

Lens Modeling in SpaceWarp

First Author,¹ Second Author,² and Third Author³

ABSTRACT

In ~~Spacewarp~~, ~~volunteers are the initial phase of Spacewarp~~ (SW1), ~~volunteers were invited to search the CFHTLS sky survey for lens candidates to look for gravitational lenses~~. Here we report on a ~~system that allows experienced volunteers to model lens candidates~~. A ~~web interface that gives experienced volunteers the opportunity to model the candidates that have been identified~~. These volunteers, ~~were also invited to model a sample of 29 simulated lenses~~ were modeled, each multiple times. The ~~to order to gauge the~~ quality of the models ~~was that~~ were being generated. These were then examined in ~~different ways~~. In particular greater detail, with particular attention being paid to the following: (i) ~~the identification of~~ image parities and time arrivals ~~were identified correctly in most cases, but not always~~; (ii) the mean convergence (equivalent to the enclosed mass) ~~was quite well constrained in the image region, and finally~~; (iii) the performance of ~~experience~~ volunteers ~~was comparable to professionals~~. ~~a volunteer vs. a professional~~. In most cases, the volunteers were found to correctly identify the image parities and time arrivals, along with a mean convergence that was contained within the image region; In all, the results could be comparable to that of professionals.

•

Subject headings:

1. Introduction

Our present picture of the formation of the largest structures in the Universe — galaxies and clusters of galaxies — is that gravitational instabilities (whose early stages are observable as fluctuations of the microwave background) caused a collapse into gravitationally bound

or should that be web application? Also, I change the tense of the opening statement. Not too sure of the correct way of introducing an abbreviation (SW1 in this case), and I was a bit liberal with my editing - Sorry. Another comment - is 'Well contained' better than 'contained' or is it just a figure of speech?

¹First place

²Second place

³Third place

structures. Dark matter would have dominated the initial collapse, but once deep wells in the gravitational potential appeared, gas would have fallen into them while radiating away its gravitational potential energy through atomic processes to form the first stars and galaxies. With time, galaxies and clusters would continue to merge and change their form. This rough outline is well-established now, but many puzzles remain to be solved before our understanding is satisfactory. To progress towards solving these, we need more knowledge of dark matter. And to study dark matter, gravitational lensing is attractive, because it depends only on mass.

There are many manifestations of gravitational lensing in astrophysics, but the most spectacular are multiple images. Gravitational lensing depends on the sky-projection mass density, and this fact has a counter-intuitive consequence: nearby galaxies do not produce multiple images, because their projected density is not high enough, but beyond 100 Mpc or so, most galaxies are potential lenses. Even so, for a potential lens to really act as a lens, a background source in the close-by region is required. So, in practice galaxies are very rarely observed as lens. The same tendency applies to clusters of galaxies, but since clusters are bigger, they are more likely lens-like.

In lensing galaxies, the observable features usually involve a single object multiply imaged. It is therefore desirable to have large numbers of lenses, and this is the strategy researchers are following. The largest single study so far uses 58 lenses (Koopmans et al. 2009). Other studies have compared lens models in more detail with stellar mass (Leier et al. 2011) or used time-delay information from lensed quasars to infer the Hubble time (Coles 2008; Paraficz & Hjorth 2010).

The trend is clear: researchers want to find many more lenses and model them in more sophisticated ways.

A good summary of the observational situation is offered by Figure 1, which shows known secure multiply-imaging lenses. The non-uniformity on the sky is not intrinsic, it just indicates the density of deep surveys up to 2013. At HST resolution, of order 1 square degree of sky must be searched to find a lens. Older ground-based surveys yield a lens per roughly 10 square degree of survey area. By the large surveys starting now (DES¹ or PanStarrs²) the expected yield is something in between. Given their unprecedented survey area, these are likely to yield thousands of lenses. Over the 2020s, LSST³ can be expected to give us

¹<http://pan-starrs.ifa.hawaii.edu>

²<http://www.darkenergysurvey.org>

³<http://www.lsst.org/>

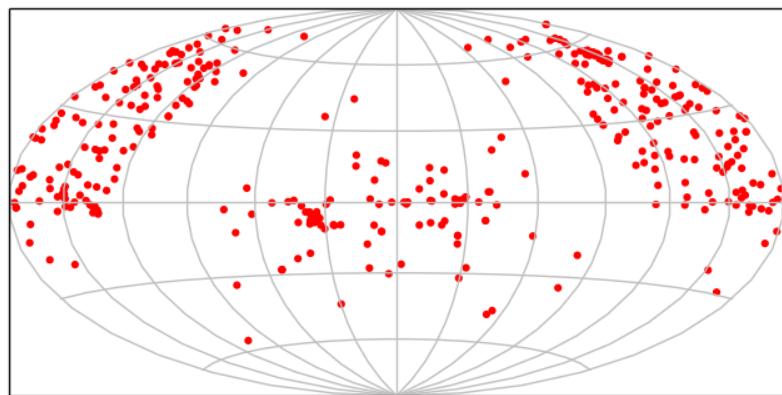


Fig. 1.— Sky distribution of 398 published lenses considered secure as of September 2013 (from the Masterlens catalog at the University of Utah, maintained by Joel Brownstein and Leonidas Moustakas.) The map is in Hammer-Aitoff projection, with North up and RA = 0 in the middle. The empty swathes to the left and right of center are the Milky Way.

ten thousand lenses.

The question now arises: can lenses be found and modeled automatically in large surveys? Work using software robots (Marshall et al. 2009) gives some interesting and unexpected results. In clean lensing system in uncrowded fields with high signal-to-noise, the robots do very well. In most situations, however, robots miss lenses (low completeness) or contaminate the results with non-lenses (low purity). Robots can be made to prioritize completeness or purity, but they cannot deliver both.

In response to the lessons learned from automated lens searches, the *Spacewarps* project⁴ has been launched. *Spacewarps* is part of the *Zooniverse* family of citizen-science projects,⁵ where members of the public are invited to analyze different kinds of scientific data which are too difficult for robots and too large for specialists. In *Spacewarps* itself, survey sky images are presented for volunteers to identify lens candidates. The first tranche of survey data, introduced from May 2013, is an area of ~ 172 square degrees (or a quarter of a percent of the sky) from the Canada-France-Hawaii Telescope Legacy Survey. The survey is divided into patches of 440×440 pixels, each of which is seen by ten volunteers. Simulated lenses are mixed in with the data, both to help train volunteers on what to look for, and to estimate completeness and purity. These data had previously been analyzed by robots Gavazzi et al. (2012); More et al. (2012). A list of candidates is being processed, but intermediate results show some very good candidates, as well as re-discoveries of candidates previously found robotically.

These encouraging results now raise the question: could modeling of the lenses also be done by volunteers? Modeling is a much more difficult task than searching for lens candidates, as it requires some expert knowledge. Nonetheless, a panel of the *Spacewarps* volunteers is quite experienced from earlier projects, having spent a thousand hours or more with data. Some of these experienced volunteers are very interested in more demanding projects. This also is a general trend (cf. Khatib et al. 2011). The present project then suggests itself.

The purpose of this study was to provide a means to model a large number of gravitational lenses by showing that gravitational lens modeling can be learned / done by volunteers. We suggest, that volunteers will be as successful as professionals with modeling if provided with an easy to use tool with visual feedback (What you see is what you get, WYSIWYG) and a minimal set of instructions. Volunteers will then crowd work (using buzz word here ;)

⁴<http://www.spacewarps.org>

⁵<http://www.zooniverse.org>

) / work collaborative on modeling lenses from several sources / groups at a central place. Since this is an iterative learning process, the more involved volunteers will quickly gain knowledge that can be passed down to new volunteers. That creates a social structure that scales well with the number of volunteers, as other projects have already shown. Finding [cit](#) people working as volunteers has been shown to be successful last but not least by Space-Warps and the whole galaxy zoo project. To test the people’s abilities, we investigated the performance of a few volunteers modeling a set of simulated lenses. We tested the ability to correctly identify lensed images and reproduce similar mass map of the lens.

2. A lens modeling program

It is well known that making a scientific contribution is a central motivation cited by citizen-science volunteers. So for a lens modeler in a citizen-science project, nobody wants that it should sacrifice scientific usefulness in order to be fun to use. That said, being aesthetically pleasing and having a short initial learning curve are also essential qualities. Also desirable is that the user is encouraged/challenged to go deeper and wider into the subject. Yet another aspect is enabling small incremental contributions by different people. SpaghettiLens tries to address all of these.

2.1. Theory: Fermat’s Principle

There are several ways to understand the formation of arcs and multiple images in gravitational lensing. We will follow some ideas originally introduced by Blandford & Narayan (1986), based on Fermat’s principle. The key to this approach is an abstract construct called the *arrival-time surface*. This surface cannot itself be observed, but several observable quantities can be derived from it.

Consider a gravitational lens. As in most astrophysical lensing, this lens is ‘thin’ along the line of sight, and effectively lies on a plane. Let x, y be coordinates on this plane. These coordinates measure ordinary physical lengths (meters, etc) on the lens plane. Let there be a point source of light at $x = 0, y = 0$ but far behind the lens plane. Now imagine a light ray coming from the source to the lens plane (x, y) and then changing its path to reach the observer. The observer would see the source apparently behind (x, y) the lens. If this happened, the light ray would have taken a longer path than having come directly to the observer. The geometrical path difference would be $\propto x^2 + y^2$ for small angles. Let us define

$$A_{\text{geom}}(x, y) = \frac{1}{2}(x^2 + y^2). \quad (1)$$

As written, this looks an area. But we can also think of it as a time delay introduced by the light ray bending — with an unwritten constant factor. That factor is basically the lens distance divided by the speed of light; the precise expression has to take the expansion of the Universe into account, and is given in the Appendix. The top part of Figure 2 shows what A_{geom} looks like: it is just a paraboloid. The minimum of the surface is at the source, and that is where, according to Fermat’s principle, the image will be. The rest of the arrival-time surface is just an imaginary surface.

The geometrical time delay (1) is not the only one. The warping of space by a gravitational field introduces a further time delay A_{grav} . The arrival time of the light ray, compared to what it would have been with no lensing, is

$$\text{arrival time} = A_{\text{geom}} + A_{\text{grav}}. \quad (2)$$

We can also express A_{grav} as an area, with the same implicit constant factor that makes it a time. The expression, however, is more complicated. We first introduce the notation $\nabla^2 f(x, y)$ to denote

$$\frac{f(x + \Delta x, y) + f(x - \Delta x, y) + f(x, y + \Delta y) + f(x, y - \Delta y) - 4f(x, y)}{\Delta x \Delta y} \quad (3)$$

in the limit of $\Delta x, \Delta y$ small. Then we write the projected mass density of the lens, which would be in kg m^{-2} , in special units (given in the Appendix) such that the surface density becomes a dimensionless field $\kappa(x, y)$. Then

$$\nabla^2 A_{\text{grav}}(x, y) = -2\kappa(x, y). \quad (4)$$

Note that $A_{\text{grav}}(x, y)$ is not given explicitly in terms of $\kappa(x, y)$, but as a differential equation that must be solved. Equations of the type (4) are, however, well known (as Poisson’s equation in two dimensions), and there are techniques for solving them. Incidentally, κ has a second meaning, as well as being a dimensionless form of the projected density: it also measures how a bundle of light rays is brought together by gravitational lensing, and is called the *convergence*.

What is the effect of mass on the arrival-time surface? If the mass distribution $\kappa(x, y)$ is circular, $A_{\text{grav}}(x, y)$ will have a peak at the center of that circle. But if the mass is not precisely in front of the source, the arrival-time surface will not be circular any more. The result is illustrated in the middle part of Figure 2. In addition to the maximum, there is now a minimum and a saddle point. According to Fermat’s principle, images appear at maximum and saddle points as well as minimum. If the mass distribution is non-circular,

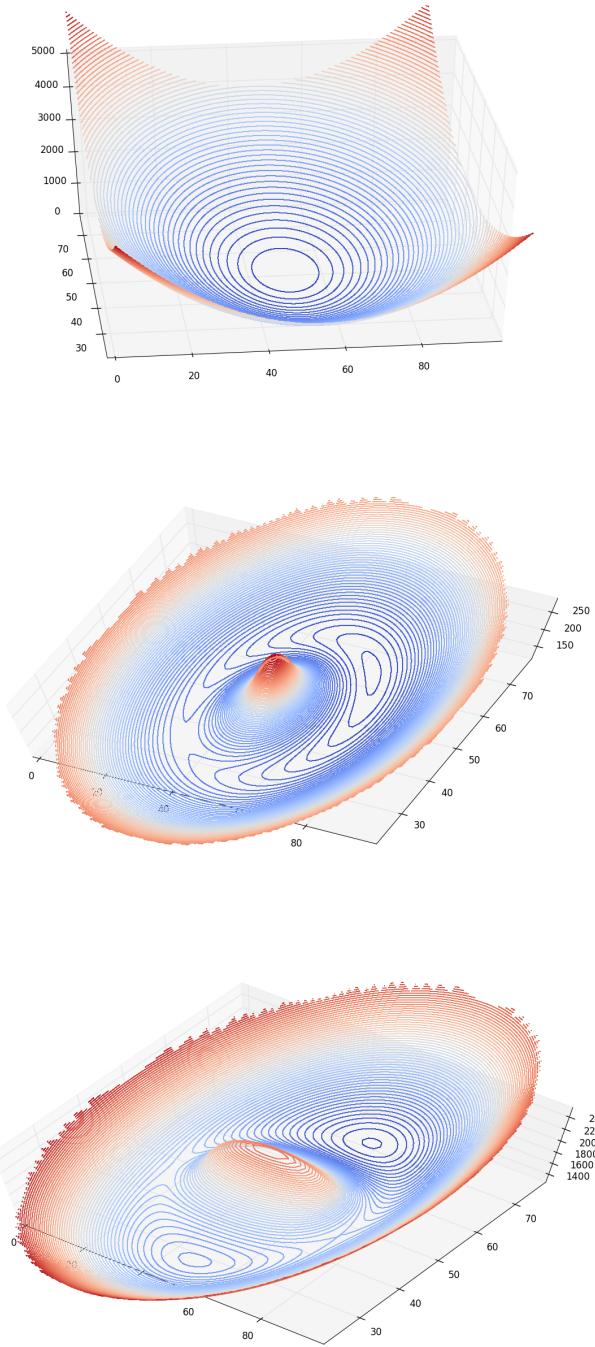


Fig. 2.— Arrival-time surfaces, with contour levels of equal arrival time. The upper surface is with no lens. The middle surface is when a circular lensing mass (offset from the source) is added; a maximum, a minimum and a saddle point can be seen. The last surface is the result of an elongated lensing mass; a maximum, two minima and two saddle points can be seen.

more images can appear. The lower example in Figure 2 shows the arrival-time surface due to an elongated mass. There is still a maximum, part of an elongated hill, and around it there are two minima and two saddle points.

So far we have considered a single point source. What if there is an extended source? To understand this, let us consider what happens if we move the original point source slightly, or equivalently, keep the original point source behind $x = 0, y = 0$ and move the lens slightly. The contours of constant arrival time will naturally move slightly, and so will the images. The movement of the contours will be most noticeable where the contours are far apart, that is where the arrival-time surface is nearly flat. As is evident from Figure 2, this is the region where the minimum and saddle points lie, or near the images. In this region, points on the source that are close together produce images that are comparatively far apart. In other words, the image is highly magnified. In summary, lower curvature in the arrival-time surface for a point source implies larger magnification of an extended source. Conversely, where the arrival-time surface is strongly curved, the image will be demagnified. We see from Figure 2 that the arrival-time surface tends to be highly curved near the maximum. Hence maximum tend to be demagnified. In practice, maxima of the arrival time are nearly always too faint to see. The minima and saddle points dominate.

In the lensed examples in Figure 2, the maximum in each case is evident, but how can we distinguish the minima from the saddle points? The key is to examine the contours of equal arrival time. Contour curves loop around a minimum or a maximum, but near a saddle point, the contours do not loop around, instead they approach and then pull away. The contour curve precisely level with a saddle point forms an X at the saddle point. In Figure 2, one such X is evident. In other cases the X contour passing through the saddle point is not itself plotted, but we can tell by the other contours where the X would be. These self-crossing contours will play a special role in the following.

2.2. SpaghettiLens

SpaghettiLens implements the ideas of the arrival-time surface into a modeling program.

The first modeling step with SpaghettiLens is for the user to call up the SpaceWarpes of interest, and make an educated guess for the topography of the arrival-time surface — specifically, the locations of the maxima, minima and saddle points that correspond to images of the brightest part of the source. Input is given by tracing a proposal for the saddle-point contours. Figure 3 shows some example inputs — and incidentally also indicates the origin of the name SpaghettiLens.

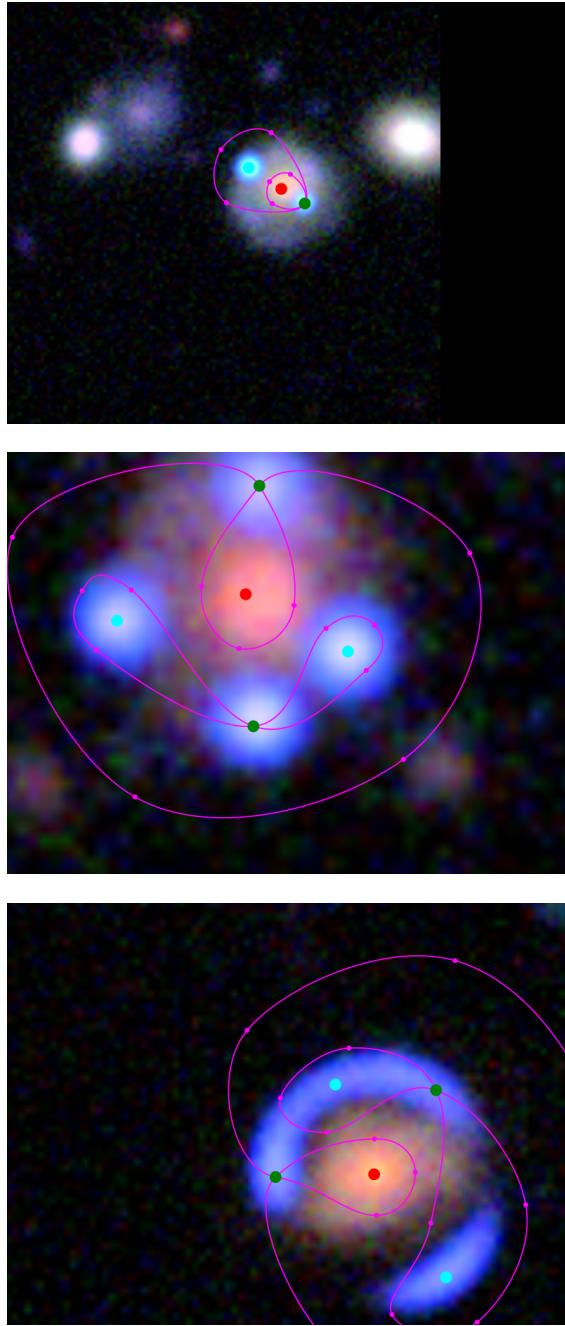


Fig. 3.— Examples of Spaghetti input. The models from these appear later in Figures 4, 9 and 11.

SpaghettiLens, as implemented so far, assumes that the lens is dominated by a single galaxy, and that the center of that galaxy is the sole maximum. Once the maximum has been identified, we wish to characterize the rest of the arrival-time surface in relation to it. From geometry, any surface that has a central maximum and goes high at the edges must have a limaçon-like contour, meaning a two-looped curve with one loop inside the other. The maximum lies within the inner loop, a minimum lies between the two loops, and a saddle point lies at the self-intersection of the contour. The top panel of Figure 3 shows an example: red marks a suggested maximum, green a saddle point, and blue a minimum. (The small pink dots are just help sketch the proposed contour.) The middle and lower panels of Figure 3 show a more complex situation. The minimum has split into three images: two new minima with a saddle point in between. In this situation the contour through the new saddle point forms a figure of eight, its two loops enclosing the two new minima. The configuration of a limaçon, possibly with a figure of eight inside, suffices to give a preliminary description of the arrival-time surface for nearly all cases where the lens is dominated by one galaxy, and the source is a single small object.

The exact placement of the pink spaghetti loops shown in Figure 3 has no significance. Their utility is just to help mark up the image as maximum, minima and saddle points. As the figure suggests, the marking up can be done very easily with a mouse. Human intuition is required to (a) identify lensed images and separate them from other background light sources, and (b) classify and order images according to arrival time. But the computer helps by allowing only valid lensing configurations to be entered, and ensuring that the odd image theorem• is taken care of.

cit

In the second modeling step, once the user has sketched a proposed spaghetti configuration, SpaghettiLens sends the input to its server-side modeling engine, called GLASS. The task of GLASS is to find a mass distribution $\kappa(x, y)$ that exactly reproduces the locations of the maximum, minima and saddle points. Now, this criterion by itself is extremely under-determined — there are infinitely many mass distributions that will reproduce a given set of maxima, minima and saddle points, but typically they (a) produce lots of extra images, and (b) look very unlike galaxies. Additional assumptions (a prior) are necessary. GLASS uses a prior based on suggestions by Saha & Williams (1997).

1. The mass distribution is built out of non-negative tiles of mass. (Sometimes these tiles are called mass pixels, but we should emphasize that they are unrelated to image pixels, and are much larger.)
2. There is a notional lens center, say (x_0, y_0) which is identified with the maximum of the arrival time. The source can have an arbitrary offset with respect to the lens center.

3. The mass distribution must be centrally concentrated, in two respects. First, the circularly averaged density must fall away like

$$[(x - x_0)^2 + (y - y_0)^2]^{-1/2}$$

or more steeply. Second, the direction of increasing density at any (x, y) can point at most 45° away from (x_0, y_0) .

4. The lens must be symmetrical with respect to 180° rotations about (x_0, y_0) . This symmetry assumption can be relaxed if the user wishes.

There are still infinitely many models that satisfy both data and prior constraints, but now they are more credible as galaxy lenses. It is then possible to generate an ensemble of models. The sampling technique used by GLASS is described in (Lubini & Coles 2012), and improves upon earlier techniques (Williams & Saha 2000; Saha & Williams 2004). Typically, ensembles of 200 models are used. That is to say, what we call a SpaghettiLens model is really the mean of an ensemble of 200 models, and its estimated uncertainty is the range covered by the whole ensemble.

In the third modeling step, having generated a model ensemble, SpaghettiLens post-processes it to present results and diagnostics to the user for inspection. This takes the form of three figures.

1. A gray scale plus contour map of the mass distribution.
2. A contour map of the arrival-time surface.
3. A synthetic image of the lensed features.

After examining this feedback, the user can archive the results for discussion, or modify their input and try again, or discard the attempt altogether.

The fourth modeling step is discussion among modelers and iteration on the model. Any archived model can be revised by another user: they can modify the spaghetti configuration slightly or drastically, or change options like the size of the mass tiles. Particularly interesting lens candidates lead to trees of models in this way. Forum discussion then prunes the tree, focusing attention on one or a few models.

3. A lens modeling challenge

Interested volunteers from the SpaceWarps forum were initially introduced to SpaghettiLens through a video tutorial and by videocon. After this introductory stage, a modeling challenge was presented. This consisted of 29 simulated lenses (sims) covering a range of lensing configurations.

The SpaceWarps sims were generated by AM, in consultation with PM and AV. To estimate the performance of the volunteers and the quality of the generated models, two analysis were done. The first analysis tested the correct identification and ordering of lensed images. The second one compared the mass distribution of the lens $\kappa(x, y)$ of the generated models to the mass distribution of the simulations.

3.1. The simulated lenses

For the sake of blind testing, the information in this section was revealed neither to RK, while choosing the challenge set of 29 sims, nor to the modelers (volunteers EB, CC, CM, JO, JW and ‘expert modeler’ PS) until the modeling stage was done.

The sims were produced using `gravlens` (Keeton 2001a,b) using the CFHTLS survey data and catalogues Coupon et al. (2009). They were of three kinds, as follows.

1. Imitating lensed quasars: having a singular elliptical isothermal lens (SIE) plus constant external shear, and a circular Gaussian source.
2. Emulating lensed galaxies: similar to the above, but with an elliptical de Vaucouleurs source.
3. Resembling cluster lenses: having a source as above, but a more complicated lens, with one dominant elliptical SIE and one or more perturbing elliptical SIEs, plus a circular NFW (Navarro et al. 1996, 1997) to represent the underlying dark matter distribution.

Formulas for the lenses appear in Keeton (2001a). The SIE lenses follow equations (33–35) of that work, with core radius set to zero. The NFW lens is in equations (48) and (50), while shear is the γ term in equation (76).

3.2. Some example models

The modelers proposed a total of 129 models for the 29 sims in the challenge. Figures 4 to 11 show eight of models in some detail.

Figure 4 shows a simple example. The top two panels show contours of convergence $\kappa(x, y)$, and of the arrival-time for a point source at the brightest point of the sim source.⁶ The spatial scale is in pixels. The information in these two panels was, of course, kept secret during the modeling challenge. The panels in the middle row and at the lower right are the results that SpaghettiLens returns to the modeler, in order for them to assess the model. These derive from the mean of an ensemble of 200 models generated by SpaghettiLens. The model mass distribution (left mid) is returned as a contour map superimposed on an intensity map. A fairly smooth mass distribution, as here, is a good sign. An irregular or checkboard pattern in the mass map usually indicates a bad model. The models arrival-time contours are shown in the right middle panel. The special contour passing through a saddle point is shown in black: this is the model version of the spaghetti sketch provided by the user. Also provided is a synthetic image of the lensed source in the bottom right panel. For this, SpaghettiLens assumes a simple conical source profile. The user can change the contrast level on the image, which (though it is not saved) amounts to adjusting the width of the cone. Finally, to the lower left of the figure, we have a comparison of the input and recovered mass profiles. The panel shows a average κ within a given radius, as a function of radius. The red curve is the true value, and where it crosses unity (dotted horizontal line) is the notional Einstein radius $\Theta_{E, \text{sim}}$. The two blue curves are the minimal and maximal mean enclosed κ from the internal ensemble in SpaghettiLens. The region between the blue curves is shaded between the radii of the innermost and the outermost images — this is the confidence region from the modeling. As we see, the shaded blue is slightly above the red curve. The Einstein Radius Θ_E of the model is estimated crossing the mean enclosed κ (not plotted) with unity. In summary: the identification of minimum and saddle point is correct, but the estimated Einstein radius is a little too high.

Figure 5 shows a lens with substructure in the form of a smaller secondary galaxy. The galaxies in such group or cluster sims were, in fact, based on galaxies visible in the images — but the modelers were not told in advance whether this was the case. The model does not include any substructure, but otherwise is not bad. The minimum and saddle point are correct, and the Einstein radius is only a little underestimated.

⁶Although the arrival-time contours represent an abstract quantity that is not directly observable, the contour map may actually resemble the appearance of the lensed arcs. The resemblance is due to a serendipitous computer-graphical effect (Saha & Williams 2001).

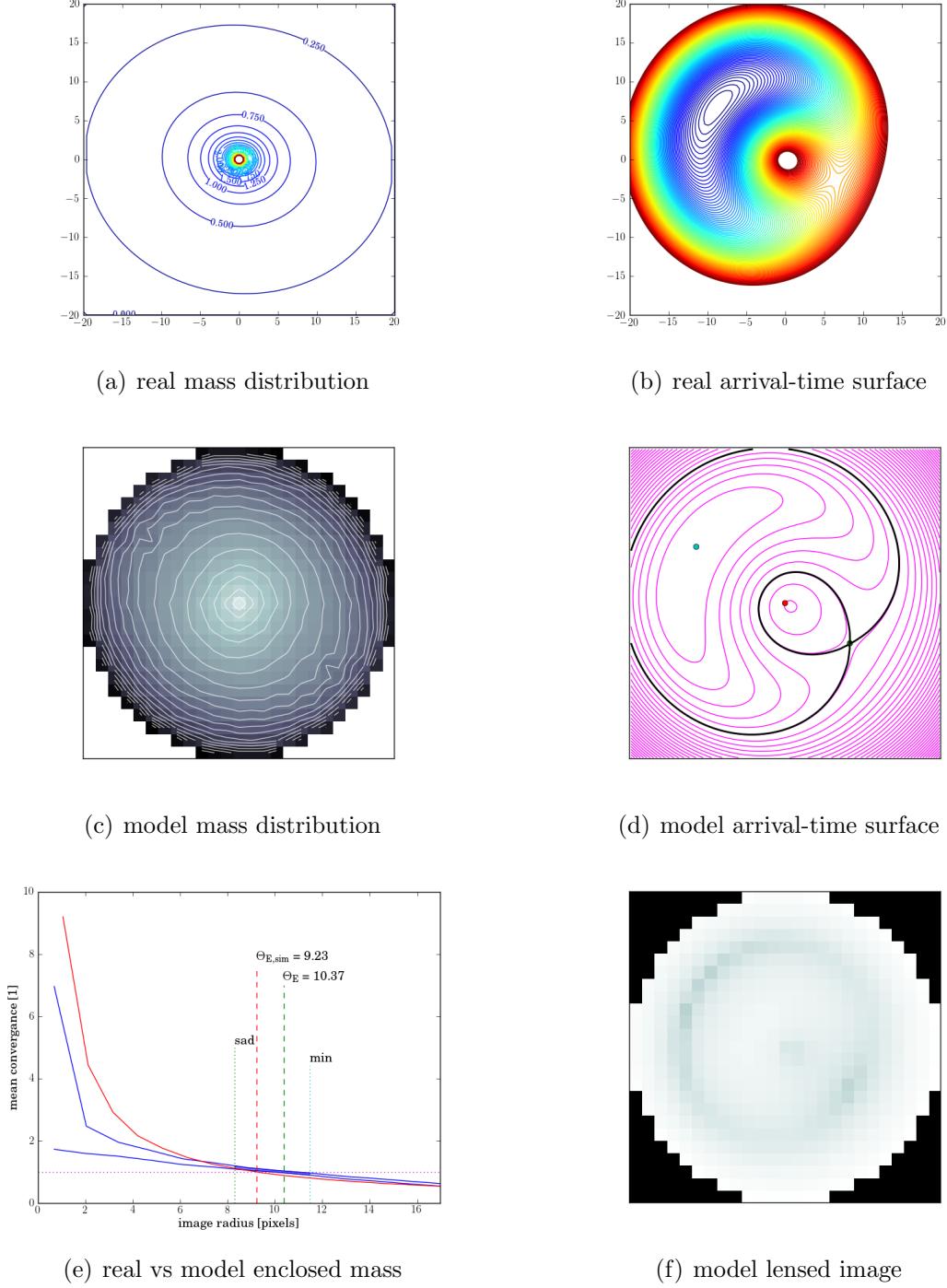


Fig. 4.— Sim ASW000102p and a model. A simple configuration; the model input is shown in the top panel of Figure 3. For details of the individual panels here, see section 3.2.

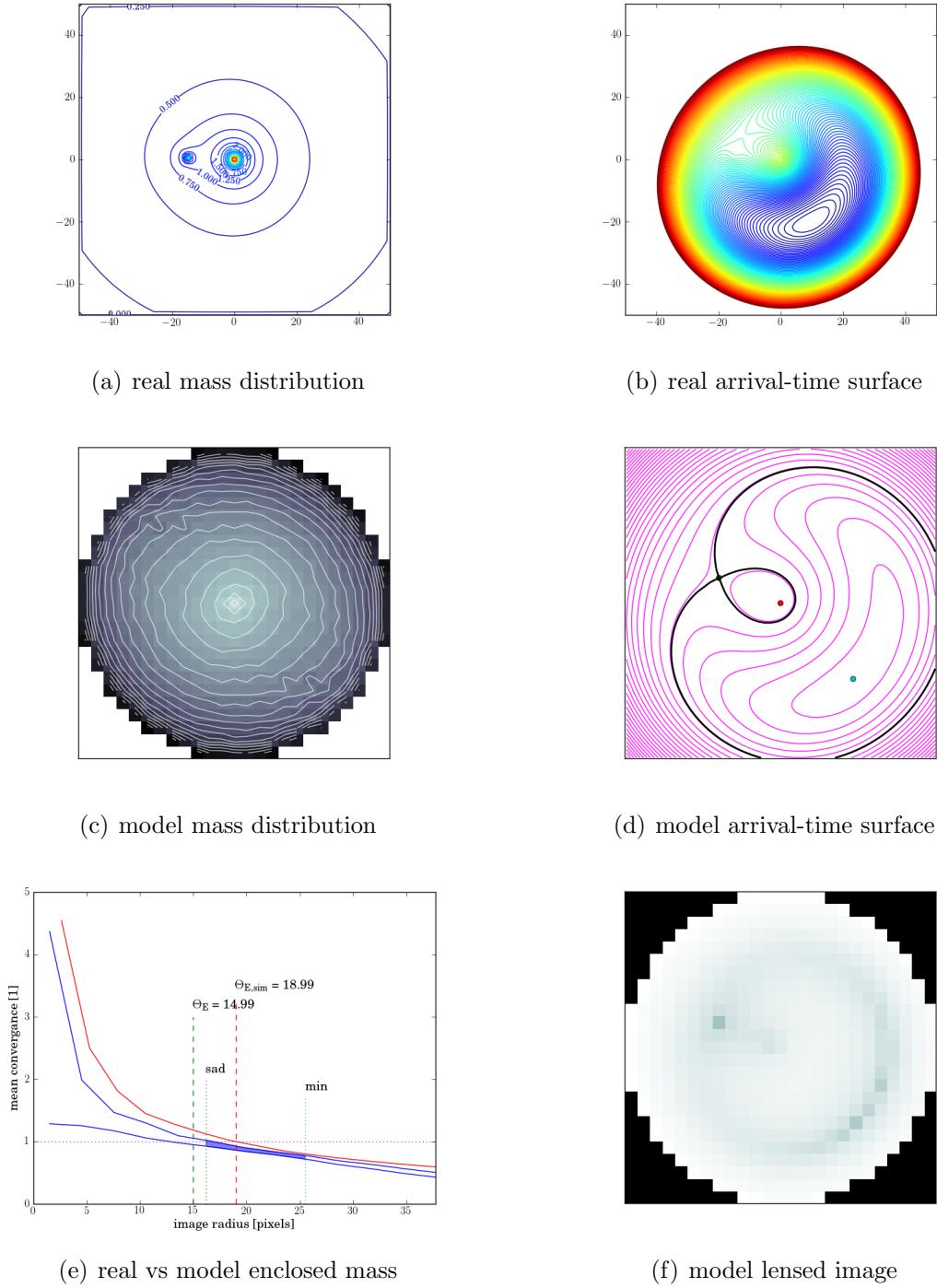


Fig. 5.— Sim ASW000195x and a model.

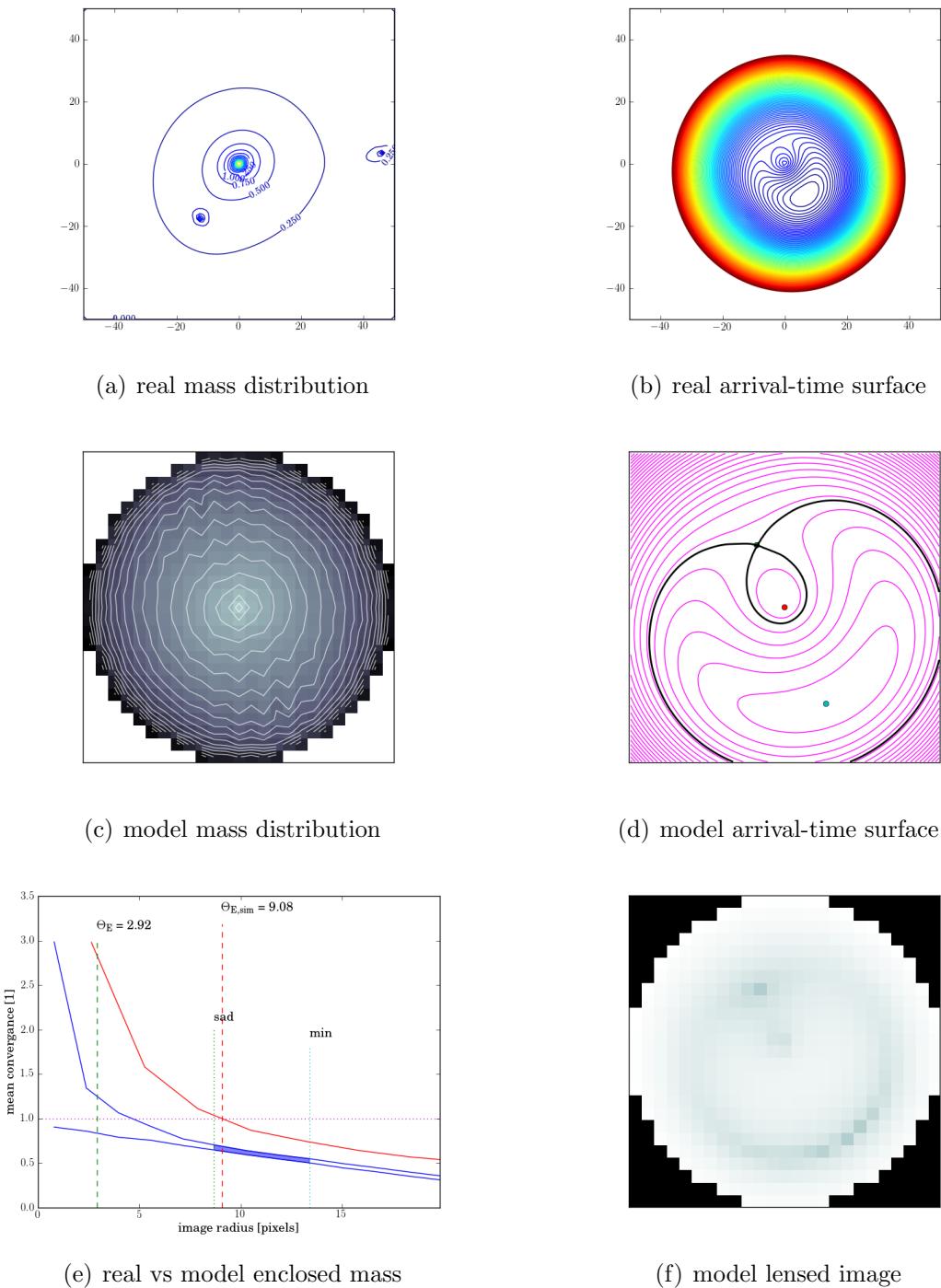


Fig. 6.— Sim ASW0000vqg and a model.

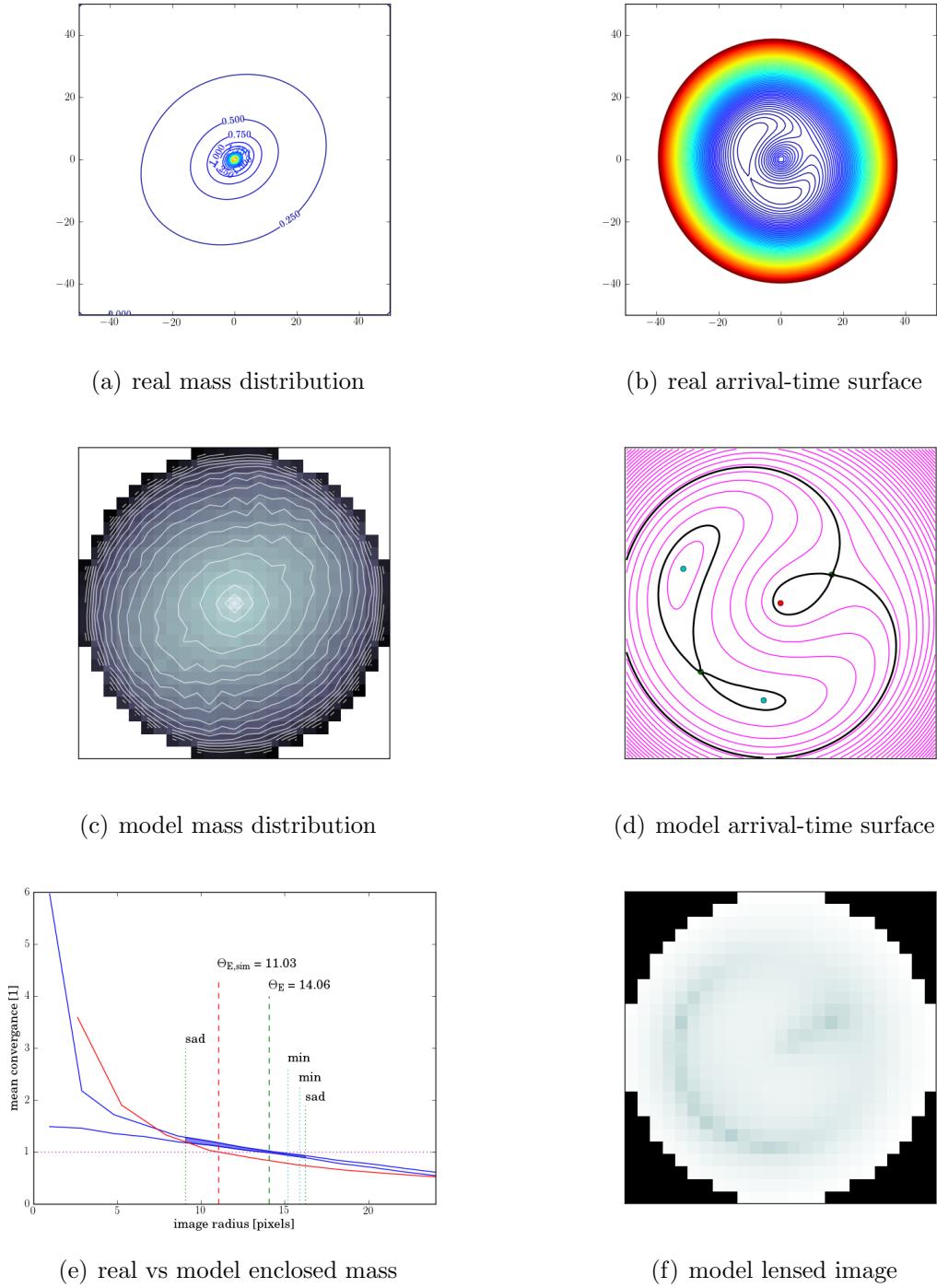


Fig. 7.— Sim ASW0004oux and a model.

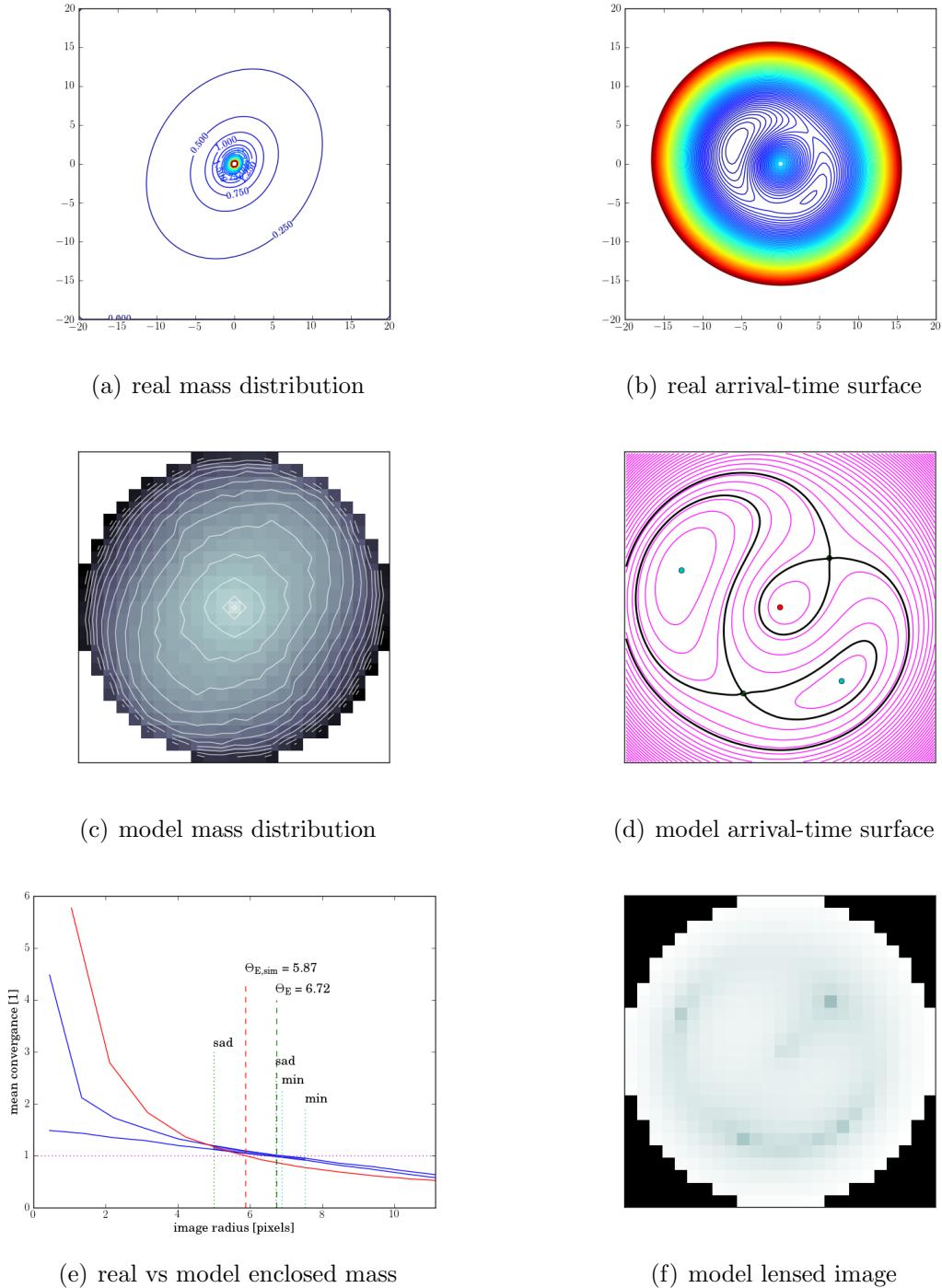


Fig. 8.— Sim ASW0001hpf and a model. Inclined quad.

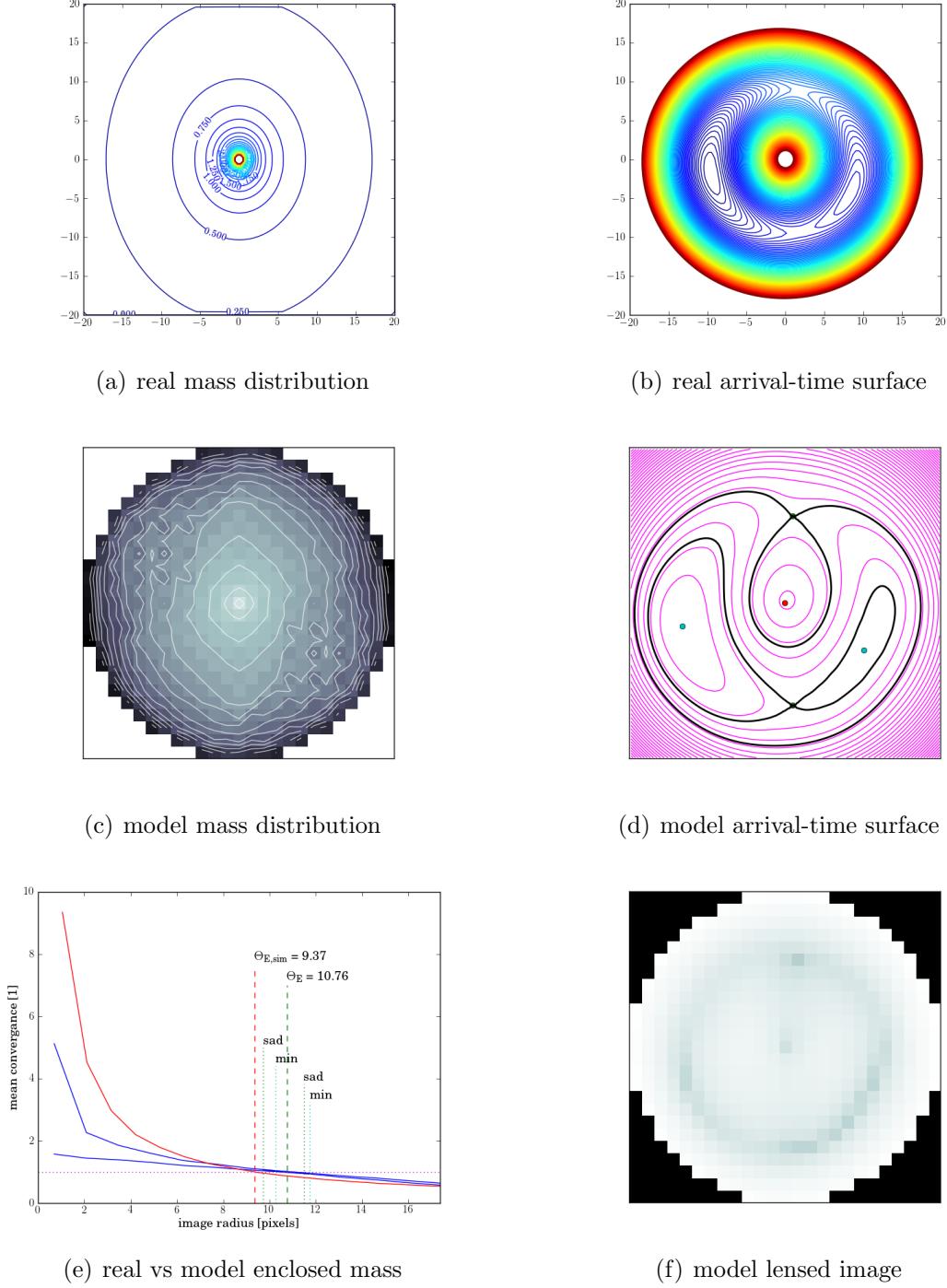


Fig. 9.— Sim ASW0000h2m and a model. The model input appears in the middle panel of Figure 3.

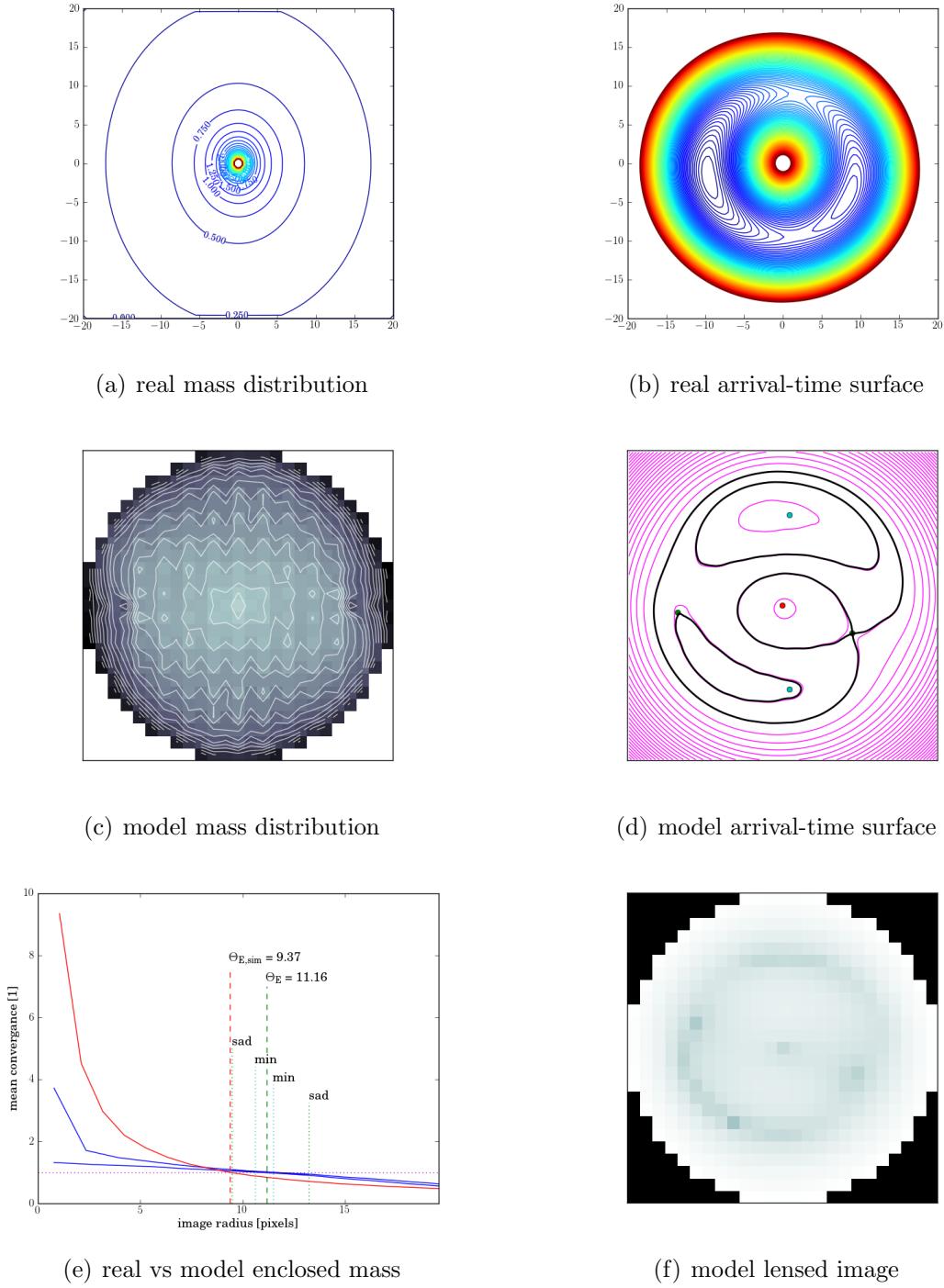


Fig. 10.— Sim ASW0000h2m and another model. The image parities in this case were incorrectly identified.

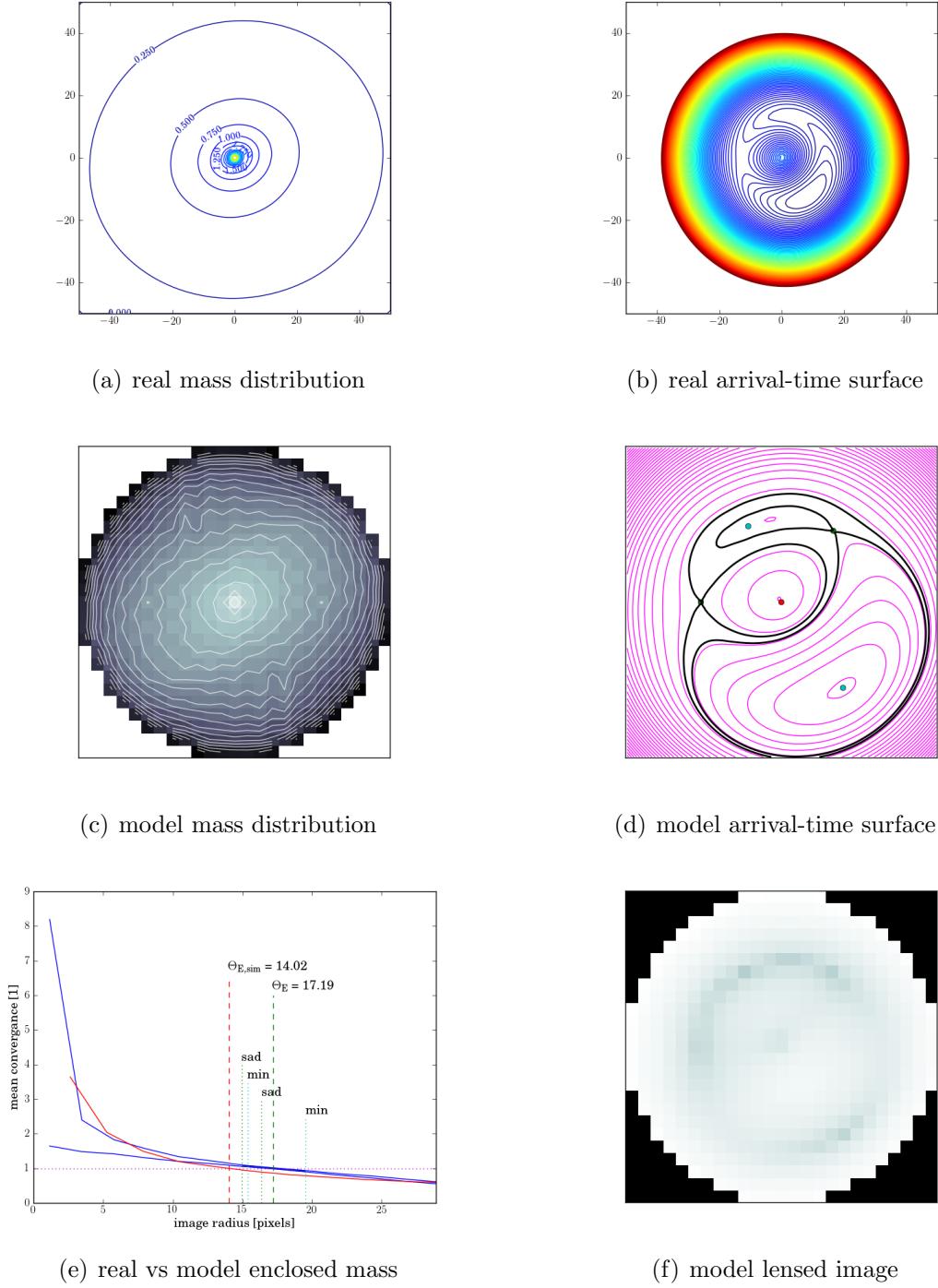


Fig. 11.— Sim ASW0002z6f and a model. The model input appears in the bottom panel of Figure 3. An incipient quad

Figure 6 shows a case where substructure leads to a poor model.

Figure 7 shows an example of an arc that has split into three images. This kind of configuration, with a counter-image close to the lensing galaxy and a more distant arc/triplet on the other side, generically arises from an elongated mass distribution when the source is displaced along the elongated direction.

Figure 8 shows another quad. This kind of configuration arises when the mass is elongated and the source is displaced at an angle to the elongation.

Figure 5 shows a fairly symmetric quad. The minima and saddle points are correctly identified, and the orientation of the ellipticity of the mass distribution is correctly reproduced. The Einstein radius is somewhat overestimated. Figure 9 shows another model of the same system. In this one, the identification of the minima and saddle points was incorrect, and mass distribution comes out elongated East-West instead of North-South. The mass distribution also appears somewhat jagged and the saddle-point contours are not as clean as in the previous examples; these are often indicators of a problem with the model. The enclosed mass is, however, none the worse — the reason is probably that in a relatively symmetrical image configuration, the Einstein radius is quite well constrained by the images in a fairly model-independent way.

To conclude this set of examples, Figure 11 shows another type of quad. Actually, in this case the brightest part of the source is only doubly imaged, but the source extends into a region that produces four images. We can also see from the real arrival-time surface that a point source is a double on the verge of splitting into a quad. The modeler interpreted the system as a quad. The appearance of the arcs, shown in the bottom panel of Figure 3, looks like an arc and counter-image such as discussed with Figure 7 above. But there is an important difference: the long arc is closer to the galaxy, as if the arc and counter-image have swapped roles. This configuration arises if the source displacement is perpendicular to the long axis of the lensing mass.

3.3. Test of image identification

A first manual evaluation tested the volunteers ability to reconstruct the arrival time surface given a survey image containing a sim. This task consists of two parts. First to correctly identify and locate the lensed images. Second the correct ordering for the identified lensed images in respect of the arrival time.

While we expected the identification of lensed images to be trivial, given the nature of

the survey images and the success of SpaceWarps, we expect the correct ordering to be more difficult. This part tests the volunteers understandings of the theory of arrival time surfaces and the odd number theorem. While we can provide the volunteers with some general rules of thumb, ordering involves imagination and guessing and therefore training could improve results in a later stage.

This first test was designed to give some feedback on the difficulties volunteers encounter, to further improve the tutorial materials.

The evaluation of the volunteers performance was done manually, comparing their input from SpaghettiLens and the resulting reconstructed arrival time surface contour line plot (arrival plot) to the arrival plot generated using simulation parameters.

The images of the system are considered to be identified correctly, if all the images have been identified and are approximated within $\pm 0.05 \cdot \text{imgage width}$. The parity is considered correct, if those identified points have the right ordering with respect to arrival time.

Additionally, ten types of errors (labeled E01 – E10, listed in Table 1) that occurred were identified. Each generated model could contain more than one error.

Table 1 presents a summary of this evaluation. We conclude that the volunteers are performing very well identifying and positioning images, with a performance of 92% (R1, $p=0.92$). Most of the problems where due to unclear arc-like structures (E01, $p=0.18$; E04, $p=0.03$; E08, $p=0.04$). Critical errors like the failure to identify all five images in a five images system (E03, $p=0.04$) or to include too many images (E10, $p=0.01$) did almost never happen. From this we conclude that the introduction materials was adequate and the volunteers understand the basics of gravitational lensing.

The assignment of the parity of the images was a more difficult task. In 59% (R2, $p=0.59$, $N=70$) of the cases the volunteers succeeded to identify the right configuration. Most of the failures are due to E06 ($N=38$, $p=32\%$), followed by E05 ($N=7$, $p=6\%$). E06 describes a situation, where the minima and saddle points of a five image configuration were exchanged (rotated by $\pm 90^\circ$), see Figure 9 for an example. E05 describes a situation, where the ordering of the saddle points was wrong (rotation by 180°). While these errors occurred, we suspect they can be avoided with better training material and some examples for the obvious cases. For more challenging cases, like very symmetrical distribution of the lensed images (for example model 7022, Figure 9), those errors should still produce plausible results, as will be explained in the next section.

		n	p
N:	Total number of models	119	1.00
R1:	images approx. on right location	110	0.92
R2:	images with correct parity	70	0.59
E01:	inaccurate in arc	21	0.18
E02:	wrong parity in 3 lens conf.	2	0.02
E03:	identified 3 of 5 imgs.	5	0.04
E04:	modeled arc with single img.	4	0.03
E05:	π rotated parity	7	0.06
E06:	$\pi/2$ rotated parity	38	0.32
E07:	missed faint img.	1	0.01
E08:	too many imgs in arc.	5	0.04
E09:	missed double img	3	0.03
E10:	too many imgs.	1	0.01

Table 1: 1. col: id of model; 2.col: places of imgs idendited correctly; 3. col: more or less correct identification of extr points. 4.col: exact identification of extremal points. 5.col: type(s) of errors ocured.... 1) inaccurate placement in an extended arc; 2) wrongly identified sad and min in 3 image configuration; 3) identified only 3 instead of 5 images; 4) tried to model an arc with a min instead of min-sad-min; 5) PI-err (rotation by 180 degrees; in 5 image configuration, exchanged the ordering of the two saddle points); 6) PI/2-err (rotation by 90 degree; sad-min-sad-min-sad); 7) missed faint image(s); 8) tried to model an arc with min-sad-min instead of only min; 9) did identify two close by images as one; 10) used 7 or more image to model a 5 image system.

3.4. Test of mass-profile recovery

The second test was to compare the mass distributions of the lens $\kappa(x, y)$ given by the sims and generated by the volunteers. To get a means of comparing the sims to the models, the total convergence⁷ $\kappa_{\text{encl}}(r)$ was calculated for both. The Einstein radius Θ_E is defined by $\kappa_{\text{encl}}(\Theta_E) = 1$ and gives a number that allows a rough comparison between a sim and a model. We also let an expert (PS) model three selected systems to compare the results from volunteers to those of a professional.

To compare the enclosed mass profile and the Einstein radius of the simulation and the models, $\kappa_{\text{encl}}(r)$ was calculated using the mass map $\kappa(x, y)$ generated in the modeling process. From the ensemble of models generated by one modeling process, the mean is taken as the resulting $\kappa(r)$ to calculate Θ_E . To estimate the errors, the extremal models are used to estimate a lower and upper limit for Θ_E . These results can be seen in Figure 12. This figure shows that this technique of estimating the error using the ensemble of all models underestimates the error significantly and should be improved for further analysis.

Figure 12 shows that the calculated Einstein radius Θ_E of the models tend to be too high. The overshoot varies from around 0.2 to 0.4 for good models. One of the reasons for this is that it is hard to get the center of the lens on spot. An offset leads to a flatter mass profile for the model compared to the simulation.

Comparing the models from volunteers and experts can be done in Figures 9, 10 and 13, where only the expert got the right configuration (Figure 9) but enclosed masses are all fairly similar. In cases with nearly symmetric image configurations, like this example, it is more difficult to identify the image parities correctly. Incorrect image-parity identification changes the orientation of the mass distribution $\kappa(x, y)$, but $\kappa_{\text{encl}}(r)$ and thus the Einstein radius Θ_E are not influenced that much.

Two additional sims were modeled by an expert, ASW0001hpf and ASW0000vqg. Looking at the results for Θ_E for those models in Figure 12, we conclude that the performance of volunteers (blue crosses) and experts (red crosses, offset) is comparable.

⁷often called enclosed mass

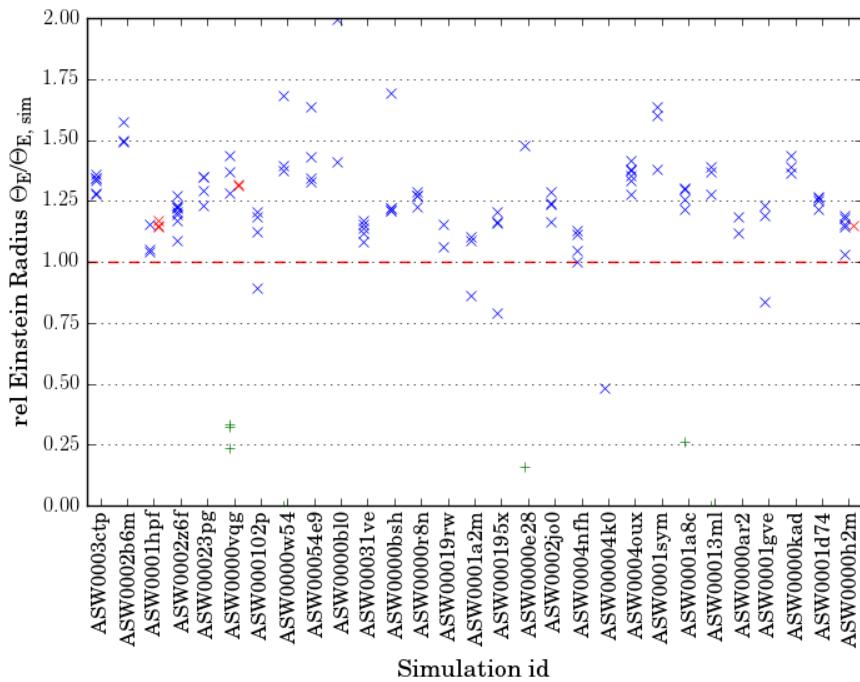


Fig. 12.— Relative Einstein radius $\Theta_E / \Theta_{E, \text{sim}}$ for models by volunteers (blue cross), models made by an expert (red cross with offset), including rejected models (green squares); binned per sim.

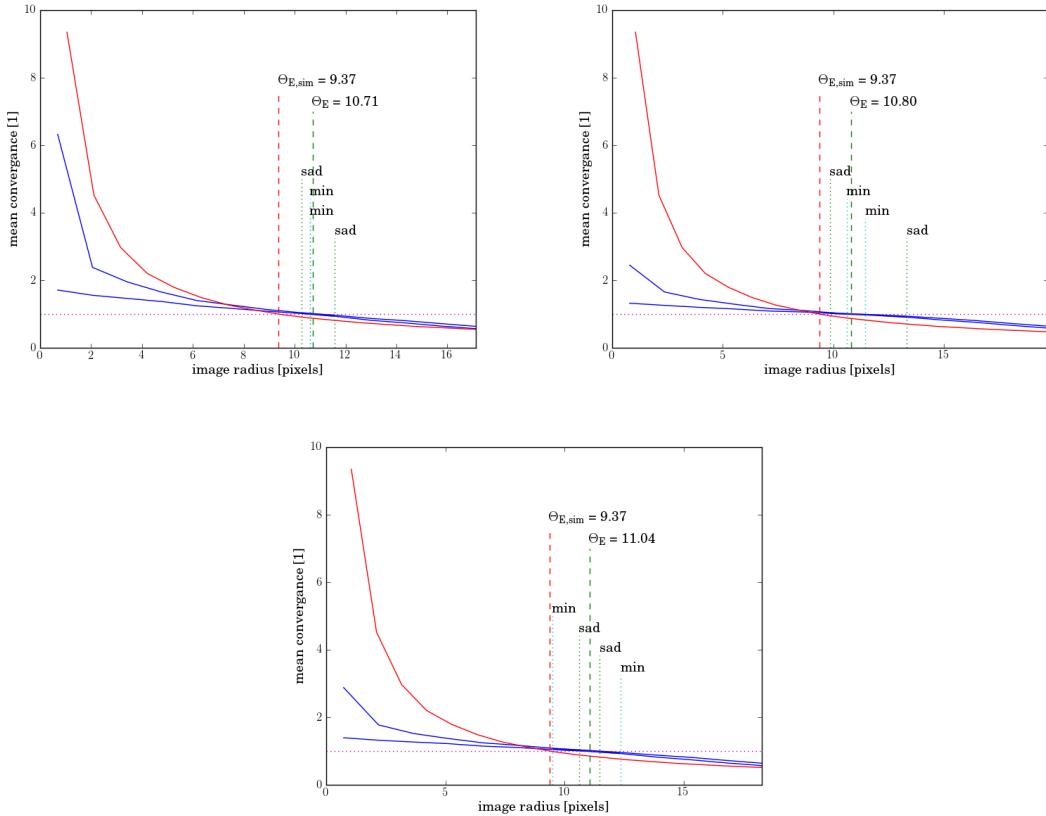


Fig. 13.— Enclosed mass $\kappa_{\text{encl}}(r)$ for models of ASW0000h2m with incorrect image identifications. Figure 9 shows a model (by an expert) with the correct identifications, while Figure 10 shows a model with error E06 (parities rotated by $\pi/2$). Shown here are the $\kappa_{\text{encl}}(r)$ results for two models with error E06 and one model (bottom) with error E05 (parities rotated by π).

4. Outlook

The lens-modeling challenge indicates that the Spacewarps collective is good, not only for identifying lens candidates for follow-up, but modeling candidates as well.

There is, however, plenty of room for improvement.

First, the particular modeling strategy implemented is not the only one possible. SpaghettiLens requires modelers to characterize the overall image structure in abstract terms based on Fermat’s principle, and the placement of mass distributions is done by the computer. In other modeling tools, the user puts down a trial mass distribution and has the machine refine it. A few of these modeling programs have also been designed with citizen science in mind, and would be interesting in the Spacewarps environment.

Second, SpaghettiLens needs some enhancements.

- Currently, SpaghettiLens does not attempt to model the source shape; the user identifies the brightest points on the image, and these are taken as images of a point-like source, whose positions must be reproduced exactly. For generating a synthetic image, a conical source profile is assumed. Fitting for the source profile to optimize resemblance to the observed lensed image after the lens model has been generated, is algorithmically straightforward and planned to be implemented. This would alleviate another problem with SpaghettiLens, which is that there is no numerical figure of merit, and assessment of a model is a judgment call based on the synthetic image, and on whether the mass distribution and the arrival-time surface show suspect features.
- SpaghettiLens has a tendency to somewhat overestimate the Einstein radius (evident from Figure 12), and it is also apparent that the models tend to be too shallow. This possible explanation is that, while the sims are steeply peaked at the centre, the pixellated mass model fixes a comparatively large area near the central at constant density. One way to solve this problem would be to introduce smaller pixels in the central region, thus enabled a steeper centre.
- Another limitation so far in SpaghettiLens is that the lens is assumed to be dominated by one galaxy, which puts most galaxy-group lenses beyond the reach of the modeler. Since complicated group lenses are some of the most interesting candidates present, removing this limitation is most desirable. From the users’ point of view, it would mean that spaghetti contours with more than one maximum can be allowed. For examples, see Figure 5c in (Rusin et al. 2001) and Figure 4b in Keeton & Winn (2003).

- At present, a single false-color composite is used as the data. An option could be added to use all available filters, individually or in combination, at the user wishes.

The third desirable avenue of improvement is to facilitate collaborative work.

- As mentioned above, the option of revising an already-archived model is already available. Desired now are tools for comparing different models of a given system, both visually and through different statistical measures. As evidenced by a current collaborative modeling effort, a particularly interesting candidate can lead to an extended discussion and dozens of models.
- Better tutorial materials are also needed, and this would address some of the problem areas found in the modeling challenge. For example, we saw in Section 3.3 that volunteers are most prone to making errors in two situations: when in identifying an arc-like structure while placing the points, and in identifying the correct ordering of the points in nearly-symmetric configurations.

A. More on Lensing Theory

In Section 2.1, for the sake of a more intuitive explanation, we suppressed some constant factors in equations (1), (2) and (4). Here we fill in the details.

First let us recall various distance formulas. Comoving distances involve integrals of the type

$$\int \frac{dz}{\sqrt{\Omega_m(1+z)^3 + \Omega_\Lambda}} \quad (\text{A1})$$

Specifically, the comoving distance D_S from observer to source, D_{LS} from lens to source, and D_L from observer to lens are given by the above integral, with limits as follows.

$$\begin{aligned} D_S & \quad \int_0^{z_S} \\ D_{LS} & = \frac{c}{H_0} \times \int_{z_L}^{z_S} \\ D_L & \quad \int_0^{z_L} \end{aligned} \quad (\text{A2})$$

Angular-diameter distances d can be related to comoving distances D by scale factors, as follows.

$$\begin{aligned} D_S & = (1 + z_S) d_S \\ D_{LS} & = \frac{1 + z_S}{1 + z_L} d_{LS} \\ D_L & = (1 + z_L) d_L \end{aligned} \quad (\text{A3})$$

Consequently, angular-diameter distances are always smaller than comoving distances. Moreover, while comoving distances can be simply added, angular distances cannot:

$$d_S \neq d_L + d_{LS} \quad (\text{A4})$$

Locations on the lens plane in angular coordinates on the sky, are expressed as

$$(x, y) = d_L(\theta_x, \theta_y) \quad (\text{A5})$$

The A areas (with the implicit constant factor that makes them time delays) and the κ

density are related to physical arrival time t and density Σ as

$$\begin{aligned} A &= \frac{c}{(1+z_L)^2} \frac{D_L D_{LS}}{D_S} \times t \\ \kappa(x, y) &= \frac{4\pi G}{c^2} \frac{D_L D_{LS}}{D_S} \times \Sigma(x, y) \end{aligned} \quad (\text{A6})$$

Letting the source be offset at (s_x, s_y) rather than at the origin, we have

$$\begin{aligned} t_{\text{geom}} &= \frac{(1+z_L)^2}{2c} \frac{D_S}{D_L D_{LS}} ((x - s_x)^2 + (y - s_y)^2) \\ \nabla^2 t_{\text{grav}} &= -(1+z_L) \frac{8\pi G}{c^3} \Sigma(x, y) \end{aligned} \quad (\text{A7})$$

We can now compare with equations (2.1) to (2.6) from Blandford & Narayan (1986).

B. Miscellaneous

Maybe discard Figure 14, since the essential information is there in Figure 12.

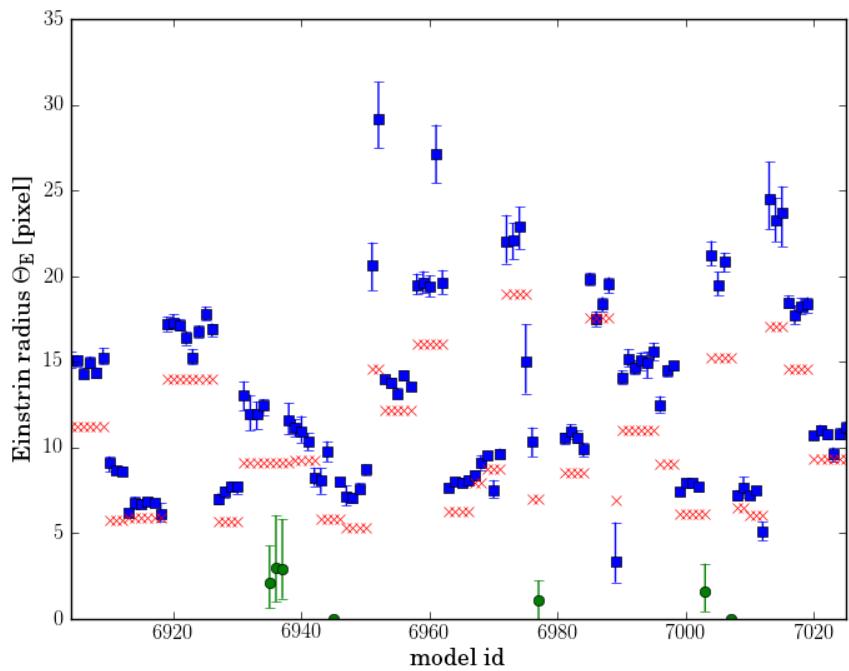


Fig. 14.— Einstein radius Θ_E for all models with estimated errors in blue squares, Θ_E of simulation in red crosses

REFERENCES

- Blandford, R. & Narayan, R. 1986, *ApJ*, 310, 568
- Coles, J. 2008, *ApJ*, 679, 17
- Coupon, J., Ilbert, O., Kilbinger, M., McCracken, H. J., Mellier, Y., Arnouts, S., Bertin, E., Hudelot, P., Schultheis, M., Le Fèvre, O., Le Brun, V., Guzzo, L., Bardelli, S., Zucca, E., Bolzonella, M., Garilli, B., Zamorani, G., Zanichelli, A., Tresse, L., & Aussel, H. 2009, *A&A*, 500, 981
- Gavazzi, R., Treu, T., Marshall, P. J., Brault, F., & Ruff, A. 2012, *ApJ*, 761, 170
- Keeton, C. R. 2001a, ArXiv: astro-ph/0102341
- . 2001b, ArXiv: astro-ph/0102340
- Keeton, C. R. & Winn, J. N. 2003, *ApJ*, 590, 39
- Khatib, F., Cooper, S., Tyka, M. D., Xu, K., Makedon, I., Popovi, Z., Baker, D., & Players, F. 2011, *Proceedings of the National Academy of Sciences*, 108, 18949
- Koopmans, L. V. E., Bolton, A., Treu, T., Czoske, O., Auger, M. W., Barnabè, M., Vegetti, S., Gavazzi, R., Moustakas, L. A., & Burles, S. 2009, *ApJ*, 703, L51
- Leier, D., Ferreras, I., Saha, P., & Falco, E. E. 2011, *ApJ*, 740, 97
- Lubini, M. & Coles, J. 2012, *MNRAS*, 425, 3077
- Marshall, P. J., Hogg, D. W., Moustakas, L. A., Fassnacht, C. D., Bradač, M., Schrabback, T., & Blandford, R. D. 2009, *ApJ*, 694, 924
- More, A., Cabanac, R., More, S., Alard, C., Limousin, M., Kneib, J.-P., Gavazzi, R., & Motta, V. 2012, *ApJ*, 749, 38
- Navarro, J. F., Frenk, C. S., & White, S. D. M. 1996, *ApJ*, 462, 563
- . 1997, *ApJ*, 490, 493
- Paraficz, D. & Hjorth, J. 2010, *ApJ*, 712, 1378
- Rusin, D., Kochanek, C. S., Norbury, M., Falco, E. E., Impey, C. D., Lehár, J., McLeod, B. A., Rix, H.-W., Keeton, C. R., Muñoz, J. A., & Peng, C. Y. 2001, *ApJ*, 557, 594
- Saha, P. & Williams, L. L. R. 1997, *MNRAS*, 292, 148

- . 2001, AJ, 122, 585
—. 2004, AJ, 127, 2604
Williams, L. L. R. & Saha, P. 2000, AJ, 119, 439