PHYSICAL REVIEW D **87,** 123512 (2013)

Robust strong lensing time delay estimation

Alireza Hojjati,[1] Alex G. Kim,[2] and Eric V. Linder[1,2,3]

[1]*Institute for the Early Universe WCU, Ewha Womans University, Seoul 120-750, Korea*
[2]*Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA*
[3]*Berkeley Center for Cosmological Physics, University of California, Berkeley, California 94720, USA*

(Received 17 April 2013; published 14 June 2013)

Strong gravitational lensing of time variable sources such as quasars and supernovae creates observable time delays between the multiple images. Time delays can provide a powerful cosmographic probe through the ''time-delay distance'' involving the ratio of lens, source, and lens-source distances. However, light curves of lensed images have measurement gaps, noise, systematics such as microlensing from substructure along an image line of sight, and no *a priori* functional model, making robust time-delay estimation challenging. Using Gaussian process techniques, we demonstrate success in accurate blind reconstruction of time delays and reduction in uncertainties for real data.

DOI: 10.1103/PhysRevD.87.123512 PACS numbers: 98.80.−k, 95.30.Sf

I. INTRODUCTION

Multiple images of a single source are dramatic evidence for the effect of gravity, specifically general relativity, on light. This strong gravitational lensing not only splits the images, but magnifies or demagnifies the source flux and induces time delays between the images. The time delays arise from both the geometric path differences along the various lines of sight and the gravitational potential differences traversed by the photons.

When the source is variable, such as from a quasar or supernova, the time delays in the flux of one image relative to another can be observed. With careful modeling of the lens mass distribution, and measurement of the angular positions of the images, the geometric factors of distances between observer and lens, observer and source, and lens and source can be extracted as a ratio called the time-delay distance. Recent advances in lens modeling [1,2] and careful, long-term flux monitoring programs such as COSMOGRAIL [3] (also see [4,5]) have matured strong lensing time delays to an incipient cosmographic probe.

This prospect is exciting for several reasons. Since time delays over cosmological distances are sensitive not just to the overall scale, or Hubble constant, but the cosmic energy density and its evolution with redshift, one can constrain (combinations of) the matter and dark energy densities and dark energy equation of state. Moreover, the time delay distance acts fundamentally differently from luminosity and angular distances measured by calibrated standard candles such as type Ia supernovae and rulers such as baryon acoustic oscillations. Hence it has distinct covariances among cosmological parameters and can be powerful in complementarity with the standard distance probes [6,7]. Finally, despite the lens mass modeling, strong lensing time delays are a geometric probe and are tied only to the late universe, unique except for supernovae (but with different systematics and covariances) among all cosmological probes.

Here we address the following important element of the use of lensing time delays: accurate estimation of the actual time delays. While great progress has been made in recent years (see, e.g., [8–10]), in large part due to heroic observing programs and improved data sets, this is not a solved problem. Mathematically, one can consider it as reconstructing a shift between multiple noisy, irregularly sampled, differentially amplified data streams. We apply a special combination of Gaussian process statistics to this task. Such a concept for strong lensing dates back to [11] and more recently has been shown to have reasonable success [8]; we introduce several new features that exhibit noticeable improvement in the state of art.

Section II outlines the challenge of reconstructing the time delays from realistic data complete with systematics such as microlensing. The Gaussian process methodology is described in Sec. III, introducing the various correlation function terms and accounting for systematics. We test the method against blinded mock data, and real data from the literature, in Sec. IV, and conclude in Sec. V.

II. TIME-DELAYED LIGHT CURVES

Fluxes received from an image at several times define a light curve, but the name is misleading since the data are not continuous but discrete, and the observations are often irregular and sparse and have measurement uncertainties. The best monitoring frequency may be every day or two, while long gaps of a few months occur due to seasonal visibility of regions of the sky from a single telescope. The cadence is often irregular, though ongoing wide area surveys such as Dark Energy Survey (DES [12]), Kilodegree Survey (KIDS [13]), and PanSTARRS [14], and in the future LSST [15], may have regular observations with periods of several days.

Apart from the sparseness, the data has photometric measurement noise. Most current observations come from small (1 meter) telescopes, and atmosphere, telescope, and detector

noise all contribute. With wide field surveys, hundreds to thousands of time-delay systems may be found, enabling choice of the cleanest for use as time-delay distance probes. Since to obtain a time-delay distance one must have a robust model of the lens mass distribution, galaxy lenses are preferred over cluster lenses due to less complex modeling. Depending on lens mass and geometry this implies time delays in the range of a few to hundred days in general.

Comparing light curves from different images involves some form of cross-correlation, looking for the time delay between them. Straightforward cross-correlation techniques tend not to work well due to the noisiness and sparseness of the data, and extrinsic contributions (see, e.g., [16]). Instead of comparing noisy data with noisy data, regression techniques attempt to reconstruct the underlying true source variation and compare the image measurements to that. We employ Gaussian processes (GP) as the regression technique. See [17] for an example of its application to (nonlensed) supernova light curves.

In addition to measurement difficulties, astrophysical systematics contribute to the challenge of time-delay estimation. Further time variations arise from microlensing caused by passage of substructure near to the line of sight. This affects images independently, breaking the (delayed) coherence between them, and can occur on all time scales. Short variations just add noise but long term variations disrupt the relation between the light curves for large portions of the data set and so can cause misestimation of the time delay. These long term variations are moderately smooth and some previous work has used low order polynomials or splines to represent them; we instead allow the data to determine their time scale.

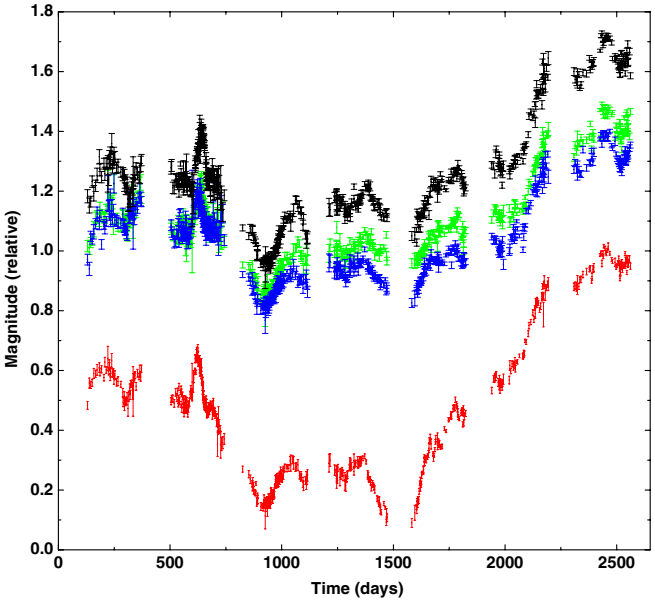

FIG. 1 (color online). Magnitudes (log flux) of four images of the quasar HE 0435 − 1223 are plotted vs time, with an arbitrary overall zeropoint.

Thus we have three elements entering into the light curves: the intrinsic variation that we want to measure, the observational noise, and the astrophysical microlensing systematic (in fact our formalism would allow multiple versions of the last two). The challenge of robust time-delay estimation is to reconstruct phase shifts of a source with unknown intrinsic flux variation, for images with independent microlensing magnifications along their lines of sight, using noisy data with irregular temporal sampling. Figure 1 shows an example of real light curves from four images of quasar HE 0435 − 1223 measured by COSMOGRAIL [18]. Conventionally observations are reported in magnitudes (logarithmic flux units).

III. TIME-DELAY ESTIMATION

For reconstructing an intrinsic function from isolated, noisy data points, Gaussian processes offer a robust, substantially model independent statistical method with well defined error characterization. See [19] for a thorough discussion of GP from a statistical point of view. The basic idea is that the function is not parametrized, but rather the data are fit to a whole family of possible curves, given by a Gaussian distribution with a mean function and a covariance kernel between points.

The key choices are the form of the mean function (which ideally does not affect the final fit but in practice a poor mean function can lead to difficulties) and the covariance kernel, together with any hyperparameters used in those functions. There is a single GP representing the true source underlying all of the images plus the microlensing of a reference image. For a mean function we adopt a constant function, then allow hyperparameters for magnifications relative to the reference image. We try different reference images to test for robustness.

For the covariance kernel we investigate three possibilities. A damped random walk (DRW) is often adopted to model the intrinsic quasar light curve [20–23]. While we are here focused on extracting accurate time delays, not modeling the quasar *per se*, it is natural to try the DRW kernel,

$$k(t_i, t_j) = \sigma^2 e^{-|t_i - t_j|/l}, \tag{1}$$

where t_i and t_j are measurement times, the hyperparameter σ adjusts the amplitude of the kernel, and l functions as a correlation length.

Another choice is a Matern function with index 3/2,

$$k(t_i, t_j) = \sigma^2\left(1 + \frac{|t_i - t_j|\sqrt{3}}{l}\right)e^{-|t_i - t_j|\sqrt{3}/l}. \tag{2}$$

The Matern function is commonly used in statistics [19] and allows for greater roughness in the variation than another common choice, the squared exponential or Gaussian,

$$k(t_i, t_j) = \sigma^2 e^{-(t_i - t_j)^2/(2l^2)}. \tag{3}$$

TABLE I. Blind analysis of time delays works for DRW, Matern, and squared exponential GPs. The input to the simulation had $\Delta t_{AB} = 15.0$ days, $\Delta t_{AC} = 25.0$ days.

Kernel	Δt_{AB}	Δt_{AC}	Δt_{BC}
DRW	14.94 ± 0.14	24.99 ± 0.09	10.0 ± 0.2
Matern	14.3 ± 0.8	25.1 ± 0.9	10.8 ± 0.9
Sq Exp	13.9 ± 1.3	25.8 ± 1.4	10.6 ± 0.7

We will compare the results for these three kernels to give extra cross-checks on the results; generally, we find that DRW works best, once guided by an initial Matern run.

We include measurement noise and an additional nugget term $\sigma_n^2 \delta_{ij}$, which acts as a zero lag dispersion, e.g., an empirical term for misestimated measurement noise or finite realization scatter. This is distinct from the GP amplitude σ in that σ accounts for the global variations of the kernel whereas the nugget term σ_n accounts for the independent dispersion of the individual data points around the predicted GP value.

The microlensing systematic has been attempted to be addressed in the literature by multiplying the light curves by a quadratic polynomial or a cubic spline over short time spans or within an observing season. This restricts the allowed variations and has the potential to lead to bias in the reconstructed time delays or simply a failed fit. We remain within the GP framework, which does not impose a specific model or timescale for the microlensing, and account for the microlensing with a GP for each image (other than the reference one) with zero mean function and a squared exponential kernel of common amplitudes σ_μ^2 and correlation lengths l_μ. To separate the microlensing GP from the quasar GP, we require a long correlation length l_μ (systems with the microlensing timescale comparable to the intrinsic variations are not useful for time-delay measurement). We have investigated various choices of priors, for example, $\pi(l_\mu) > 50$ days, $\pi(l_\mu) >$ season, or $\pi(l_\mu) > 3l$; all give equivalent results.

We emphasize that neither the intrinsic quasar light curve nor the microlensing actually have to be (and may not be) true GPs in themselves; all we want to test is whether robust time delays can be estimated from this approach.

In summary, the light curve predictions for our full GP regression take the form

$$\vec{y}_1 \sim GP_Q(\vec{\theta}_{\rm Qhp}; t - t_1) \quad (4)$$

$$\vec{y}_2 \sim GP_Q(\vec{\theta}_{\rm Qhp}; t - t_2) + GP_{\mu 2}(\vec{\theta}_{\mu \rm hp}) + \Delta m_2 \quad (5)$$

$$\vec{y}_3 \sim GP_Q(\vec{\theta}_{\rm Qhp}; t - t_3) + GP_{\mu 3}(\vec{\theta}_{\mu \rm hp}) + \Delta m_3 \quad (6)$$

and so forth for each image, where $\vec{\theta}_{\rm Qhp}$ is the hyperparameter vector for the quasar GP, $\vec{\theta}_{\mu \rm hp}$ is for the microlensing GP, and Δm represents the magnification relative to the reference image 1.

The GP likelihood is [19]

$$2 \ln p(Y|\vec{\theta}) = -Y^T K^{-1} Y - \ln |K| - N_d \ln 2\pi, \quad (7)$$

where Y is the vector of magnitude data, with N_d the total number of data points, $\vec{\theta}$ represents the fit parameters, e.g., time delays, and K is the full kernel (the sum of the quasar GP, microlensing GP, measurement noise, and nugget) with $|K|$ being its determinant. The likelihood is maximized for the most likely values of the time delays and magnifications, which we find using the function minimizer routine Minuit [24] and have validated using a Monte Carlo analysis.

In principle, we can combine all light curves at once, compare two at a time, or any number of light curves. Simultaneous analysis of more than two curves allows a consistency check in the form of the triangle equality, e.g. $\Delta t_{AC} = \Delta t_{AB} + \Delta t_{BC}$, and is our baseline approach. Using more light curves also has the advantage of the leverage of more images on simultaneously constraining the underlying source light curve. Analysis using just a pair has fewer hyperparameters and may deliver smaller statistical errors, but at the risk of bias. We carry out cross-checks by trying different numbers of light curves in the analysis, finding that the results from the pair analyses can provide useful initial conditions to the simultaneous fit. One can also use portions of data, such as selected observation seasons, to cross-check the consistency of the results or to reduce the impact of microlensing as has been done in the literature before. We find the results from our approach to be robust to the number of data points used in the analysis.

In summary, when fitting N light curves we have the $N - 1$ time-delay parameters that are our goal, the $N - 1$ magnifications Δm, and the hyperparameters σ^2, σ_n^2, σ_μ^2, l, l_μ.

IV. TESTS AND RESULTS

A. Blind mock data

To test the accuracy and robustness of the method we initially created blinded mock data sets. To preserve realistic sampling and data quality, one author took light curve data from one image of quasar HE 0435 − 1223, realized three new light curves using random Gaussian distributions with mean zero and standard deviation equal to the data errors, and shifted each of the resulting light curves vertically by various magnifications and horizontally by time delays. The shifted data were then resampled onto the original time sampling using linear interpolation. Another author, unaware of the simulated time-delay and magnification values, was given the final data points with error bars and carried out the GP fit.

TABLE II. Time-delay estimations are compared between our GP analysis and values in the literature using different reconstruction methods. A question mark represents time-delay estimates not provided by the literature, a ⋯ indicates there is no fourth image.

Kernel	Δt_{AB}	Δt_{AC}	Δt_{AD}	Δt_{BC}	Δt_{BD}	Δt_{CD}
HE 0435 − 1223 GP-DRW	-9.5 ± 0.3	-1.9 ± 0.4	-15.6 ± 0.3	8.1 ± 0.3	-6.0 ± 0.3	-13.6 ± 0.4
HE 0435 − 1223 GP-Mat	-9.6 ± 1.1	-1.5 ± 1.1	-14.0 ± 0.9	8.1 ± 1.1	-5.0 ± 1.1	-12.3 ± 1.1
HE 0435 − 1223 Lit(1) [3]	-8.4 ± 2.1	-0.6 ± 2.3	-14.9 ± 2.1	7.8 ± 0.8	-6.5 ± 0.7	-14.3 ± 0.8
HE 0435 − 1223 Lit(2) [27]	-8.8 ± 2.4	-2.0 ± 2.7	-14.7 ± 2.0	6.8 ± 2.7	-5.9 ± 1.7	-12.7 ± 2.5
WFI J2033 − 4723 GP-DRW	35.1 ± 0.9	-24.9 ± 0.4	⋯	-59.2 ± 2.1	⋯	⋯
WFI J2033 − 4723 GP-Mat	36.0 ± 1.5	-26.3 ± 1.7	⋯	-62.0 ± 2.3	⋯	⋯
WFI J2033 − 4723 Lit [26]	35.5 ± 1.4	$-27.1 + 4.1/-2.3$	⋯	?	⋯	⋯
B1608 + 656 GP-DRW	31.8 ± 2.4	-1.3 ± 1.5	-51.0 ± 6.2	-33.1 ± 2.7	-72.0 ± 4.5	-43.1 ± 3.6
B1608 + 656 GP-Mat	31.7 ± 2.1	-2.4 ± 2.2	-50.4 ± 6.9	-35.0 ± 4.0	-77.5 ± 7.1	-44.4 ± 5.4
B1608 + 656 Lit [28]	$31.5 + 2.0/-1.0$	?	?	-36.0 ± 1.5	$-77.0 + 2.0/-1.0$	?

The results are shown in Table I, with the true values of 15.0 and 25.0 day delays recovered within the 68% confidence level by each of the three covariance functions. Several other tests with different time delays had similar results. The DRW kernel gives results that are significantly more precise, but due to its allowance of high level of variations we find that it works best when we first run a GP with a Matern kernel, and use that result as a prior with 10 times the Matern time-delay uncertainties when running DRW.

We find that the magnification and nugget terms are both important to include. Time delays are also tested for robustness by choosing different reference curves and different multiplicities (i.e. fitting for the AB time delay in isolation, or simultaneously fitting the GP to more than two light curves). Quoted values reflect the central values and uncertainties from the configuration that has the best reduced χ^2 and the smallest errors. These uncertainties are marginalized over all the other parameters and hyperparameters; the distributions are sufficiently Gaussian that the 68% C.L. error bars are symmetric.

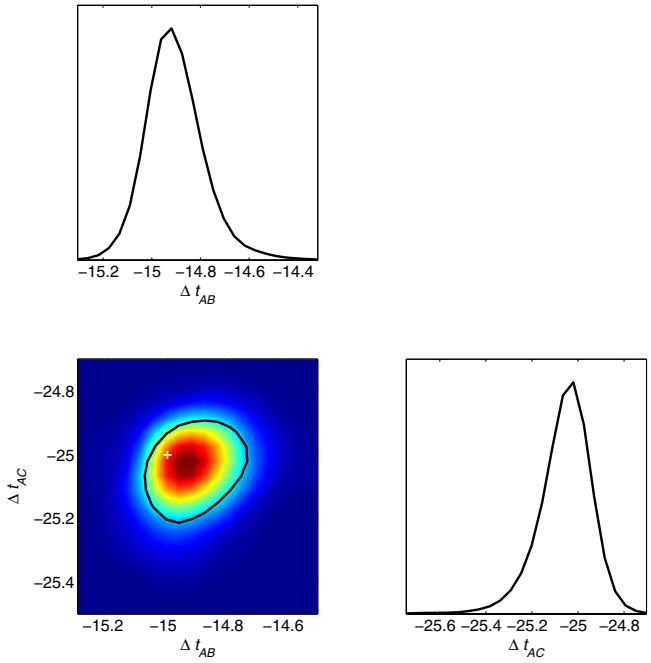

FIG. 2 (color online). Marginalized one-dimensional and two-dimensional likelihood contours are illustrated for the two time delays in the mock data case. The fiducial value is marked with a white plus sign.

Figure 2 shows the one-dimensional and two-dimensional joint likelihood contours for the time-delay parameters in the mock data case using the DRW GP. As a comparison, these results are obtained using COSMOMC [25] as a generic Monte Carlo sampler, and are wholly consistent with the Minuit results. For all the parameters and hyperparameters we impose a very wide flat prior and let data decide their values. The only constraint is on the microlensing correlation length, which as discussed should not be too small and hence mix with the actual correlation length of the GP kernel.

A larger and more sophisticated series of data challenges is forthcoming as part of the LSST Dark Energy Science Collaboration strong lensing working group. This will provide large, sophisticated mock data sets and an excellent opportunity for testing further development of robust time-delay estimation.

B. Actual data

The second part of testing the GP method involves using public data sets from COSMOGRAIL and other literature sources [3,4,26–28] as inputs for time-delay estimation. These results can then be compared to the literature results obtained using a variety of different methods.

We use the two sets of COSMOGRAIL light curves publicly available at [18], for quasars HE 0435 − 1223 and WFI J2033 − 4723, and the radio light curves of quasar B1608 + 656, courtesy of Chris Fassnacht. Table II compares the results we obtain from our GP analysis using the DRW and Matern kernels with those published in the literature. We also have tested the square exponential

kernel but this gives weaker uncertainties. The values from our analysis and the literature are consistent with each other, with the GP analysis tending to have smaller uncertainties. Note the true values of the time delays are not known, but the consistency offers an indication of robustness.

The GP analysis not only estimates the time delays, a key input for cosmography through time-delay distances, but provides information on the intrinsic quasar variability, the variations around the best-fit GP light curve, and the microlensing systematics through the hyperparameters such as the correlation lengths, GP amplitudes, and nugget.

We find that there is no significant correlation between the parameters. The nugget term is usually important and has a value comparable to the errors on the data points. We also find that including the microlensing term is useful even when there is no significant microlensing in the system.

The quasar HE 0435 − 1223 (Fig. 1) has a long observation period with distinct features in the intrinsic variability, making it fairly straightforward to compute the time delays. The bottom curve has significant microlensing variation which leads to large microlensing amplitude σ_μ. The microlensing correlation length ($\sim$700 days) is completely separated from the quasar GP correlation length ($\sim$100 days). There is strong agreement between our results, those of Literature 1 [3] that uses only the first two observation seasons, and those of Literature 2 [27]. Our uncertainties are smaller by a factor of several.

The quasar WFI J2033 − 4723 has a relatively shorter observation time but distinct features in the light curves. There is no significant long-range microlensing and hence σ_μ is very small indicating that including microlensing terms may not be necessary (but this is not known *a priori*).

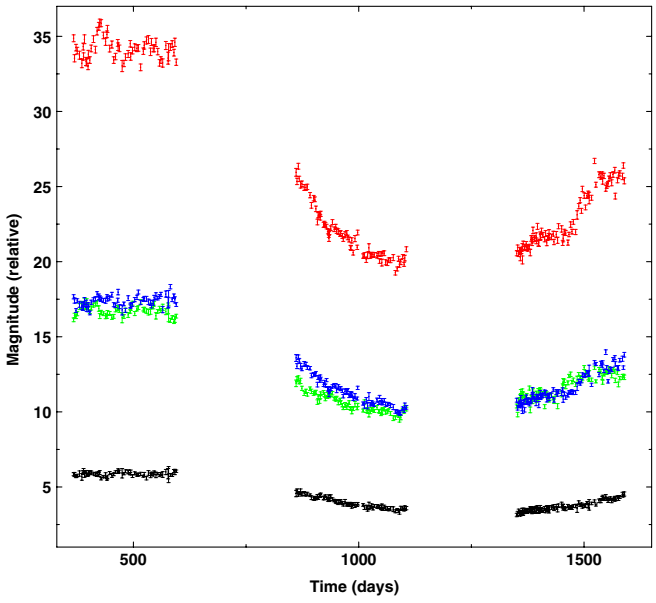

FIG. 3 (color online). Magnitudes (log flux) of four images of the quasar B1608 + 656 are plotted vs time, with an arbitrary overall zeropoint.

Again, despite using several hyperparameters, our marginalized uncertainties are smaller than the results from [26].

The quasar B1608 + 656 (light curves shown in Fig. 3) is an example of a challenging system with large data gaps, relatively small intrinsic variability, and significant microlensing, all of which make it hard to estimate the time delays of its images. While we have successfully derived the time delays between all the images, including the cases not presented in the literature, the error bars are relatively large. This is in part due to the featureless light curves (especially the bottom curve, D in Table II, which is almost flat) and also due to the fact that our errors are marginalized over other parameters. For example, fitting the nugget term increases the errors by at least a factor of two, while its presence is relatively unimportant for this system. We find that the probability distributions for some of the time delays in the DRW case have a smaller secondary peak, so comparison with the Matern results is useful to ensure robustness.

V. CONCLUSIONS

Accurate estimation of strong lensing time delays is an essential element in the use of time-delay distances as a novel cosmological probe. The complementarity, substantially geometric nature, and disjoint systematics of this technique make its use a goal worth striving for.

We have explored Gaussian processes as a regression method that is effectively model independent and we demonstrated robust results for both blind mock data and actual literature data, in many cases reducing the uncertainties of the time-delay estimations. Noisy data, gaps in the observations, and extrinsic microlensing variations can all be handled by the method.

Robustness arises not just from the technique itself, but the ability to use multiple light curves simultaneously, and test results against different combinations. Several possibilities exist for further improvement. For example, one could weight the estimations derived from different combination of curves or one could remove unnecessary hyperparameters to reduce estimation uncertainty while checking that the best fit does not shift.

Future data challenges will provide an opportunity to further develop the technique, providing important training and assessment of the reconstruction method. And of course one could obtain better real data. Forthcoming surveys will find many more suitable lensing systems, allowing choice of the cleanest or best observed (with low photometric uncertainties, better cadence with fewer gaps, etc.).

While time-delay estimation is just one element in the development of strong lensing distances as a new cosmological probe, its improvement is key to this promising technique for mapping the Universe. Future work includes applying our GP reconstruction method to studies of lensed supernovae or other variable sources.

ACKNOWLEDGMENTS

We thank Chris Fassnacht for providing light curve data, Chris Kochanek for discussions on the DRW approach to intrinsic light curves, and Arman Shafieloo for discussions on GP code methodology. A. H. acknowledges the Berkeley Center for Cosmological Physics for hospitality. This work has been supported by World Class University Grant No. R32-2009-000-10130-0 through the National Research Foundation, Ministry of Education, Science and Technology of Korea and the Director, Office of Science, Office of High Energy Physics, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

[1] M. Oguri, Astrophys. J. **660**, 1 (2007).
[2] S. H. Suyu, P. J. Marshall, R. D. Blandford, C. D. Fassnacht, L. V. E. Koopmans, J. P. McKean, and T. Treu, Astrophys. J. **691**, 277 (2009).
[3] F. Courbin *et al.*, Astron. Astrophys. Suppl. Ser. **536**, A53 (2011).
[4] S. H. Suyu, P. J. Marshall, M. W. Auger, S. Hilbert, R. D. Blandford, L. V. E. Koopmans, C. D. Fassnacht, T. Treu, and Astrophys. J. **711**, 201 (2010).
[5] R. Fadely, C. R. Keeton, R. Nakajima, and G. M. Bernstein, Astrophys. J. **711**, 246 (2010).
[6] E. V. Linder, Phys. Rev. D **70**, 043534 (2004).
[7] E. V. Linder, Phys. Rev. D **84**, 123529 (2011).
[8] M. Tewes, F. Courbin, and G. Meylan, arXiv:1208.5598.
[9] M. Tewes *et al.*, arXiv:1208.6009.
[10] S. H. Suyu *et al.*, Astrophys. J. **766**, 70 (2013).
[11] W. H. Press, G. B. Rybicki, and J. N. Hewitt, Astrophys. J. **385**, 404 (1992).
[12] Dark Energy Survey Collaboration, arXiv:astro-ph/0510346.
[13] J. T. A. de Jong, G. A. Verdoes Kleijn, K. H. Kuijken, and E. A. Valentin, Exp. Astron. **35**, 25 (2013).
[14] N. Kaiser *et al.*, Proc. SPIE Int. Soc. Opt. Eng. **7733**, 77330E (2010).
[15] LSST Dark Energy Science Collaboration, arXiv:1211.0310.
[16] L. V. E. Koopmans, A. G. de Bruyn, E. Xanthopoulos, and C. D. Fassnacht, Astron. Astrophys. **356**, 391 (2000).
[17] A. G. Kim *et al.*, Astrophys. J. **766**, 84 (2013).
[18] CosmoGrail Collaboration, http://www.cosmograil.org.
[19] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning* (MIT Press, Cambridge, MA, 2006).
[20] B. C. Kelly, J. Bechtold, and A. Siemiginowska, Astrophys. J. **698**, 895 (2009); Astrophys. J. **732**, 128(E) (2011).
[21] Y. Zu, C. S. Kochanek, and B. M. Peterson, Astrophys. J. **735**, 80 (2011).
[22] Y. Zu, C. S. Kochanek, S. Kozlowski, and A. Udalski, Astrophys. J. **765**, 106 (2013).
[23] R. Andrae, D.-W. Kim, and C. A. L. Bailer-Jones, arXiv:1304.2863.
[24] Minuit, http://seal.web.cern.ch/seal/work-packages/mathlibs/minuit/home.html.
[25] A. Lewis and S. Bridle, Phys. Rev. D **66**, 103511 (2002).
[26] C. Vuissoz *et al.*, Astron. Astrophys. **488**, 481 (2008).
[27] C. S. Kochanek, N. D. Morgan, E. E. Falco, B. A. McLeod, J. N. Winn, J. Dembicky, and B. Ketzeback, Astrophys. J. **640**, 47 (2006).
[28] C. D. Fassnacht, E. Xanthopoulos, L. V. E. Koopmans, and D. Rusin, Astrophys. J. **581**, 823 (2002).