Research Statement

Nan Li

Postdoctoral Researcher at
The University of Chicago and Argonne National Laboratory
KICP Associate Fellow

Current Research

Gravitational lensing effects can be observed on all scales in the Universe: weak lensing on the scale of Mpc, strong lensing on the scale of kpc, and microlensing on the scale of pc. Applications of lensing in astrophysics and cosmology specifically include a broad range of problems, e.g., reconstructing the lens mass distribution, identifying faint or dark substructures, detecting galaxies at high redshift, measuring the Hubble constant through time delays, mapping the large-scale mass field, and constraining cosmological parameters. Gravitational lensing is broadly applicable, and is now a critical element of the astrophysics toolkit.

Cosmological strong gravitational lensing, in particular, probes the properties of the dense cores of dark matter halos over decades in mass and offers the opportunity to study the distant universe at flux levels and spatial resolutions that are otherwise unavailable. Strongly lensed *variable* sources offer yet further scientific opportunities. A major challenge in realizing the potential of strong lensing is in understanding the statistical context of both individual systems that receive extensive follow-up study, and the larger samples of strong lenses that are now emerging from survey efforts, This problem is rapidly becoming significant, with DES now producing extensive data, and LSST, Euclid, and WFIRST coming. Motivated by this challenge, my collaborators and I have developed an image-simulation pipeline to generate realistic strong gravitational lensing signals from group- and cluster-scale lenses. This project has been the primary focus of my work in the past two years, and is the genesis of my research ambitions for the next few years.

The core simulation pipeline is PICS — Pipeline for Images of Cosmological Strong-lensing [1]. The goal is to produce simulated images that are indistinguishable (even to the eye, expert or otherwise) from real observations in a variety of imaging surveys. PICS uses a low-noise and unbiased density estimator based on Delaunay Tessellations to calculate the density field from particle distributions from N-body (or hydro) simulations, and lensed images are produced by ray-tracing images of actual galaxies from deep Hubble Space Telescope observations. Other galaxies, similarly sampled, are added to fill in the light cone. The pipeline adds lens member galaxies and foreground stars onto the lensed images. The entire image ensemble is then observed using a realistic point spread function, including bright stars with appropriate detector artifacts. Finally, noise is added, including non-gaussian elements such as noise window-paning from mosaiced observations, residual bad pixels and cosmic rays, and the like.

The current priority application of the pipeline is strong lensing arc statistics. It is well known that the frequency of strongly lensed arcs on the sky reflects the abundance, concentration, and astrophysical properties of massive lenses[2, 3], and new galaxy cluster catalogs from the South Pole Telescope offer an unprecedented opportunity to test simulations against reality.

To further understand the factors affecting arc statistics in the SPT sample, I produced $\sim 10^5$ simulated images by applying PICS to the Argonne Outer Rim simulation, one of the largest simulations currently available. The goal is to assess, across the full redshift range of the SPT cluster catalog, how important secondary projected halos are to the total strong lensing cross section. Mulitple-lens-planes with multiple-source-planes and single-lens-plane with multiple-source-plane are both enable in the PICS pipeline, allowing a thorough exploration of the importance of additional halos. I am working on comparing not only the absolute numbers of lenses in the cases of multiple-lens-planes and single-lens-plane but also the normalized redshift distributions of both cases.

Finally, the statistical properties of lenses (including the cases of multiple-lens-planes and single-lens-plane), with a given threshold of observability, will be compared to that of the optical follow-up results from the South Pole Telescope (SPT) cluster catalog. The observational work was started by my collaborators in mid-2015, and will be finished by mid-2017. Furthermore, we will compare the statistics of other lensing features to that of the follow-up results as well, for instance, the distribution of length to width ratios, radial distances, and azimuths of lensed arcs; the focus of this analysis will be guided by the observational results, which will inform how 'observable' various possible statistical measures are in practice.

An auxiliary use for the pipeline includes testing the preservation of morphological measurements, such as the Gini coefficient, under strong gravitational lensing. In particular, as the new generation of space-based large survey telescopes come online, such as the Wide-Field Infrared Survey Telescope (WFIRST), the number of observed strong lensing systems is expected to expand into the thousands. Such systems provide more detailed views of the internal structure of galaxies at higher redshift than would otherwise be possible; however, the challenge is extracting useful morphological information from such data. It will be necessary to develop morphological measurements that are conserved under gravitational lensing and some elements of the pipeline described above have been used to test the reliability of image plane measurements of the Gini coefficient [4, 5].

The pipeline can also be applied in the weak lensing regime. For instance, I am currently investigating the concentration-mass relation of the halos in the "Q-Continuum Simulation"[6]. The goal is to model the bias between the 2D concentration-mass relation measured via galaxy-galaxy weak lensing and the 3D concentration-mass relation measured via the mass measured from the particle data from the simulation directly. To obtain the 2D concentration-mass relation, I am producing hundreds of thousands of simulated shear maps of halos with the mass range $10^{12}\ M_{\odot}$ to $10^{15}\ M_{\odot}$. The model will be utilized for calibrating the results of mass reconstruction of cluster samples in next generation surveys.

Research Plan

With the capabilities of next-generation telescopes, first with LSST, and then Euclid and WFIRST, astrophysics and cosmology are stepping into the *big data* era. There will be billions of objects and thousands or tens of thousands of strong lenses. To explore the enormous datasets, we must develop sufficiently accurate and efficient pipelines to identify objects and extract features from the observations, for instance, galaxy-scale strong lenses with distorted galaxies or time delays from multiple lensed supernovae and quasars. Regarding the accuracy and reliability of data analysis pipelines, realistic simulations are necessary; such simulations are critical for understanding both uncertainties and systematic bias in the analysis procedures. Strong lensing presents as a subtle and complex signal, and so simple object-finding algorithms are insufficient for defining samples. To find and study strong lenses in these upcoming big data surveys, one can adopt either crowdsourcing or machine learning, or both. Combining the advantages of humans and machines is likely the best way to take on these challenges.

Machine Learning and Gravitational Lensing

So far, the majority of lens finding is based on human effort because human eyes are still the best tool to identify the subtle and complex shapes of lensed images of extended sources (galaxies). Currently, the number of observations is tolerably large ($100 \sim 1000$ lensing systems) for humans to examine, however, next-generation imaging surveys will produce billions of images. It is impossible for only a few people to find all the lensing systems in such enormous datasets. There are two ways to deal with this situation. One is to ask for help from citizen scientists, such as is accomplished in *SpaceWarps*[1]. The alternative is to design an automated approach based on machine learning, a potentially powerful tool to perform pattern recognition. I am working on applying these techniques to improve algorithms for automated

[1] `http://spacewarps.org`

lens finding. Likewise, there is a similar problem of extremely large datasets in lens modeling (tens of thousands of systems, where currently even single lens models are published as entire papers that represent significant researcher effort). I expect that machine learning can eventually lead to automated lens modeling as well.

Interactive Lens Modeling

Lens modeling is a technique to reconstruct the mass distribution of lenses. It is a critical processing step for converting lensing signals to useful constraints for astrophysics and cosmology. As with other regression problems, setting a proper starting point for lens modeling is the key issue. I am developing an interactive Web-APP, named *Hoopla*[2] for estimating appropriate starting points of lens modeling procedures. This tool would be useful for lens modelers to save time by finding appropriate initial conditions, and will help to improve the efficiency of the lens modeling process. Furthermore, it is not only designed for professional lens modelers but also for citizen scientists. I am working to embed a lightweight MCMC sampler into Hoopla to convert a roughly guessed model to an accurate model so that citizen scientists can model detected lensing systems by dragging and clicking only, without knowing any sophisticated mathematics or physics. When complete, Hoopla will extend into the citizen scientists context.

Simulations of Time Delays in Galaxy Scale Lensing

Simulations of galaxy-scale lensing systems is one of my high-priority projects in the future, specifically the simulations of lensing time delays in galaxies. I am a member of the LSST Dark Energy Science Collaboration (LSST-DESC), and working on simulations with the LSST-DESC strong lensing working group. The **Twinkles**[3] project is one of my focus areas. My interests and responsibilities are: (a) Using CatSim and GalSim to produce the images of host galaxies of point sources (quasars and supernovae) and line of sight galaxies; (b) According to the mass distributions of lenses, applying ray-tracing to generate lensed images of the host galaxies of qusars and supernovae; (c) Beyond inputting fits files into PhoSim (a photon ray-tracing code), designing a new approach to parameterize strongly lensed galaxies. The purpose of this simulation is to test a number of software instruments that will be applied to make high accuracy cosmological measurements of type Ia supernovae and strong gravitational lensing time delays with LSST data, e.g., the pipeline for using strong lensing time delays to constrain the equation of state for dark energy.

References

[1] N. Li, M. D. Gladders, E. M. Rangel, M. K. Florian, L. E. Bleem, K. Heitmann, S. Habib, and P. Fasel. PICS: Simulations of Strong Gravitational Lensing in Galaxy Clusters. , 828:54, September 2016.

[2] M. B. Bayliss, M. D. Gladders, M. Oguri, J. F. Hennawi, K. Sharon, B. P. Koester, and H. Dahle. The Redshift Distribution of Giant Arcs in the Sloan Giant Arcs Survey. , 727:L26, January 2011.

[3] B. Xu, M. Postman, M. Meneghetti, S. Seitz, A. Zitrin, J. Merten, D. Maoz, B. Frye, K. Umetsu, W. Zheng, L. Bradley, J. Vega, and A. Koekemoer. The Detection and Statistics of Giant Arcs behind CLASH Clusters. , 817:85, February 2016.

[4] M. K. Florian, N. Li, and M. D. Gladders. The Gini Coefficient as a Morphological Measurement of Strongly Lensed Galaxies in the Image Plane. *ArXiv e-prints*, November 2015.

[5] M. K. Florian, M. D. Gladders, N. Li, and K. Sharon. The Gini Coefficient as a Tool for Image Family Idenitification in Strong Lensing Systems with Multiple Images. , 816:L23, January 2016.

[6] K. Heitmann, N. Frontiere, C. Sewell, S. Habib, A. Pope, H. Finkel, S. Rizzi, J. Insley, and S. Bhattacharya. The Q Continuum Simulation: Harnessing the Power of GPU Accelerated Supercomputers. , 219:34, August 2015.

[2] `http://linan7788626.github.io/pages/Hoopla/index.html`
[3] `https://github.com/DarkEnergyScienceCollaboration/Twinkles`