Fig. 3. Notice that the total event rate is not the same, so in this figure only the standard model is normalized to unity, while other models are normalized to the standard model. The total event rate ratio is \(1 : 1.041 : 1.015\). The detailed event rate comparison is shown in the figure and Table 2. Because all the \(e_{\text{f}}\) distribution follows \(P(\leq e_{\text{f}}) \propto e_{\text{f}}^2\) at the small \(e_{\text{f}}\) tail and \(P(> e_{\text{f}}) \propto e_{\text{f}}^{-3}\) at the long \(e_{\text{f}}\) tail (Mao & Paczynski 1996), the ratios of the three models on the log scale tend to be constant at both small and large \(e_{\text{f}}\). This can be seen as follows: if the \(e_{\text{f}}\) distribution for model \(m\) is a power-law, for example, the \(e_{\text{f}} \to \infty\) part that \(p_m(\log e_{\text{f}}) = A_m e_{\text{f}}^{-3}\), the fraction of \(> e_{\text{f}}\) events will be

\[
\int_{e_{\text{f}}}^{\infty} p_m(\log e_{\text{f}}) \, \text{d} \log e_{\text{f}} = \frac{\ln 10}{3} A_m e_{\text{f}}^{-3}.
\]

So \(p_1(> e_{\text{f}}) : p_2(> e_{\text{f}}) = p_1(\log e_{\text{f}}) : p_2(\log e_{\text{f}}) = A_1 : A_2\), we can read the event rate ratio between two models directly from the log-plot. The relation is valid on both small and large \(e_{\text{f}}\) power-law tails.

At the \(e_{\text{f}} > 300\) days tail, the two more realistic models only produce \( \sim 59\% \) and \( \sim 73\% \) events of the standard model. In other words, if an individual event shows long \(e_{\text{f}}\), the lens would be more likely to be a massive object, because in these models, low-mass objects cannot produce as many long \(e_{\text{f}}\) events as in the standard model. In addition, at the \(e_{\text{f}} < 3\) days tail, Model B and Model C produce a little more \(\sim 1\% \) and \(12\%\) events than the standard model. However, the difference may not be significant for the current event database since these extreme events only accounted for less than 1%.

### 3.2 \(\theta_E\) and \(\pi_E\) Distribution

From Eqs. 3 and 5, we can see that \(\theta_E\) and \(\pi_E\) are both a function of \(M_L, D_L\) and \(D_S\) but unrelated to the velocity distribution. However, the event rate (Eq. 6) depends on the velocity distribution, so under our three models, the event rate distribution as a function of \(\theta_E\) and \(\pi_E\) would also be different.

In Fig. 2, the second and third rows show the \(\theta_E\) and \(\pi_E\) distribution, respectively. The shapes of \(\theta_E\) and \(\pi_E\) distributions are similar (actually, they are strictly the same) simply because \(\theta_E = \kappa M_L \pi_E\) (Eq. 1) and we fix the \(M_L\) in the simulation. In all models, small \(\theta_E\) and small \(\pi_E\) are dominated by bulge-source bulge-lens (B-B) events since these events have small \(\pi_{\text{rel}}\) and large \(\theta_E\) and large \(\pi_E\) are dominated by bulge-source disc-lens (B-D) events because of their large \(\pi_{\text{rel}}\). For Models B and C, the disc is dynamically warmer than that in the standard model, leading to a higher probability of

### Table 1. Parameters adopted in the simulation

| Parameter        | Value       | Unit       | Remarks                  | Reference     |
|------------------|-------------|------------|--------------------------|---------------|
| \(R_0\)          | 8.3         | kpc        | Distance to Galactic center | Gillessen et al. (2009) |
| \(\sigma_{\text{bar}}\) | 30          | deg        | Bar angle                | Cao et al. (2013); Wegg & Gerhard (2013) |
| \((x_0, y_0, z_0)\) | (1.59, 0.424, 0.424) | kpc | Bar scale length along each axis | Robin et al. (2003) |
| \(n_{B,0}\)      | 13.7        | stars/kpc³ | Bulge stellar density coefficient | Robin et al. (2003) |
| \((\sigma_{B,x}, \sigma_{B,y}, \sigma_{B,z})\) | (120, 120, 120) | km/s | Bulge velocity dispersion | Zhu et al. (2017) |
| \((R_d, z_d)\)    | (2.5, 0.325) | kpc | Disc scale length and scale height | Binney & Tremaine (2008) |
| \(V_c\) or \(V_o\) | 220         | km/s       | Disc mean rotation velocity | Binney & Tremaine (2008) |
| \(V_{\phi}(R_0)\) | 0           | km/s       | Disc mean perpendicular velocity |                |
| \(\sigma_{D,\phi}(R_0)\) | 33          | km/s       | Solar neighbor velocity dispersion |                |
| \(\sigma_{D,z}(R_0)\) | 18          | km/s       | Solar neighbor velocity dispersion |                |
| \((V_{D}, V_{c,L})\) | (V_c + 12, 7) | km/s | Motion of solar system | Schönrich et al. (2010) |
| \(D_{SB}\)       | [5.6, 11.0] | kpc        | Distance range of bulge component |                |
| \(D_{S,L}^m, D_{L,D}^m\) | [0.1, 16.6] | kpc        | Distance range of disc component |                |
| \(\mathbb{M}_L\)  | 0.5         | M⊙         | Lens mass                |                |
| \(\gamma\)       |             |            | The factor of detectable source number when varying the distance |                |

**Figure 1.** The \(V_\phi\) distribution of three models at a given distance \(R\) from the Galactic center. Notice that the local rotation speed is fixed at 220 km/s and the Sun is at \(R = 8.3\) kpc.
B-D as well as D-D events. The comparisons of the three models are shown in Fig. 4. B-B events are the same in these three models, so the distribution of $\theta_E < 0.2$ mas and $\pi_E < 0.06$ are mostly the same. On the other hand, at $\theta_E > 1$ mas and $\pi_E > 0.3$ part, the two advanced models predict much more events than the standard model ($\sim 4.6 \times$ for model B and $\sim 2.7 \times$ for model C). These parts of the distributions are dominated by small $D_L$ events (leading to large $\pi_{rel}$ according to Eq. 3). For these nearby events, Models B and C have warmer dynamics, so the event rate is much higher than that in the standard model.

However, these tail events are rare and only denote < 5% of all events, and they do not have much influence on the statistical results. By contrast, for specific individual events have large $\theta_E$ or $\pi_E$, the model selection might bring a large difference in the parameter space.

In addition, from Fig. 4 we find that, similar to $t_E$, $\theta_E$ and $\pi_E$ also follow the power-law asymptotic behavior in the tails, $p(\log \theta_E) \propto \theta_E^3$ and $p(\log \pi_E) \propto \pi_E^3$ when they are very small, and $p(\log \theta_E) \propto \theta_E^3$ and $p(\log \pi_E) \propto \pi_E^3$ when they take large values. The identical asymptotic power-law indices of ($\theta_E$, $\pi_E$) and $t_E$ are somewhat a coincidence because $t_E$ depends on kinematics while ($\theta_E$, $\pi_E$) do not. With a simple toy model as in Mao & Paczynski (1996), we can analytically derive the ±3 powers (for more details, see the appendix).

We define “extreme” events for the typical lens mass $0.5M_\odot$ are those whose $t_E > 200$ days or $\theta_E > 1$ mas or $\pi_E > 0.3$, which show great differences in different models. We find that for (standard, B, C) models, the $t_E$ extreme events only contribute (0.13%, 0.08%, 0.10%) of all events, while $\theta_E$ and $\pi_E$ extreme events contribute (1.8%, 4.1%, 2.8%) and (1.0%, 2.7%, 1.8%) of all events.
This indicates that $\theta_E$ and $\pi_E$ might be somewhat more sensitive to the Galactic model than $t_E$.

### 3.3 Bayesian Analysis

Below we use three examples to illustrate the impact of Galactic models on individual events, where 1) only $t_E$ is available, 2) $t_E$ and $\pi_E$ are available, 3) $t_E$ and $\theta_E$ are available.

#### 3.3.1 $t_E$ is the Only Observable

In reality, for an individual event, the timescale value and its error are obtained by modeling the observation light curve. As mentioned above, if $t_E$ is the only observable, other parameters like $M_L$ and $D_L$ can only be obtained by Bayesian analysis. The Bayesian analysis here is actually an inverse problem of calculating the $t_E$ distribution in Section 3.1, the difference is that the weight of each event $w_i$ needs an extra $t_E$ prior,

$$w_i = \Gamma_i \mathcal{L}_i(t_E), \quad \mathcal{L}_i(t_E) = \frac{\exp \left[-(t_{E,i} - t_E)^2/2\sigma_{t_E}^2\right]}{\sqrt{2\pi}\sigma_{t_E}}, \quad (17)$$

where $t_E$ and $\sigma_{t_E}$ is the mean value and standard deviation of the timescale from an observation, and $\mathcal{L}_i(t_E)$ is the likelihood of $t_E$. Here we directly give some typical values $t_E = 5, 20, 50, 200$ days and $\sigma_{t_E} = 0.1 t_E$. Here we use the Kroupa mass function for brown dwarfs and main-sequence stars (Kroupa 2001), where

$$\frac{d\pi(M_L)}{d\log M_L} \propto \begin{cases} M_{L}^{0.7}, & 0.013 < M_L/M_\odot < 0.08 \\ M_{L}^{-1.3}, & 0.08 < M_L/M_\odot < 0.5 \\ M_{L}^{-0.3}, & 0.5 < M_L/M_\odot < 1.3 \end{cases} \quad (18)$$

In addition, we also consider the distribution of white dwarfs (WD), neutron stars (NS) and black holes (BH) (Karolinski & Zhu in prep.), where

$$\frac{d\pi_{WD}}{dM_L} \propto N(0.65M_\odot, 0.16^2M_\odot^2), \quad 0.3 < M_L/M_\odot < 1.4,$$

$$\frac{d\pi_{NS}}{dM_L} \propto N(1.5M_\odot, 0.2^2M_\odot^2), \quad 1.1 < M_L/M_\odot < 2.5,$$

$$\frac{d\pi_{BH}}{dM_L} \propto 10^{-\frac{M_L}{10^{15}M_\odot}}, \quad 2.5 < M_L/M_\odot < 80.$$ (21)

and their number densities are taken to be 3.1%, 0.2% and 0.1% of normal stars. The fractions and parameters are derived or speculated from Giannicchel et al. (2012) (WD), Kizilhan et al. (2013) and Sartore et al. (2010) (NS), and Olejak et al. (2020) (BH).

We also generate $4 \times 10^9$ events for each Galactic model, and then substitute different $t_E$ prior to obtain the $D_L$ and $M_L$ posterior probability distributions (see Figs. 5 and 6). We can see that for the same model, longer $t_E$ decreases the bulge-source probability and increases the disc-source probability. This is consistent with the result in Section 3.1, for which disc-source events contribute more at long $t_E$.

The differences of $D_L$ and $M_L$ between different models is only a few percent and mostly emerges on long $t_E$ events, which indicates that the Bayesian analysis is insensitive to Galactic models if $t_E$ is the only observable. The conclusions made by previous work with simple models may still be mostly valid, however, if $\theta_E$ or $\pi_E$ was reliably measured and was involved in the analysis, the model dependence could be significant (see next two Sections 3.3.2 and 3.3.3).

#### 3.3.2 A Black Hole Candidate: OGLE3-ULENS-PAR-02

Here we apply the Bayesian analysis procedure on a black hole candidate OGLE3-ULENS-PAR-02 found by Wyrzykowski et al. (2016). The coordinate of this event is $(RA, Dec)_2000 = (17:57:23.14, -28:46:32.0)$. The light-curve analyses show the parameters are (we choose the $u_0 > 0$ solution here): $t_E = 296.1^{+7.6}_{-7.4}$ days, $\pi_{E,E} = -0.051^{+0.002}_{-0.002}$, and $\pi_{E,N} = 0.033^{+0.001}_{-0.001}$, where $\pi_{E,E}$ and $\pi_{E,N}$ are the eastern and northern components of $\pi_E$. 

---

**Figure 3.** The event rate as a function of $t_E$ of the three models and their ratios (bottom right). $N_0$ is the total event rate of the standard model. Here we fix $M_L = 0.5M_\odot$. Different Galactic components are shown in different panels. Each model is presented in different colours, and the event rate ratio of each component compared to that in the standard model is also shown in the bottom right panel. Note that the ratio of the standard model (green) is normalized to unity, and the total probability of the standard curve (upper right panel) is unity.

**Figure 4.** The event rate as a function of $\theta_E$ and $\pi_E$ for the three models and their ratios to the standard model. $N_0$ is the total event rate of the standard model. Here we fix $M_L = 0.5M_\odot$. Each model is presented in a different colour. The ratio of the standard model (green) is normalized to unity, and the total probability of the standard curves are unity. Notice that the ratios at large $\theta_E$ and $\pi_E$ are the same, and are much larger than those for $t_E$. 

---

**References:**

1. Giannicchel et al. (2012)
2. Kizilhan et al. (2013)
3. Sartore et al. (2010)
4. Olejak et al. (2020)
5. Wyrzykowski et al. (2016)
In this case, the weight for each simulated event becomes

\[ w_i = \Gamma_i \mathcal{L}_i(t_E) \mathcal{L}_i(\pi_{E,E}) \mathcal{L}_i(\pi_{E,N}), \]

where \( \mathcal{L}_i(t_E) \), \( \mathcal{L}_i(\pi_{E,E}) \) and \( \mathcal{L}_i(\pi_{E,N}) \) are the likelihood of \( t_E \), \( \pi_{E,E} \) and \( \pi_{E,N} \). In order to show the model dependence on the velocity distribution, we use a log-uniform mass function in \( M_L \in [10^{-2}, 10^3] \, M_\odot \) as the \( M_L \) prior in all models. We simulated \( \sim 2 \times 10^{10} \) events and the results are shown in Table 4 and Fig. 7. We find the disc-lens probability dominates in all models, and the masses are \( 2.3^{+4.1}_{-1.1} \, M_\odot \), \( 3.2^{+5.4}_{-1.8} \, M_\odot \) and \( 4.6^{+10.8}_{-3.0} \, M_\odot \) for the standard model, Model B and C, respectively.

In Table 4, for disc-lens events, both Model B and Model C estimate a larger mass than the standard model. This is consistent with our discussion in Section 3.1, that under a specific lens mass, Model B and Model C cannot produce as many long \( t_E \) events as the standard model, therefore the Bayesian analysis for a given long \( t_E \) event tends to be more massive. In addition, the largest \( M_L \) part is dominated by B-D and D-D events whichever model we choose. And the fraction of the disc-source component increases in both Model B and Model C, which leads to the higher estimation of \( M_L \).

Note that even though the model used in Wyzykowski et al. (2016) is similar to our standard model, the parameters are quite different. In addition, they assumed the event is 100% a B-D event, so the mass they estimated \( M_L = 9.3^{+5.7}_{-4.3} \, M_\odot \) is different from ours (but still within \( \sim 1 \sigma \)), which again, proves the importance of model selection.

3.3.3 A Free Floating Planet Candidate: OGLE-2017-BLG-0560

We also applied the Bayesian approach to a free-floating planet candidate OGLE-2017-BLG-0560 at the coordinate (RA, Dec)\( \equiv (17:51:51.33, -30:27:31.4) \) (Mróz et al. 2019). This is an ambiguous event because the mass occasionally located at the critical mass between a brown dwarf and a planet (\( \sim 13 M_J \), where \( M_J = 0.95 \times 10^{-3} M_\odot \) is the mass of Jupiter). From light-curve modeling and source star identifying, the parameters of this event are yielded: \( t_E = 0.905^{+0.005}_{-0.006} \) days, and \( \pi_E = 38.7^{+1.6}_{-1.8} \) mas. The proper motion of the source star is also available in Gaia DR2 (Gaia Collaboration et al. 2018), where \( \mu_{S,RA} = -2.19 \pm 0.30 \) mas/yr, \( \mu_{S,Dec} = -11.7 \pm 0.24 \) mas/yr and their correlation \( \rho_{\mu_{S,RA}, \mu_{S,Dec}} = -0.77 \).
0.276. Such a proper motion strongly prefers a bulge-source event to a disc-source event ($\sim 1 \times 10^{-10}$ in any model), so the Bayesian analysis below only includes bulge-source components.

Here we generate the source velocity from Gaia proper motion probability distribution rather than from the Galactic model. This is equivalent to generate from the Galactic model and then multiply a $\mu_S$ likelihood on the weight, but can reduce the computational work.

We used a log-uniform MF for $M_L \in [10^{-5}, 10^1] M_\odot$ and simulated $\sim 4 \times 10^9$ events to acquire the posterior $M_L$ and $D_L$ distribution, the weight of each event in this case is

$$w_i = \Gamma_l L_i(t_E) L_i(\theta_E),$$

where $L_i(\theta_E)$ is the likelihood of $\theta_E$. Table 5 and Fig. 8 show the results in detail.

![Figure 6](image)

**Figure 6.** The posterior probability distribution of $M_L$ under the given Galactic model and $t_E$ prior. Each column represents a Galactic model, and each row uses the same $t_E$ prior. The vertical dashed lines denote the mean and 90% confidence interval of the overall distribution. The fraction of each component (B-B, B-D, D-B, D-D) and the $>2.5 M_\odot$ probability are shown on each subfigure, too.

| Model | Component | B-B | B-D | D-B | D-D | All |
|-------|-----------|-----|-----|-----|-----|-----|
| Standard | $P_{\text{Comp}}$ | 11.2% | 47.8% | 28.0% | 13.1% | Total |
|        | $D_L$ | 7.1±0.5 kpc | 3.7±1.4 kpc | 7.6±0.8 kpc | 5.2±1.0 kpc | 5.4±2.1 kpc |
|        | $M_L$ | 1.0±0.3 $M_\odot$ | 5.0±2.4 $M_\odot$ | 1.7±0.6 $M_\odot$ | 1.9±0.9 $M_\odot$ | 2.3±1.1 $M_\odot$ |
|        | $P(>3M_\odot)$ | 0.0% | 78.8% | 1.1% | 21.7% | 40.8% |
| Model B | $P_{\text{Comp}}$ | 9.7% | 51.1% | 23.3% | 15.8% | Total |
|        | $D_L$ | 7.1±0.6 kpc | 3.6±1.4 kpc | 7.6±0.8 kpc | 3.8±1.7 kpc | 4.6±2.9 kpc |
|        | $M_L$ | 1.0±0.4 $M_\odot$ | 5.3±2.5 $M_\odot$ | 1.7±0.6 $M_\odot$ | 4.2±2.6 $M_\odot$ | 3.7±1.8 $M_\odot$ |
|        | $P(>3M_\odot)$ | 0.0% | 80.7% | 1.3% | 63.6% | 51.7% |
| Model C | $P_{\text{Comp}}$ | 5.8% | 64.4% | 15.1% | 14.7% | Total |
|        | $D_L$ | 7.0±0.7 kpc | 3.3±1.5 kpc | 7.6±0.8 kpc | 3.3±1.4 kpc | 3.8±1.2 kpc |
|        | $M_L$ | 1.0±0.4 $M_\odot$ | 6.1±2.6 $M_\odot$ | 1.7±0.6 $M_\odot$ | 5.9±3.7 $M_\odot$ | 4.6±3.0 $M_\odot$ |
|        | $P(>3M_\odot)$ | 0.0% | 83.0% | 0.8% | 76.8% | 64.9% |
Figure 7. The posterior probability distribution of OGLE3-ULENS-PAR-02. The fraction of each Galactic component is labeled on the top of $D_L$ plot, and the probability of $M_L > 3M_\odot$ is shown on the $M_L$ plot. The black vertical dashed lines denote the mean value and 68.3% confidence interval of the overall distribution. We also label the mean and ±1σ values for $D_L$ and $M_L$ from Wyrzykowski et al. (2016) with the pink dashed lines.

Table 5. Bayesian Results of the Free-floating Planet Candidate OGLE-2017-BLG-0560

| Model | Component | B-B | B-D | All |
|-------|-----------|-----|-----|-----|
|       | $P_{\text{Comp}}$ | 83.3% | 16.7% | - |
| Standard | $D_L$ | $7.9^{+0.7}_{-0.6}$ kpc | $4.4^{+0.9}_{-0.5}$ kpc | $7.7^{+0.8}_{-1.5}$ kpc |
|       | $M_L$ | $16.0^{+1.4}_{-0.7}$ M$_\odot$ | $1.7^{+1.1}_{-0.9}$ M$_\odot$ | $13.0^{+3.8}_{-4.4}$ M$_\odot$ |
|       | $P(<13M_\odot)$ | 40.7% | 95.7% | 49.9% |
| Model B | $P_{\text{Comp}}$ | 48.9% | - | 51.1% |
|       | $D_L$ | $7.9^{+0.7}_{-0.6}$ kpc | $5.5^{+1.2}_{-1.1}$ kpc | $6.8^{+1.4}_{-1.0}$ kpc |
|       | $M_L$ | $16.0^{+1.4}_{-0.7}$ M$_\odot$ | $2.8^{+1.2}_{-1.1}$ M$_\odot$ | $6.2^{+2.3}_{-1.4}$ M$_\odot$ |
|       | $P(<13M_\odot)$ | 40.4% | - | 97.0% |

For the standard model, our result is qualitatively consistent with Mróz et al. (2019). In our case, the lens is either a Jupiter-mass planet in the disc ($M_L = 1.7^{+3.1}_{-0.7}$ M$_\odot$, $D_L = 4.4^{+2.0}_{-1.3}$ kpc) or a brown dwarf in the bulge ($M_L = 16.0^{+36.6}_{-8.4}$ M$_\odot$, $D_L = 7.9^{+0.7}_{-0.6}$ kpc), the probability ratio is $P_{\text{bulge}} : P_{\text{disc}} \approx 83\% : 17\%$. However, Model B and Model C show different results, Model B shows a higher probability of disc lens at $M_L \sim 10^{-3} M_\odot$ or 1M$_\odot$ and $D_L \sim 3$ kpc because the sharper and higher structure on the peak of the $r_E$ distribution (top left panel in Fig. 3) can contribute many events. Bulge-lens events are the same for all of the three models, but Model B and Model C have a dynamically warmer disc, making it easier to produce short $r_E$ events in the disc for a fixed lens mass. So for the disc-lens component, the two models shows higher lens mass estimation of $M_L = 2.5^{+2.2}_{-1.2}$ M$_J$ and $M_L = 10.0^{+18.2}_{-5.7}$ M$_J$. However, a warmer disc also increases the event rate, leading to a significant higher disc-source probability ($P_{\text{bulge}} : P_{\text{disc}} \approx 49\% : 51\%$ in B and 26% : 74% in C) and making the overall $M_L$ posterior smaller and the probability of being a planet ($P(<13M_\odot)$ higher. The smoother transition of the velocity distribution between the disc and the bulge also makes the $D_L$ more likely to be located in the inner part of the disc.
Figure 8. The $D_L$, $M_L$ and combined posterior probability distribution of the free-floating planet candidate OGLE-2017-BLG-0560. The fraction of each Galactic component is labeled on the top of $D_L$ plot, and the probability of $M_L < 13M_J$ is shown on the $M_L$ plot. The black vertical dashed lines denote the mean value and 68.3% confidence interval of the overall distribution. We also label the mean and ±1σ posterior $M_L$ values from Mróz et al. (2019) (the result of prior 2 in it) with the pink dashed lines.

4 DISCUSSION AND CONCLUSION

The two examples in Sections 3.3.2 and 3.3.3 show the influence of the Galactic model selection on the understanding of individual events. It changes the estimation of $D_L$ and $M_L$, and often more significantly, $P_{\text{bulge}} : P_{\text{disc}}$ (for the importance of the planet bulge-disc ratio, see Zhu et al. 2017). By comparing the Sections in 3.3 we can see, generally, the more observables are measured and more precisely they are, the greater the influence of the Galactic model in Bayesian approach.

On the other hand, if enough microlensing events are observed, it is possible to constrain the Galactic model as well. Wood & Mao (2005) tested the Han & Gould (2003) model with the OGLE-II observations (Sumi et al. 2006) by the $t_E$ distribution, and found a consistency at the ~52% confidence level. However the sample is small at that time. Shan et al. (2019) used the Spitzer microlensing events to test two commonly used Galactic models, and the models showed good consistency. The event sample they used has good constraints on physical parameters, but is of moderate size (includes 13 published Spitzer events) and no extreme events (e.g., $t_E > 300$ days or $t_E < 3$ days) are included. The models may show some difference if any "tail" events have happened. The current Galactic model are derived from the visible part of the Milky Way galaxy, while microlensing is also sensitive to dark objects (e.g., stellar remnants, free-floating planets, and perhaps primordial black holes) in the galaxy. They may have different density profiles, velocity distributions and mass functions, and microlensing is the unique way to detect and identify them.

The Galactic model plays an critical role in the Bayesian analysis of microlensing events. Sometimes different Galactic model choices can have a big impact on the results, especially at the tails. However, before a feasible way of measuring $\theta_E$ and $\pi_E$ is found, the Bayesian approach is still the main way for us to learn the lens properties.

Even the “realistic” Galactic models in this paper are far from perfect. For example, our mass function has no dependence on the density profile. In reality the time evolution and even different stellar populations need to be concerned. Moreover, our model did not consider the bar model differences. But after only roughly considering the two known effects of velocity distribution in our Galaxy, that is radial dependence and asymmetric drift, the results have already shown a considerable difference. This paper provides an idea on how the Galactic model influence the results. In the future, a more refined model like that in galaxia (Sharma et al. 2011) or a non-parametric model needs to be develop. We suggest that researchers should be more careful when selecting a Galactic model.

ACKNOWLEDGEMENTS

H.Y., S.M., X.Z. and W.Z. acknowledge support by the National Natural Science Foundation of China (Grant No. 11821303 and 11761131004). We thank Wei Zhu for the insightful discussions. We thank the efforts of the galpy software developers.
DATA AVAILABILITY

The data underlying this article will be shared on reasonable request to the corresponding author.

REFERENCES

Alcock C., et al., 1993, Nature, 365, 621
An J. H., et al., 2002, ApJ, 572, 521
Beaulieu J.-P, et al., 2006, Nature, 439, 437
Bennett D. P., Becker A. C., Calitz J. J., Johnson B. R., Laws C., Quinn J. L., Rhie S. H., Sutherland W., 2002, arXiv Astrophysics e-prints,
Bennett D. P., et al., 2014, ApJ, 785, 155
Binney J., Tremaine S., 2008, Galactic Dynamics: Second Edition
Bovy J., 2015, ApJS, 216, 29
Cao L., Mao S., Nataf D., Rattenbury N. J., Gould A., 2013, MNRAS, 434, 595
Dwek E., et al., 1995, ApJ, 445, 716
Gaia Collaboration et al., 2018, A&A, 616, A1
Giammichele N., Bergeron P., Dufour F., 2012, ApJS, 199, 29
Gillesen S., Eisenhauer F., Trippe S., Alexand er T., Genzel R., Martins F., Ott T., 2009, ApJ, 692, 1073
Gould A., 2000a, ApJ, 535, 928
Gould A., 2000b, ApJ, 542, 785
Gould A., Loeb A., 1992, ApJ, 396, 104
Han C., 2008, ApJ, 681, 806
Han C., Gould A., 1995, ApJ, 447, 53
Han C., Gould A., 2003, ApJ, 592, 172
Kent S. M., Dame T. M., Fazio G., 1991, ApJ, 378, 131
Kiraga M., Paczynski B., 1994, ApJ, 430, L101
Kiziltan B., Kottas A., De Yoreo M., Thorssett S. E., 2013, ApJ, 778, 66
Koshimoto N., Bennett D., 2019, arXiv e-prints, p. arXiv:1905.05794
Kroupa P., 2001, MNRAS, 322, 231
Lam C. Y., Lu J. R., Hosek Matthew W. J., Dawson W. A., Golovich N. R., 2020, ApJ, 889, 31
Mao S., 2012, Research in Astronomy and Astrophysics, 12, 947
Mao S., Paczynski B., 1991, ApJ, 374, L37
Mao S., Paczynski B., 1996, ApJ, 473, 57
Mao S., et al., 2002, MNRAS, 329, 349
Mróz P., et al., 2018, AJ, 155, 121
Mróz P., et al., 2019, A&A, 622, A201
Olejak A., Belczynski K., Bulik T., Sobolewska M., 2020, A&A, 638, A94
Robin A. C., Reylé C., Derrière S., Picard S., 2003, A&A, 409, 523
Sartore N., Rapacz J., Treves A., Turolla R., 2010, A&A, 510, A23
Schöning R., Binney J., Dehnen W., 2010, MNRAS, 403, 1829
Shan Y., et al., 2019, ApJ, 873, 30
Sharma S., Bland-Hawthorn J., Johnston K. V., Binney J., 2011, ApJ, 730, 3
Shu F. H., 1969, ApJ, 158, 505
Sumi T., et al., 2006, ApJ, 636, 240
Sumi T., et al., 2011, Nature, 473, 349
Suzuki D., et al., 2018, ApJ, 869, L34
Ulaski A., Szymanski M., Kaluzny J., Kubiak M., Krzeminski W., Mateo M., Preston G. W., Paczynski B., 1993, Acta Astron., 43, 289
Wegg C., Gerhard O., 2013, MNRAS, 435, 1874
Wood A., Mao S., 2005, MNRAS, 362, 945
Wyrzykowski L., et al., 2016, MNRAS, 458, 3012
Zang W., et al., 2020, ApJ, 891, 3
Zhu W., et al., 2017, AJ, 154, 210

APPENDIX A: THE POWER LAW TAIL BEHAVIORS IN $\theta_E$ AND $\pi\theta_E$ DISTRIBUTIONS

Consider the simplest model that the source is at a fixed distance $D_S$, and the lenses are uniformly distributed along the line of sight with constant mass $M$. The lens velocity $V$ is independent of $D_L$. The Einstein radius of a microlensing event is
$$R_E^2 = \frac{4GM D_L}{c^2} (D_S - D_L) = \frac{4GM D_S}{c^2} x(1-x),$$
where we define $x \equiv D_L/D_S$. The event rate is
$$d\Gamma \propto \frac{R_E^2}{l_E} n_L(D_L) dD_L p(V) d^3V \propto R_E dD_L V_L p(V) d^3V \propto R_E dD_L,$$
where $n_L(D_L)$ is the lens distance distribution which is a constant, $p(V)$ is the velocity distribution, and $V_L$ is the transverse velocity. We omit the velocity term because both $\theta_E$ and $\theta_L$ are pure geometric parameters that are independent of $V$. Furthermore, in our simple model, $V$ is also independent of the distance $D_L$, the integration will give a constant. Then the total event rate is given by
$$\Gamma_{\text{tot}} = \int_{0 < D_L < D_S} \int_{0 < x < 1} d\Gamma \propto \int_{0 < x < 1} \sqrt{x(1-x)} \, dx.$$
(A3)

For the $\theta_E$ distribution, we are interested in the probability that $\theta_E < \theta_0$,
$$P(\theta_E < \theta_0) = \int_{\theta_0 < \theta_E} d\theta_E.$$
(A4)
The $\theta_E < \theta_0$ condition puts constraint on $x$ as
$$P(\theta_E < \theta_0) = P(\theta_E < \theta_0) = P(\sqrt{\frac{\theta_E}{\theta_0}} < 1 < \theta_0) = P \left[ x > (\theta_0^2 + 1)^{-1} \right],$$
(A5)
where $\pi_{\text{rel}} \equiv \frac{\sqrt{D_S}}{\sqrt{M}}$ and we define
$$\beta \equiv \sqrt{\frac{D_S}{kM}}$$
(A6)

Therefore
$$P(\theta_E < \theta_0) \propto \int_{(\beta^2 \theta_0^2 + 1)^{-1}}^{1} \sqrt{x(1-x)} \, dx \propto \frac{1}{4} \left[ \beta\theta_0 (\beta^2 \theta_0^2 - 1) + \arctan(\beta\theta_0) \right].$$
(A7)

Normalizing the result, we find
$$P(\theta_E < \theta_0) = \frac{2}{\pi} \left[ \beta\theta_0 (\beta^2 \theta_0^2 - 1) + \arctan(\beta\theta_0) \right].$$
(A8)

Expanding the result at $\theta_0 = 0$, we have
$$P(\theta_E < \theta_0) \approx \frac{16}{3\pi} (\beta\theta_0)^3 - \frac{48}{5\pi} (\beta\theta_0)^5 + \ldots.$$ 
(A9)
Similarly, for $\theta_0 \to \infty$,
$$P(\theta_E > \theta_0) = 1 - P(\theta_E < \theta_0) \approx \frac{16}{3\pi} (\beta\theta_0)^{-3} \frac{48}{5\pi} (\beta\theta_0)^{-5} + \ldots.$$ 
(A10)

Thus we obtain the asymptotic behaviors for the log $\theta_E$ distribution as follows:
$$p(\log \theta_E)\bigg|_{\theta_E \to 0} = \frac{dP(\log \theta_E)}{d\log \theta_E} \bigg|_{\theta_E = 0} \approx \frac{16 \ln 10}{\pi} \beta^3 \theta_E^{-3},$$
(A11)
$$p(\log \theta_E)\bigg|_{\theta_E \to \infty} = \frac{dP(\log \theta_E)}{d\log \theta_E} \bigg|_{\theta_E = \infty} \approx \frac{16 \ln 10}{\pi} \beta^{-3} \theta_E^{-3}.$$ 
(A12)
Similarly, for $\pi_\text{E}$ we have

$$P(\pi_\text{E} < \pi_0) = P\left(\frac{\pi_{\text{rel}}}{\sqrt{\text{K}M} < \pi_0}\right)$$

$$= P\left(\frac{\beta}{D_S} \sqrt{\frac{1}{x} - 1 < \pi_0}\right) = P\left[x > \left(\frac{D^2_S}{\beta^2} \pi_0^3 + 1\right)^{-1}\right],$$

(A13)

and

$$P(< \pi_\text{E}) \approx \frac{16}{3\pi} \left(\frac{D_S}{\beta}\right)^3 \pi_\text{E}^3, \quad (\pi_\text{E} \to 0); \quad (A14)$$

$$P(> \pi_\text{E}) \approx \frac{16}{3\pi} \left(\frac{D_S}{\beta}\right)^{-3} \pi_\text{E}^{-3}, \quad (\pi_\text{E} \to \infty). \quad (A15)$$

And the asymptotic behaviors for $\log \pi_\text{E}$ are given by

$$p(\log \pi_\text{E})|_{\pi_\text{E} \to 0} = \frac{dP(< \pi_\text{E})}{d\log \pi_\text{E}}|_{\pi_\text{E} \to 0} \approx \frac{16 \ln 10}{\pi} \left(\frac{D_S}{\beta}\right)^3 \pi_\text{E}^3;$$

(A16)

$$p(\log \pi_\text{E})|_{\pi_\text{E} \to \infty} = \frac{dP(< \pi_\text{E})}{d\log \pi_\text{E}}|_{\pi_\text{E} \to \infty} \approx \frac{16 \ln 10}{\pi} \left(\frac{D_S}{\beta}\right)^{-3} \pi_\text{E}^{-3}. \quad (A17)$$

Although we derived the asymptotic behavior from a simple toy model, the ±3 power-law behaviors on the tails appear to be valid on many other situations. For example, for $P(< \theta_\text{E})$, the coefficient $\beta$ will be a function of the mass $M$ if the mass has a distribution $f(M)$, and the asymptotic behavior at $\theta_\text{E} \to 0$ will be

$$P(< \theta_\text{E}) \approx \sum_i \frac{16}{3\pi} \beta^3 (M_i) f(M_i) \theta_\text{E}^3 \propto \theta_\text{E}^3, \quad (A18)$$

where the tail behavior $P(< \theta_\text{E}) \propto \theta_\text{E}^3$ still holds. In Fig. 2, the bulge-lens probability distributions (red and green) do not follow the asymptotic −3 power-law behaviors at large $\theta_\text{E}$ ($\pi_\text{E}$) because the bulge does not extend to the solar neighborhood.

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