

UC Santa Cruz

UC Santa Cruz Electronic Theses and Dissertations

Title

Prospects for Finding Primordial Black Holes with the Rubin Observatory

Permalink

<https://escholarship.org/uc/item/9g81m0j9>

Author

Wood, Duncan Gerner

Publication Date

2025

Copyright Information

This work is made available under the terms of a Creative Commons Attribution-NonCommercial License, available at <https://creativecommons.org/licenses/by-nc/4.0/>

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA
SANTA CRUZ

**PROSPECTS FOR FINDING
PRIMORDIAL BLACK HOLES
WITH THE RUBIN OBSERVATORY**

A dissertation submitted in partial satisfaction of the
requirements for the degree of

DOCTOR OF PHILOSOPHY

in

PHYSICS

by

Duncan Wood

March 2025

The Dissertation of Duncan Wood
is approved:

Professor Steven Ritz, Chair

Professor Connie Rockosi

Professor Robert Johnson

Professor Jason Nielsen

Peter Biehl
Vice Provost and Dean of Graduate Studies

Copyright © by

Duncan Wood

2025

Table of Contents

List of Figures	vi
Abstract	xv
Dedication	xvi
Acknowledgments	xvii
1 Introduction	1
1.1 Cosmology	2
1.1.1 Λ CDM	4
1.1.2 Dark Matter	7
1.2 Vera C. Rubin Observatory	9
1.2.1 Scientific goals	9
1.2.2 Camera design	11
1.2.3 Data management	12
1.3 Primordial Black Holes as Dark Matter	13
1.3.1 Production methods	14
1.3.2 Existing constraints	14
1.3.3 Rubin’s potential contributions	17
1.4 Microlensing	18
1.4.1 Derivation of microlensing curve from General Relativity	18
1.4.1.1 Null geodesics of the Schwarzschild metric	18
1.4.1.2 Deflection angle	21
1.4.1.3 The lens equation	24
1.4.1.4 Magnification	26
1.4.1.5 Lightcurves	27
1.4.1.6 Astrometric shifts	28
1.4.2 Parallax effects	28

2	Camera work	32
2.1	Charge Coupled Devices (CCDs)	32
2.1.1	Detection and integration of photons	34
2.1.2	Readout operation	35
2.1.3	Bias correction	37
2.2	Multi-modal behavior of LSST Camera amplifiers	40
2.2.1	Bias shifts	42
2.2.2	Bias shift detection algorithm	45
2.2.3	Association with bright pixels	46
3	Projected constraints on primordial black holes with the Legacy Survey of Space and Time	52
3.1	Modeling the Milky Way	53
3.1.1	LSST SIM Catalog	53
3.1.2	Milky Way density profile of dark matter	54
3.2	Generating microlensing events	55
3.2.1	Simulation with PopSyCLE	55
3.2.2	Monte-Carlo approach	56
3.2.3	Estimating the density of events at one pointing	57
3.2.3.1	Using differential rates as Monte Carlo probabilities	58
3.2.3.2	Implementation in Python	60
3.2.3.3	Resource usage/complexity comparison to simulation	62
3.2.4	Extension to generic cadences	63
3.2.4.1	rubinsim for simulating proposed LSST strategies	63
3.2.4.2	MicrolensingMetric for detection efficiency	64
3.3	Results	65
4	A novel microlensing detection algorithm	72
4.1	Time series analysis	72
4.1.1	Evenly sampled	73
4.1.2	Unevenly sampled	75
4.2	Precursor dataset: NOIRLab Source Catalog	78
4.2.1	Compatibility with Rubin lightcurves, criteria on measurement frequency	81
4.2.2	Footprints of objects in consideration	83
4.3	Background Characterization for LSST	85
4.4	A new microlensing discovery algorithm: Weighted Overlapping Outlier Detection	88
4.4.1	Algorithm specification	92
4.4.2	Table of default parameters	96

4.4.3	Examples of detections and near detections, descriptions of parameters	97
4.4.4	Extended rejection techniques	99
4.4.4.1	More sophisticated cuts on fit parameters	104
4.4.4.2	Moving baselines	104
4.4.4.3	Repeat offenders	106
4.4.4.4	Achromaticity	106
4.4.4.5	Symmetry	107
4.4.4.6	Dynamic context	107
4.5	Detection efficiency with synthetic microlensing events	108
4.5.1	Ensuring consistency between injected and detected events	110
4.5.2	Exclusions for different masses of PBH	112
5	Results and Discussion	119
5.1	Anticipated benefits and challenges with Rubin data	119
5.2	Limitations of the NSC precursor data	121
5.2.1	Pointings	121
5.2.2	Timescales	121
5.3	Rubin’s sensitivity through the years	123
5.4	Future work	125
5.4.1	Finding Stellar Origin Black Holes	125
5.4.2	Search other datasets with the these methods	126
A	Search algorithm implementation notes	128
A.1	Windowing for WMA and WMS	128
	Bibliography	132

List of Figures

- 1.1 A composite of existing PBH constraints as a fraction of Dark Matter, taken from [1]. The red regions are from gravitational wave experiments, and the blue regions are from other measurements. Dashed lines show projected sensitivities from upcoming searches. 15
- 1.2 A schematic diagram of the lensing geometry. The upper path has deflection angle α_- , image location θ_- , and closest approach radius $r_{\min-}$, and the lower path has α_+ , θ_+ , and $r_{\min+}$. Only one parameter is shown per pair to maintain visibility. In reality, α_{\pm} is an extremely small angle, and the objects are very nearly aligned. The dashed lines are the null geodesics between the observer and source. The images will appear at the angle θ_{\pm} away from the lens, and the true position of the source is β away from the lens. . . . 25

1.3	A point source point lens microlensing curve with parallax (red) and without parallax (blue), from a bulge source and a $0.1M_{\odot}$ lens halfway to the bulge from Earth, with an atypically ($\sim 4\sigma$) long crossing time of 370 days. The left shows the trajectory in impact parameter \mathbf{u} , and the right shows the corresponding change in magnitude over time, given in Modified Julian Days (MJD).	31
1.4	A median timescale event of ~ 6 days from a $0.1M_{\odot}$ PBH, showing the marginal effect of parallax on the typical event.	31
2.1	The bucket brigade operation of charge shifting in a CCD.	36
2.2	Visualizing the changes in bias level in response to different sequencer settings. Each datapoint is derived from one image, where the x and y positions are the mean of the bias level over all segments associated with one of the two ASPIC chips. Multimodal structure persists in different ways through all tested settings.	41
2.3	A sudden shift in bias from the amplifier in an LSST Camera CCD, seen as a brightness change in row 67. This example is not associated with any bright pixels. Although no photons are detected here, they would be measured with gain of $O(1)$ ADU/photon, making the bias level $\sim 25k$ times higher than the marginal contribution of a photon.	43

2.4	Figure 2.3 as a time series near the shift. This is different from the raw image as the parallel (left-right) structure has been removed. The blue dots show the individual pixel values, and the green lines are a moving average.	44
2.5	Visualizing the kernel used in bias shift detection.	47
2.6	Bias shifts associated with a bright pixel. The vertical stripes are parallel structure.	48
2.7	Distribution of the fraction of images with shifts for all amps in the focal plane. A few problematic sensors dominate the number of shifts.	49
2.8	Shiftest sensors - distribution of the amplifiers (equivalently, CCD segments) exhibiting the most bias shifts, partitioned by sensors on the x-axis and segments in different colors.	50
2.9	Distribution of shift locations and magnitudes for a sensor with shifts in many different rows in different images. Only two segments display shifts.	51
2.10	Distribution of shift locations and magnitudes for a sensor with shifts in nearby rows, although not perfectly consistent.	51

- 3.1 Demonstrating the stability of the total on-sky rate calculation estimated from using small subsets of the available sources. A “full” rate is calculated with 336265 sources, corresponding to the number of expected events for a particular simulation (though any number could have been used here). Subsets of these sources were taken with sizes specified by the x-axis, and 200 random subsets were taken for each size. The central line is the (signed) fractional difference between the average total rate of these 200 subsets and the full rate, and the error bars are the average fractional distance of each subset from the full rate. Since there are many other systematic and modeling uncertainties in this simulation, $\lesssim 10\%$ error would be acceptable. The errors from the subset sizes used here are considerably better. 61
- 3.2 Showing the $t_E \propto M^{1/2}$ relationship in simulated events, without considering detection efficiency. Events were cutoff below 1 hour and above 20 years. For easier comparison, total number of events is normalized across different masses, though in reality the rate varies. Color corresponds to the density of events at a crossing time. 68

3.3	Total number of in-principle detectable microlensing events across the sky for populations of delta-function mass distribution PBHs comprising 100% of Milky Way dark matter. This includes sources from the Magellanic Clouds, but not PBH lenses. Including a subhalo with these lenses would increase the low-mass sensitivity, although these events are mostly too fast for Rubin’s cadence. . . .	69
3.4	Same as 3.3, excluding the most distant sources corresponding to the Magellanic Clouds.	70
3.5	Absolute theoretical detection efficiency from the <code>MicrolensingMetric</code> in <code>opsim</code>	71
3.6	Comparison of theoretical detection efficiency for difference cadences, normalized to the <code>baseline_v3.6</code> run. Each alternative survey strategy has PBH masses with modest 1 – 4% increases in detection efficiency. This is particularly pronounced in the low-mass ($10^{-2}M_{\odot}$) range, which are centered at lower crossing times.	71
4.1	Visualizing the windowed approach to computing the convolution of an unevenly sampled dataset (y_i, t_i)	78
4.2	Spatial distribution of NSC sources on the sky, with galactic center at (0,0)	84

4.3 Spatial distribution of a random subset of LSST SIM sources in galactic coordinates. In contrast to the NSC data, there will be many galactic sources sampled in the densest regions. The small patches in the lower left are the Large and Small Magellanic Clouds. 84

4.4 The Weighted Moving Average and Weighted Moving Scatter computed continuously for synthetic data. WMS is visualized as an envelope around the WMA. This continuous version is not computed during the analysis, but is a generalization of the discrete formula described below, found by substituting $\kappa_{ij} \rightarrow \kappa(t - t_j)$, giving functions $A(t)$ and $S(t)$. The discrete version is trivially equivalent to evaluating the functions at the same times as the datapoints, t_j . One must avoid interpreting the envelope as a confidence interval on the moving average. It is only a measure of the consistency between nearby measurements. 91

4.6 A background event, rejected for a poor fit (high condition number) to the PSPL curve in Round 2. 101

4.7 Even without particularly dense sampling, a good fit is inferred to this synthetic injected microlensing curve with $t_E = 40$ days and $u_{\min} = 1$ 102

4.8	A lightcurve with a consistent bright excursion (away from 0), but not enough context to constrain the fitting procedure, resulting in a high condition number.	103
4.9	$1M_{\odot}$ synthetic events and real backgrounds plotted by the condition number and p-value of their fits to the PSPL lightcurve in Round 2. The color histogram shows the density of the large population of synthetic events. One could draw a curve delineating the two populations, while preserving $> 90\%$ of synthetic events, at the risk of overfitting and generalizing poorly. Note that this plot includes all events that pass Round 1, including those synthetic events with poorly inferred curve parameters.	105
4.10	The closest background to acceptance, seen as the furthest top-left red point in Figure 4.9. Just the excursion is shown on top, and the full set of datapoints below.	109
4.11	Before performing Round 2 cuts on fit quality, we can see a strong correlation between distance of the inferred peak time from actual peak time with the condition number on the fitting procedure. . .	113

4.12 Validation of inferred microlensing peak time versus injected peak time after the Round 2 cut on condition number. Population standard deviation is a distance $\sim 5\%$ of the crossing time, FWHM of 3.6%. 114

4.13 Validation of inferred microlensing crossing time versus injected crossing time after the Round 2 cut on condition number. 115

4.14 95% confidence exclusions for LSST, given a number of detections corresponding to the upper bound of a background rate in NSC data, assuming they are log-evenly distributed in crossing time. There's no reason to expect backgrounds this high, but it's impossible to know what will be found in the novel LSST dataset. But even in the worst case, there will be some exclusion power in the slightly sub-solar mass range. 117

4.15 95% confidence exclusions for LSST, given 0 detections over ten years, shown in green. Exclusions are superimposed on a collection of existing bounds from [1]. Despite the strongest exclusions coming from $\sim 10^{-1}M_{\odot}$, the strongest “new” exclusions are in the $1 - 100 M_{\odot}$ range. In any case, LSST provides independent coverage to areas which previously were only excluded by gravitational wave measurements. The dashed region around $10M_{\odot}$ is due to the expected background of Stellar Origin Black Holes, which is described in Section 4.3. Although it will be difficult, if not impossible, to distinguish these from PBH’s of the same size with these methods, Rubin should provide novel insights into this currently poorly constrained population. 118

5.1 PBH exclusions over the course of the LSST, with time on a log scale and mass in units of solar masses. 124

Abstract

Prospects for Finding
Primordial Black Holes
with the Rubin Observatory

by

Duncan Wood

New tools are presented to generate simulated catalogs of microlensing events in the Milky Way from populations of primordial black holes tracing the dark matter halo. These Monte Carlo methods are orders of magnitude faster than the state-of-the-art simulations, and reduce the computational requirements from a large computer cluster to a laptop for full-sky surveys.

A new statistic and method is demonstrated for highly efficient detection of microlensing events in multi-color star surveys. The background filtering is sufficiently strong to reject all lightcurves in a subset of NOIRLab Source Catalog data, while maintaining high efficiency on simulated events injected on the same data.

These insights are combined to predict the exclusions on PBH dark matter the Legacy Survey of Space and Time will create over the ten-year survey.

Dedicated to my parents, in return for their lifelong dedication to my health and happiness.

And in memory of Zach, who dared me to live up to his overestimation of my abilities.

Acknowledgments

Thank you to my advisor Steve Ritz for supporting me all the way from the Summer before my first term at UCSC. I am lucky to have found someone else who values breadth of study in addition to depth.

Thank you to my defense committee: Connie Rockosi, Jason Nielsen, and Robert Johnson, for the substantial labor involved in critiquing a dissertation.

Thank you to everyone in the UCSC Physics Department and SCIPP for creating a supportive and fun environment to learn. In particular, thank you to Michael Dine, Stefania Gori, and Wolfgang Altmannshofer for your letters and enjoyable QFT courses, and thanks to Jeremy Parker, Ben Miller and Amy Radovan for your administrative support.

Thank you to the LSST Camera Team, for your helpful comments and encouragement over the years, especially Seth Digel, Andy Rasmussen, Yousuke Utsumi, Jim Chaing, and Aaron Roodman.

I couldn't be any happier with the quality and quantity of rich friendships in my life. They pulled me away from my computer when I needed a break, brought me food when I worked through meals, and kept me tethered to reality. Their support has been integral to this effort. A complete list would be impossible, but a few people that went out of their way for me during this program are below.

Thank you to Evan, for your comradery throughout this program, from day

one. Dawn patrol, to coffee, to AdS/CFT study sessions were the highlights of my time at Santa Cruz.

Thank you to my colleagues Adrian and Renée. I didn't realize how much easier research could be with friends until you joined the group.

Thank you to Chanslor, Tyler, Nick, Bryce, and Rebecca for your consistent companionship through every stage of our lives.

Thank you to Brenda for always reminding me to “follow your dreams, sweetie!”

Thank you to Aris for keeping things in perspective.

Thank you to the many friends who supported me in Santa Barbara, especially my housemates Shane, Hanna, Christian, Percy, and sometimes Julia, and my creative partners Lindsey, Yen and Erin.

Thank you to Sebastian and Hanna for your friendship throughout school, and sharing your space during the writing process. I just realized while writing this that our time living together has book-ended my graduate career.

Finally, thank you to my first research advisor Keith Andrew for dedicating time to cultivating my interest in physics in high school. I see those early research experiences with you and Shane, Sam, and Akhil as the first step on the path that has led me here.

Chapter 1

Introduction

The microscopic nature of dark matter has been a major topic of research for many decades. One candidate theory is that it consists of black holes created in the early universe, known as primordial black holes (PBHs) [2]. The upcoming Rubin Observatory [3] and its ten-year Legacy Survey of Space and Time (LSST) will have a wide-ranging impact on astrophysics and cosmology, and it will have unique capabilities to characterize this dark matter. In this dissertation, I present work that predicts its ability to detect PBHs through gravitational microlensing events.

The first step is predicting the number of these events we can expect to see over the course of the ten-year survey. This is accomplished with analytical models of the dark matter density in our galaxy, and a simulated star catalog for Rubin. In

addition to the total number of events, I use Monte Carlo methods to generate catalogs of events that trace the parameter space predicted from the models. I then use simulated surveys to estimate the detection efficiency for these events.

The second aspect is a new detection algorithm to find these rare events in stellar lightcurves. It is particularly well-suited to make use of multi-color data, a unique aspect of the survey. By running this analysis pipeline on data from the NOIRLab Source Catalog (NSC) [4], I demonstrate strong background rejection for non-microlensing variations. I also show very high detection efficiency on real lightcurve data by superimposing the effects of the previously generated microlensing events.

This introduction provides a brief review of relevant background topics including the field of cosmology, the Vera C. Rubin Observatory, primordial black holes, and microlensing.

1.1 Cosmology

Cosmology is the study of nature at the largest distance scales. It asks questions about the universe as a whole. The field as we know it was arguably born with the publication of the general theory of relativity by Einstein in 1915 [5],

where he presented his field equations for gravity:

$$G_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu} \tag{1.1}$$

Einstein’s insights into the connection between space on the left side and matter on the right provided a framework to understand how the components of the universe affect its evolution through cosmic history. Because of this connection, the study of the largest scales is deeply intertwined with the physics at the smallest scales. A universe composed entirely of light would evolve differently under gravity from one composed entirely of ordinary matter, in a way that follows from matter-energy equivalence. The interplay between baryons, radiation, and dark matter left its mark in Large Scale Structure through Baryon Acoustic Oscillations [6–8], and it can be seen in the Small Scale Structure of galaxies [9].

The practice of cosmology also relies heavily on an understanding of microscopic physics. Doppler shifting of atomic emission spectra allows us to calculate the relative motion of distant light sources, and through Hubble’s Law [10], infer their distances. Nuclear physics gives insight into the consistent explosions of Type Ia supernova, and lets us use them as “standard candles” to calibrate distances. Laser interferometry allowed us to measure gravitational waves on Earth [11], giving us an alternate window into the cosmos from electromagnetic radiation, via gravity’s effect on electromagnetic radiation.

The core motivation for the work presented here is to test a model of one

of the components of the universe: dark matter. Although there’s much we do understand about dark matter on the cosmic scale, we know little to nothing about its microscopic properties despite decades of research. It could be ultra-light axions, and it could be black holes an order of magnitude heavier than the sun. Proponents of “modified gravity” believe it has no microscopic properties, and that the evidence displays a misunderstanding of gravity on certain scales [12]. However, this study on Primordial Black Holes (PBHs) requires no new particles, and relies on the same gravity we were presented with over a hundred years ago. I discuss PBHs as a candidate for dark matter in 1.3, but first, a bit of pedagogy.

1.1.1 Λ CDM

The Λ Cold Dark Matter model (Λ CDM) is known as the Standard Model of cosmology [13]. With just 6 parameters, it accurately reproduces many of the observed features of the cosmos, from present-day Large Scale Structure all the way back to the oldest observable light, the Cosmic Microwave Background. It is a particular realization of the Friedmann–Lemaître–Robertson–Walker (FLRW) solution to Einstein’s Field Equations [14, 15]:

$$ds^2 = dt^2 - a(t)d\Sigma^2 \tag{1.2}$$

Here, $a(t)$ is a scale factor, and contains all the time dependence of the metric. Σ is the spatial metric for the comoving geometry. Observations constrain this metric

to very close to flat on relevant scales, making $d\Sigma^2 \approx dx^2 + dy^2 + dz^2$. The FLRW metric is the most general metric which is both *homogeneous* and *isotropic*. This means every point in the comoving space is equivalent, and any time dependence is the same in all spatial directions.

The scale factor $a(t)$ is constrained by a choice of stress-energy tensor $T_{\mu\nu}$, and the EFE. In Λ CDM, we take the content of the universe to be a perfect fluid, with stress tensor

$$T_{\mu\nu} = (\rho + p)u_\mu u_\nu + pg_{\mu\nu} \tag{1.3}$$

with energy density ρ and pressure p , using units where $c = 1$. u_μ is the 4-velocity of the fluid. These are decomposed into contributions from radiation, matter, and dark energy as $\rho = \rho_r + \rho_m + \rho_\Lambda$ and $p = p_r + p_m + p_\Lambda$. Each constituent has a different equation of state, relating the energy to the pressure:

1. Radiation: $p_r = \frac{1}{3}\rho_r$
2. Matter: $p_m = 0$
3. Dark energy: $p_\Lambda = -\rho_\Lambda$.

The density and pressure, like the scale factor, are only functions of time. This follows from the homogeneity and isotropy assumption. By considering Equations 1.2 and 1.3 as an ansatz to the Einstein Field Equations, one can derive the

Friedmann equations for a universe with curvature parameter k [16]:

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{k}{a^2} \quad (1.4)$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3p) \quad (1.5)$$

The first is found by looking at the 00 component of the EFE, and the second is found by looking at its trace. The ratio \dot{a}/a is known as the Hubble expansion rate, denoted H , and H_0 is the present day value. The equations can be combined to give:

$$\dot{\rho} = -3H(\rho + p) \quad (1.6)$$

Considering a general perfect fluid with equation of state $p = w\rho$, we can solve the above ODE:

$$\dot{\rho} = -3(\dot{a}/a)\rho(1+w) \quad (1.7)$$

$$\frac{d\rho}{da} = -\frac{3(1+w)}{a}\rho \quad (1.8)$$

For $w \neq -1$, this has the solution $\rho = \rho_0 a^{-3(1+w)}$. For $w = 1$, we see $\dot{\rho} = 0$, and ρ must be a constant.

If we consider a universe containing different species of fluids which don't interact with each other, each of these densities will follow its own Equation 1.8, and the total density ρ will be a linear combination of these solutions:

$$\rho(a) = \rho_{0,r} a^{-4} + \rho_{0,m} a^{-3} + \rho_{0,\Lambda} \quad (1.9)$$

We can rewrite the first Friedmann equation in terms of this solution, introducing a few variables [17]:

$$\rho_c \equiv \frac{3H_0^2}{8\pi G} \quad (1.10)$$

$$\Omega_i \equiv \frac{\rho_i}{\rho_c} \quad (1.11)$$

$$\frac{H^2}{H_0^2} = \left(\frac{\dot{a}}{a}\right)^2 \frac{1}{H_0^2} = \Omega_{0,r}a^{-4} + \Omega_{0,m}a^{-3} + \Omega_{0,k}a^{-2} + \Omega_{0,\Lambda} \quad (1.12)$$

This ODE can be solved to find the scale factor a as a function of time, given the density parameters for these components of the universe. The Ω_i are not directly observable. Using astronomical data, cosmologists fit these parameters to the values that best match the measurements. In the late 1990s, two independent groups led by Riess and Perlmutter [18, 19] used Type Ia supernova measurements to demonstrate that not only was Ω_Λ nonzero, it was likely the dominant component of the universe at present day. This component was dubbed “Dark Energy” by Michael Turner in 1998 [20], and like Dark Matter, its microscopic properties remain unknown.

1.1.2 Dark Matter

Discussion about unseen mass contributions to the cosmos far predates Dark Energy, and even predates Relativity [21]. Lord Kelvin in 1904 was the first to be credited with finding kinematic evidence in the motions of stars, in that their velocity dispersions were inconsistent with those implied by their density

and mass [22]. In 1933, Fritz Zwicky's studies of the Coma Cluster revealed that the velocity dispersion of individual galaxies in the cluster was over an order of magnitude higher than would be predicted by the virial theorem [23].

The next genre of kinematic analysis would be that of rotation curves of galaxies, meaning the average velocity of stars as a function of their distance from the center of the galaxy. Although anomalies were reported in M31 as early as 1939 by Babcock in his Ph.D. dissertation [24], the idea that dark matter halos were responsible for the flattening of galactic rotation curves reached widespread popularity with Vera Rubin and Kent Ford's observations of M31 with a newly developed spectrograph [25].

As computers became more capable, it became feasible to simulate structure formation in the universe. The results of the simulations serve as a test of the cosmological models by comparing to observed distributions of galaxies on the sky, for example in [26, 27]. Simulations where dark energy and cold dark matter dominate are able to accurately reproduce the so-called Large Scale Structure seen in observations.

All the evidence listed so far is based on present-day measurements of the local universe. But the imprint of dark matter can be seen in the oldest electromagnetic signals available, the Cosmic Microwave Background (CMB). Anisotropies in the CMB were first revealed by COBE in 1992 [28, 29] and later characterized more

precisely by WMAP [30] and Planck [6]. The scale and amplitude of these temperature fluctuations reveal the interaction between dark matter, baryons, and radiation in the very early universe, which we call Baryon Acoustic Oscillations (BAO). These features set the density ratios of the three major components of the universe at the moment where the primordial plasma recombined into atoms. This is a critical measurement for distinguishing the dark matter content from baryonic matter on cosmic scales in the context of Λ CDM.

1.2 Vera C. Rubin Observatory

Astronomers have been anticipating the power of the Rubin Observatory for many years. It will produce a dataset which is the first of its kind, a movie in six colors across 10 years covering $\sim 20,000$ square degrees of the sky. It has required decades of research, design, construction, optimization, and verification. As a result of all this work, it is expected to start producing data in 2025.

In this section I will give a brief introduction to the motivation, design, and capabilities of the observatory. For thorough references, see [3,31].

1.2.1 Scientific goals

The list of science cases for such a dataset can fill a large book [31]. After 10 years it is expected to contain 20 billion galaxies, 17 billion resolved stars, and 6

million solar system objects. As impressive as those numbers are already, much of the power of the survey lies in the fact that we will not measure these objects once, but hundreds of times each. We will see the dynamic nature of the universe around us — a movie, not a photo. And while the most powerful results will come from the full survey, the discoveries in the final few years will be marginal compared to the wealth of previously inaccessible power gained in the first few years.

The four main pillars organizing the scientific goals of the survey are: Taking an Inventory of the Solar System, Mapping the Milky Way, Exploring the Transient Optical Sky, and Probing Dark Energy and Dark Matter. All these topics warrant a different observing strategy for optimizing the power of the survey. For a simple example, a solar system survey would want to spend a lot of time observing within the plane of the solar system. A Milky Way survey would spend a lot of time in a different plane — that of the galaxy. A transient survey might spend more time on the galactic center, where the density of stars is higher. And a dark energy/dark matter survey would want to find distant galaxies, and preferentially point away from all of these planes. But the Legacy Survey of Space and Time (LSST) plans to take data that's useful for all these topics! Designing a survey strategy - or *cadence* - that achieves this balance is a nontrivial task. The details are still under deliberation. In Section 3.3, I present a comparison of the capabilities of different

proposed survey strategies to detect Primordial Black Holes of different masses based on simulations of the full 10-year survey.

1.2.2 Camera design

At 3.2 gigapixels, the LSST camera is the largest digital camera ever built [32]. The pixels are $10\ \mu\text{m} \times 10\ \mu\text{m}$, which is around the scale needed to Nyquist sample the PSF on a night with ideal seeing, around 0.67 arcsec, given the plate scale of $50\ \mu\text{m}/\text{arcsec}$ [33]. The massive number of pixels is necessary to cover the large $9.6\ \text{deg}^2$ field of view, which enables rapid revisit times of a few days for objects across the sky. The photometry is extremely precise, with detector noise near the level of atmospheric noise on a clear night, $\lesssim 9\ e^-$.

The focal plane is composed of 189 $4\text{K} \times 4\text{K}$ CCD sensors¹, grouped into 3×3 *rafts* which share electronics including readout and power supplies. This parallel readout configuration allows a fast readout time of 2 seconds for the entire focal plane. High quantum efficiency from 320 to 1080 nm is achieved with $100\ \mu\text{m}$ thick deep depleted silicon. High voltage back-bias reduces diffusion between pixels to minimize the point-spread function (PSF) of the detector. Six bandpass filters are available with an automatic changer: u, g, r, i, z and y. Having data in several colors is important for characterizing transient phenomena. In particular, microlensing events are *achromatic* - the amplification of light is the same across

¹Described in more detail in Section 2.1

the spectrum. This is a key feature that lets us distinguish those events from other astrophysical variations.

1.2.3 Data management

The camera will produce $\sim 10 - 20$ TB² of raw image data every night for 10 years. That's challenging enough to move straight to archives around the world for scientists to access. Every image must undergo Instrument Signature Removal (ISR) to correct biases, dark currents, cosmic rays, and many other effects, in order to produce a Calibrated Exposure fit for science.

In addition, there will be real-time processing at the base facility to generate alert streams, enabling rapid followup with other facilities. This includes Difference Imaging Analysis (DIA), which compares the brightness of objects to baselines estimated from previous measurements. This technique gives us a shortcut to examine dynamic objects in the sky, including sources undergoing microlensing events. Reducing the search domain to only those showing sufficient changes will expedite searches for not only microlensing, but also variable stars, supernova, solar system objects, and more. These real-time alerts will include postage-stamp images of the transient objects, catalog information, and history of the objects.

Annual Data Releases will be produced containing all data from the beginning

²For years the plan was to take 2×15 second exposures for each visit, in part to mitigate cosmic ray events. It seems more likely now that the observatory will take 1×30 s exposures instead, halving the data requirements, and reducing read noise.

of the survey to the cutoff date of the release. They will contain data products including:

1. Calibrated exposures
2. Coadded images produced from several exposures
3. Object, Source, and FaintSource tables, constructed with a stack of exposures. Includes shapes, positions, fluxes, kinematics, and limited classification.
4. DIASources from the alert production
5. Orbits for Solar System Objects
6. An atmospheric model for every exposure

Since these releases are cumulative, containing all previous data every year, the processing can improve with every release, and remain self-consistent for every Data Product within the same release.

1.3 Primordial Black Holes as Dark Matter

Primordial Black Holes (PBHs) are, in a sense, the least speculative proposal for dark matter. They are unique in that they don't require undiscovered particles. They are simply black holes produced in the very early universe. For this reason,

and in part due to potential excesses in mass gap object detections with LIGO-Virgo [34], they have seen renewed interest in recent years.

Though we don't have evidence that the conditions were right for PBH production during inflation (or any indication about the details of inflation), we have plenty of evidence for black holes, both as stellar remnants and supermassive black holes [11, 35, 36].

1.3.1 Production methods

For a review of the various models capable of producing PBH's, see [2]. The general picture is that the extreme densities during inflation allows overdensities to overcome expansion pressure and gravitationally collapse. There are numerous studies proposing and exploring the PBH production of various inflation models. These models have a great deal of freedom to produce PBHs anywhere between 10^{-38} and $10^5 M_\odot$, though as described below, the ability to account for present day dark matter is much more limited.

1.3.2 Existing constraints

On the light mass end, we are constrained by Hawking radiation. Even in his first paper on the topic [37], Hawking notes that PBHs produced lighter than $\sim 10^{-18} M_\odot$ would have evaporated by now. Directly above this limit, light black

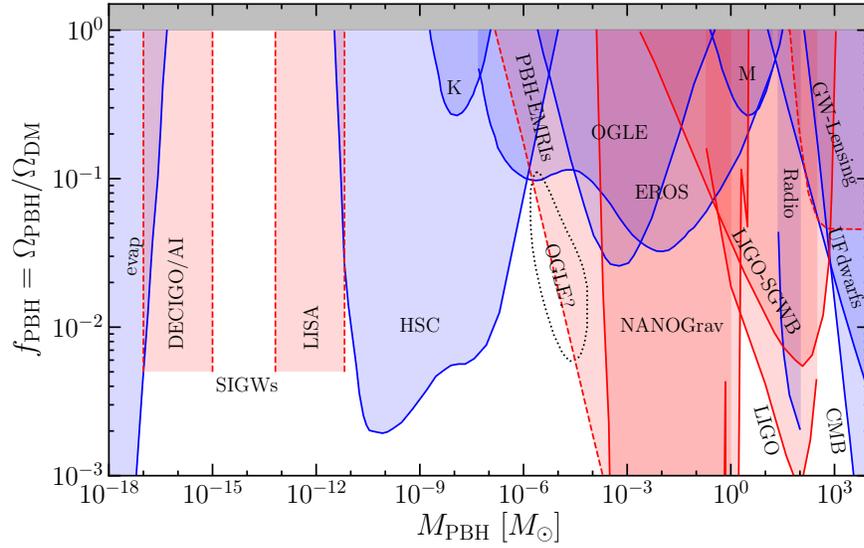


Figure 1.1: A composite of existing PBH constraints as a fraction of Dark Matter, taken from [1]. The red regions are from gravitational wave experiments, and the blue regions are from other measurements. Dashed lines show projected sensitivities from upcoming searches.

holes with $M \lesssim 10^{-17} M_\odot$ would produce an abundance of positrons [38]. Annihilation with nearby electrons would produce detectable γ -rays, particularly at the energy of the annihilated particles, $m_e = 511$ keV. Measurements of this emission line from INTEGRAL constrain the abundance of PBHs to $f_{DM} < 10^{-3}$ below $\sim 10^{-17} M_\odot$, and stronger for smaller masses [39].

On the high mass end, strong constraints are provided by anisotropies in the CMB, though these are highly dependent on modeling of the accretion history of the PBHs [40]. These limit the abundance around and above $10^3 M_\odot$.

In the vicinity of $10^{-11} - 10^{-6} M_\odot$, the Hyper-Supreme Cam/Subaru observations of the Andromeda Galaxy (M31) [41] place strong limits via microlensing, the subject of this dissertation. This mass sensitivity window is set by models of the M31 halo, velocity dispersions of the stars and PBHs, the distance to M31, the 2-minute sampling cadence of the observations, and the 7-hour duration of the program. As we will see in Chapter 3, for a fixed halo mass distribution and velocity profile, microlensing events have timescales t_E proportional to the square root of the individual PBH mass. This means faster observations are more sensitive to lighter lenses.

The slower cadence, longer duration surveys OGLE [42], EROS [43] and MACHO [44] constrain the higher masses $10^{-7} - 10^0 M_\odot$ by measuring the Magellanic clouds and the Milky Way Galactic Bulge.

1.3.3 Rubin’s potential contributions

Unlike MACHO, EROS, and OGLE, Rubin’s Wide-Fast-Deep survey isn’t tailored to microlensing event discovery. The survey strategy balances a great deal of scientific goals. Nonetheless, the consistent revisiting of many sources over such a long period with strong photometric repeatability is the ideal condition for discovering long duration events.

One distinct and powerful advantage for Rubin in this domain is the broad quantum efficiency of the camera from IR to UV wavelengths, and its 6 available filters. Previous major microlensing surveys have only used one or two filters. As previously mentioned, the achromaticity of microlensing events is an important feature which can be leveraged to strongly reject background astrophysical variations. It is central to the approach described in Chapter 4. With only one filter, one learns nothing about the color variations in a transient event, and must rely on the quality of fit to an ideal microlensing curve, and contextual information like whether the source has displayed variations in the past. In contrast, with multiple filters, certain variable stars can be kept as potential lensing sources, since intrinsic variation is typically chromatic to some degree, with each band changing a different amount. If one were to be lensed during a period of stability, the event would be distinguished by its profile in color space. With such rare events, background rejection becomes indispensable.

1.4 Microlensing

1.4.1 Derivation of microlensing curve from General Relativity

1.4.1.1 Null geodesics of the Schwarzschild metric

The most general static spherically-symmetric solution to the Einstein Field Equations in (1,3) dimensions is the Schwarzschild metric, first presented in [45]:

$$ds^2 = \left(1 - \frac{r_s}{r}\right) dt^2 - \left(1 - \frac{r_s}{r}\right)^{-1} dr^2 - r^2 d\theta^2 - r^2 \sin^2 \theta d\phi^2 \quad (1.13)$$

with $r_s \equiv 2GM$ and units where $c = 1$.

We want to understand what happens to light rays traveling through this metric. Light follows null geodesics, paths with $ds^2 = 0$. Without loss of generality, let's just consider paths along the equatorial plane, with $\theta = \pi/2$. Then we're left with:

$$0 = \left(1 - \frac{r_s}{r}\right) dt^2 - \left(1 - \frac{r_s}{r}\right)^{-1} dr^2 - r^2 d\phi^2 \quad (1.14)$$

We can use the geodesic equation to find some conserved quantities:

$$\frac{d^2 x^\mu}{d\lambda^2} + \Gamma_{\nu\rho}^\mu \frac{dx^\nu}{d\lambda} \frac{dx^\rho}{d\lambda} = 0 \quad (1.15)$$

with the Christoffel symbols $\Gamma_{\nu\rho}^\mu$ and an affine parameter λ .

From [46], the definition of the Christoffel symbols is:

$$\Gamma_{\mu\nu}^{\lambda} = \frac{1}{2}g^{\lambda\sigma} \left(\frac{\partial g_{\sigma\nu}}{\partial x^{\mu}} + \frac{\partial g_{\sigma\mu}}{\partial x^{\nu}} - \frac{\partial g_{\mu\nu}}{\partial x^{\sigma}} \right) \quad (1.16)$$

We will compute the nonzero Christoffel symbols below. Since the Schwarzschild metric is expressed diagonally above, the symbols simplify first because the inverse metric is simply given by the inverse of the diagonal elements, and second because the sum over σ reduces to just the $\sigma = \lambda$ terms.

$$\Gamma_{tr}^t = \Gamma_{rt}^t = \frac{1}{2} \left(1 - \frac{r_s}{r} \right)^{-1} \left(\frac{r_s}{r} \right) \quad (1.17)$$

$$\Gamma_{rr}^r = -\frac{1}{2} \left(1 - \frac{r_s}{r} \right) \left(\left(1 - \frac{r_s}{r} \right)^{-2} \left(\frac{r_s}{r^2} \right) \right) \quad (1.18)$$

$$= -\frac{1}{2} \frac{r_s}{r^2} \left(1 - \frac{r_s}{r} \right)^{-1} \quad (1.19)$$

$$\Gamma_{tt}^r = \frac{1}{2} \frac{r_s}{r^2} \left(1 - \frac{r_s}{r} \right) \quad (1.20)$$

$$\Gamma_{\phi\phi}^r = -\frac{1}{2} \left(1 - \frac{r_s}{r} \right) (2r) = -r \left(1 - \frac{r_s}{r} \right) \quad (1.21)$$

$$\Gamma_{\phi r}^{\phi} = \Gamma_{r\phi}^{\phi} = -\frac{1}{2r^2} (-2r) = \frac{1}{r} \quad (1.22)$$

We can plug these in to get constraints on the geodesics in each coordinate. From now on we will repurpose the coordinate symbols to refer to functions of an affine parameter λ describing the path taken by a light ray: $t = t(\lambda), r = r(\lambda), \phi = \phi(\lambda)$.

The geodesic equation for ϕ is:

$$\frac{d^2\phi}{d\lambda^2} + \frac{2}{r} \frac{d\phi}{d\lambda} \frac{dr}{d\lambda} = 0 \quad (1.23)$$

We can multiply both sides by r^2 to see this is equivalent to a total derivative equal to zero:

$$\frac{d^2\phi}{d\lambda^2}r^2 + \frac{d\phi}{d\lambda} \left(2r \frac{dr}{d\lambda} \right) = 0 \quad (1.24)$$

$$\frac{d}{d\lambda} \left(\frac{d\phi}{d\lambda} r^2 \right) = 0 \quad (1.25)$$

This implies that $\boxed{\frac{d\phi}{d\lambda} r^2 \equiv L}$ is a constant along the geodesic.

Similarly, for the t equation:

$$\frac{d^2t}{d\lambda^2} + \frac{r_s}{r^2 \left(1 - \frac{r_s}{r} \right)} \frac{dt}{d\lambda} \frac{dr}{d\lambda} = 0 \quad (1.26)$$

$$\frac{d^2t}{d\lambda^2} \left(1 - \frac{r_s}{r} \right) + \frac{r_s}{r^2} \frac{dt}{d\lambda} \frac{dr}{d\lambda} = 0 \quad (1.27)$$

$$\frac{d}{d\lambda} \left(\frac{dt}{d\lambda} \left(1 - \frac{r_s}{r} \right) \right) = 0 \quad (1.28)$$

Implying another constant $\boxed{\frac{dt}{d\lambda} \left(1 - \frac{r_s}{r} \right) \equiv E}$. Now we can refactor these into the metric for the light path, dividing by $d\lambda^2$, and using $\frac{dr}{d\lambda} = \frac{dr}{d\phi} \frac{d\phi}{d\lambda} = \frac{dr}{d\phi} \frac{L}{r^2}$.

$$0 = \left(1 - \frac{r_s}{r} \right) \left(\frac{dt}{d\lambda} \right)^2 - \left(1 - \frac{r_s}{r} \right)^{-1} \left(\frac{dr}{d\lambda} \right)^2 - r^2 \left(\frac{d\phi}{d\lambda} \right)^2 \quad (1.29)$$

$$0 = \left(1 - \frac{r_s}{r} \right)^{-1} E^2 - \left(1 - \frac{r_s}{r} \right)^{-1} \frac{L^2}{r^4} \left(\frac{dr}{d\phi} \right)^2 - \frac{L^2}{r^2} \quad (1.30)$$

$$\frac{L^2}{r^4} \left(\frac{dr}{d\phi} \right)^2 = E^2 - \frac{L^2}{r^2} \left(1 - \frac{r_s}{r} \right) \quad (1.31)$$

$$\left(\frac{dr}{d\phi} \right)^2 = r^4 \frac{E^2}{L^2} - r^2 \left(1 - \frac{r_s}{r} \right) \quad (1.32)$$

E and L being dependent on an arbitrary parametrization of the coordinates via the λ derivative makes it unclear how to physically interpret them. But since

just their ratio appears in the above differential equation, we can use the chain rule to reduce these to something independent of λ :

$$\frac{L}{E} \equiv b = \frac{d\phi/d\lambda}{dt/d\lambda} \frac{r^2}{1 - r_s/r} \quad (1.33)$$

$$= \frac{d\phi}{dt} \frac{r^2}{1 - r_s/r} \quad (1.34)$$

We can understand the *impact parameter* b by going back to Equation 1.32, and considering the moment of closest approach, when $r = r_{\min}$. At this point, since r is at a minimum along its trajectory, its derivative vanishes along λ . There must still be motion along ϕ , so $dr/d\phi = 0$. Then we have:

$$\frac{1}{b^2} = \frac{1}{r_{\min}^2} - \frac{r_s}{r_{\min}^3} \quad (1.35)$$

$$b^2 = r_{\min}^2 (1 - r_s/r_{\min})^{-1} \quad (1.36)$$

So assuming r_{\min} is meaningfully larger than r_s , b is on the same order as r_{\min} .

We can see that for $r_s = 0$, $b = r_{\min}$.

1.4.1.2 Deflection angle

Now, to find the angle at which light is deflected by the black hole, we must evaluate $\int_{\infty}^{r_{\min}} dr [d\phi/dr]$. First rearrange Equation 1.32:

$$\frac{d\phi}{dr} = \left(r^4/b^2 - r^2 + rr_s \right)^{-1/2} \quad (1.37)$$

$$\Delta\phi = 2 \int_{\infty}^{r_{\min}} dr \left(r^4/b^2 - r^2 + rr_s \right)^{-1/2} \quad (1.38)$$

The factor of 2 comes from integrating the function along a curve from $r = \infty$ back to $r = \infty$ - we must integrate the path coming toward the minimum approach distance, then back out. Since it's symmetric, we just multiply 2.

It turns out this is an intricate integral, resulting in an analytic solution involving the Weierstrass elliptic function, which Mathematica will produce after a long evaluation time. And it makes sense that this is something complicated. After all, this is an exact solution so far. It has to include all the chaotic funny business one could imagine when a light ray gets close to the photon sphere. Fortunately, in the case of microlensing, we have a small parameter: $r_s/r_{\min} \ll 1$. To start, there is a hard constraint from the black hole shadow limiting light rays to $r_{\min} > \sqrt{27}$ [47]. Numerical evaluation shows the approximation below is off by $\sim 20\%$ at that limit. But from the geometric arguments in the next section, for microlensing, this small parameter is more like 10^{-6} in the least generous cases, and generally much smaller³.

³This is for a $100M_{\odot}$ lens 10 pc away from a bulge source.

We can organize terms in powers of r_s :

$$\Delta\phi = 2 \int_{\infty}^{r_{\min}} dr \left(r^4/r_{\min}^2 (1 - r_s/r_{\min}) - r^2 + rr_s \right)^{-1/2} \quad (1.39)$$

$$= 2 \int_{\infty}^{r_{\min}} dr \left(r^4/r_{\min}^2 - r^4 r_s/r_{\min}^3 - r^2 + rr_s \right)^{-1/2} \quad (1.40)$$

$$= 2 \int_{\infty}^{r_{\min}} dr \left(r^4/r_{\min}^2 - r^2 - r_s \left(r^4/r_{\min}^3 - r \right) \right)^{-1/2} \quad (1.41)$$

$$= 2 \int_{\infty}^{r_{\min}} dr \frac{1}{r} \left(r^2/r_{\min}^2 - 1 - \frac{r_s}{r} \left(r^3/r_{\min}^3 - 1 \right) \right)^{-1/2} \quad (1.42)$$

$$= 2 \int_{\infty}^{r_{\min}} dr \frac{1}{r} \left(r^2/r_{\min}^2 - 1 \right)^{-1/2} \left(1 - \frac{r_s}{r} \left(\frac{r^3/r_{\min}^3 - 1}{r^2/r_{\min}^2 - 1} \right) \right)^{-1/2} \quad (1.43)$$

Near $r \rightarrow \infty$, the outer $1/r$ term pushes the integrand to 0. Near $r \rightarrow r_{\min}$, the second term is well approximated by a series expansion of $(1-x)^{-1/2} \approx 1 + \frac{1}{2}x$.

$$\Delta\phi \approx 2 \int_{\infty}^{r_{\min}} dr \frac{1}{r} \left(r^2/r_{\min}^2 - 1 \right)^{-1/2} \left(1 + \frac{r_s}{2r} \left(\frac{r^3/r_{\min}^3 - 1}{r^2/r_{\min}^2 - 1} \right) \right) \quad (1.44)$$

The first term in this integral evaluates to $2 \arctan(\sqrt{r^2/r_{\min}^2 - 1})_{\infty}^{r_{\min}} = -\pi$.

This corresponds to no change in direction with no black hole, $r_s = 0$.

The first order term in r_s is:

$$\alpha \equiv \pi - \Delta\phi \approx -r_s \int_{\infty}^{r_{\min}} dr \frac{1}{r^2} \left(\frac{r^3/r_{\min}^3 - 1}{(r^2/r_{\min}^2 - 1)^{3/2}} \right) \quad (1.45)$$

$$(1.46)$$

Substitute $x = r/r_{\min}$, then evaluate the integral over x by substituting $u = \sec(x)$:

$$\alpha \approx -\frac{r_s}{r_{\min}} \int_{\infty}^1 \frac{dx}{x^2} \left(\frac{x^3 - 1}{(x^2 - 1)^{3/2}} \right) \quad (1.47)$$

$$\approx -\frac{r_s}{r_{\min}} \left(\frac{2x^2 - x - 1}{x\sqrt{x^2 - 1}} \right)_{\infty}^1 \quad (1.48)$$

$$\alpha \approx \boxed{\frac{2r_s}{r_{\min}}} \quad (1.49)$$

This matches the result of Einstein first presented much more concisely in [5].

1.4.1.3 The lens equation

The deflection angle is all the gravitational input we need. The rest is geometry and optics. We will proceed in the Point Source-Point Lens (PSPL) approximation⁴.

Call the true angular distance between the source and the lens β , and the position of the deflected image from the lens θ . Of course, these are really solid angles $\vec{\beta}$ and $\vec{\theta}$, but for simplicity and without loss of generality, consider the angle restricted to the plane made by the observer, source, and lens, since they are not perfectly collinear.

We can make a few observations given the geometry of Figure 1.2. Since the deflection angle is small, $r_{\min} \approx D_L \theta$. We also see $(\theta - \beta)D_S \approx \alpha D_{LS}$. We can

⁴A point source refers to anything unresolved by a given optical system. All stars (except the Sun) are point sources for ground-based telescopes like Rubin, but if a galaxy is lensed, its shape comes into play.

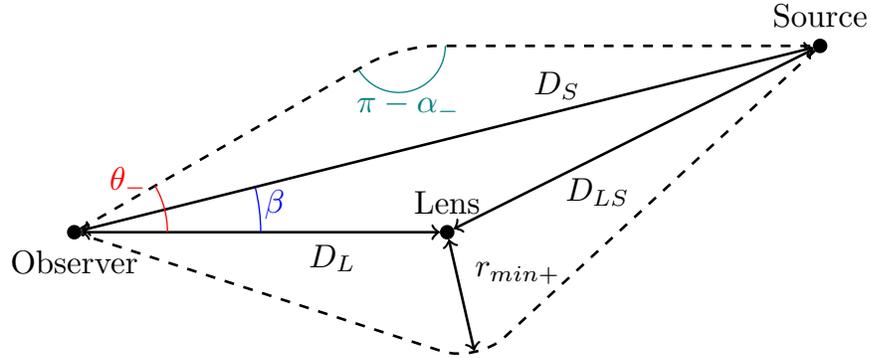


Figure 1.2: A schematic diagram of the lensing geometry. The upper path has deflection angle α_- , image location θ_- , and closest approach radius $r_{\min-}$, and the lower path has α_+ , θ_+ , and $r_{\min+}$. Only one parameter is shown per pair to maintain visibility. In reality, α_{\pm} is an extremely small angle, and the objects are very nearly aligned. The dashed lines are the null geodesics between the observer and source. The images will appear at the angle θ_{\pm} away from the lens, and the true position of the source is β away from the lens.

solve for β to get:

$$\beta = \theta - \alpha \frac{D_{LS}}{D_S} \quad (1.50)$$

$$= \theta - \frac{2r_s D_{LS}}{D_L D_S \theta} \quad (1.51)$$

$$= \theta - \frac{\theta_E^2}{\theta} \quad (1.52)$$

where we have defined the *Einstein angle* $\theta_E \equiv \sqrt{2r_s \frac{D_{LS}}{D_S D_L}}$. This parameter sets the angular scale of the lensing event. Now we have a quadratic equation for θ given β , and it has two solutions:

$$\theta_{\pm} = \frac{1}{2} \left(\beta \pm \sqrt{\beta^2 + 4\theta_E^2} \right) \quad (1.53)$$

1.4.1.4 Magnification

The deflection of light distorts areas when mapping from the source plane to the image plane. As a result, although the light intensities and frequencies stay the same, the fluxes reaching the observer do not. They vary proportionally to the area of their images, as described in depth in Section 5.2 of [48].

Because of the circular symmetry of Equation 1.53, we can think of β and θ as radial coordinates from the lens, and a corresponding angle ϕ about the lens, which is unaffected by the transformation. An infinitesimal solid angle in the source plane is $\beta d\phi d\beta$, and in the image plane $\theta d\phi d\theta$. Thus, after cancelling the

$d\phi$, the ratio of the areas is

$$\mu = \left| \frac{d\theta}{d\beta} \frac{\theta}{\beta} \right| \quad (1.54)$$

$$\mu_{\pm} = \left| \frac{1}{2} \left(1 \pm \beta (\beta^2 + 4\theta_E^2)^{-1/2} \right) \frac{\theta_{\pm}}{\beta} \right| \quad (1.55)$$

$$\mu_{\pm} = \frac{1}{4} \left(1 \pm \beta (\beta^2 + 4\theta_E^2)^{-1/2} \right) \left(\frac{1}{\beta} \sqrt{\beta^2 + 4\theta_E^2} \pm 1 \right) \quad (1.56)$$

$$\mu_{\pm} = \frac{1}{4} \left(\beta (\beta^2 + 4\theta_E^2)^{-1/2} + \frac{1}{\beta} \sqrt{\beta^2 + 4\theta_E^2} \pm 1 \pm 1 \right) \quad (1.57)$$

$$\mu_{\pm} = \frac{1}{4} \left(u (u^2 + 4)^{-1/2} + \frac{1}{u} \sqrt{u^2 + 4} \pm 2 \right) \quad (1.58)$$

$$(1.59)$$

where we substitute the dimensionless impact parameter $u = \beta/\theta_E$.

Since there are two images, to get the total flux from the source, we simply add the magnifications.

$$\mu = \frac{1}{2} \left(u (u^2 + 4)^{-1/2} + \frac{1}{u} \sqrt{u^2 + 4} \right) \quad (1.60)$$

$$\mu = \frac{1}{2} \left(\frac{u^2 + (u^2 + 4)}{u \sqrt{u^2 + 4}} \right) \quad (1.61)$$

$$\boxed{\mu = \frac{u^2 + 2}{u \sqrt{u^2 + 4}}} \quad (1.62)$$

1.4.1.5 Lightcurves

If we assume the lens and source have linear trajectories on the sky, the impact parameter can be parametrized as:

$$u(t) = \sqrt{u_{\min}^2 + \left(\frac{t - t_0}{t_E} \right)^2} \quad (1.63)$$

where u_{\min} is the smallest value of u along the trajectory, which is reached at the peak time t_0 , and the Einstein crossing time t_E is the time it takes for the lens to traverse one Einstein angle θ_E in the frame of the source. Computing the magnification as a function of time using this impact parameter gives the blue curves in Figure 1.4.

1.4.1.6 Astrometric shifts

Even though the two images from a single lens are rarely resolved, the shift from the source position can be described in terms of the centroid of the two deflected images, given in [35] as:

$$\delta(\mathbf{u}) = \frac{\mathbf{u}}{u^2 + 2} \theta_E \quad (1.64)$$

For a bulge source lensed by a large $100M_{\odot}$ PBH halfway to the bulge with $u_{\min} = 1$, this corresponds to a shift of ~ 5 mas. Rubin’s astrometric accuracy spec of 50 mas [33] indicates this will not be a relevant effect for most events, though if it is pronounced, it can be combined with inference on the lightcurve to better constrain the Einstein angle θ_E .

1.4.2 Parallax effects

Assuming the motions of sources and lenses are linear trajectories on the sky on these time and distance scales is reasonable considering the rotation curves of

the Milky Way and the distance to the bulge and Magellanic Clouds. However, the motion of Earth around the sun is a relevant effect to microlensing in certain regimes. Ref [35] contains a thorough and generic presentation of the relevant kinematic equations.

Copied from the above reference, the concise form reads:

$$\mathbf{u}(t) = \mathbf{u}_0 + (t - t_0)\dot{\mathbf{u}}_0 + \pi_E \delta\boldsymbol{\gamma}(t) \tag{1.65}$$

where \mathbf{u} is the 2D impact parameter on the sky, $\dot{\mathbf{u}}_0$ is the linear trajectory assuming a stationary observer, for example centered at the sun, and $\delta\boldsymbol{\gamma}(t)$ is the projection of Earth’s orbit onto a plane perpendicular to the vector toward the source star.

Figure 1.3 shows an extreme example of the effect of parallax on a microlensing curve, with an especially long duration event for the size of the lens, whereas Figure 1.4 shows a typical case, where there is little effect.

The algorithm described in Section 4.4 does assume a PSPL curve with no parallax for the fitting in Round 2. But Round 1 is model-independent, and only looks for achromatic brightening, which is the case with or without parallax. The information gained in pronounced parallax effects is a good thing for parameter inference. Without it, we only infer one degree of freedom from the curve — the crossing time ⁵. With an additional feature to fit in a model with parallax,

⁵The distribution of the minimum impact parameter u_{\min} is uniform across all relevant PBH masses, and doesn’t help infer lens masses.

combined with a known parallax from a source star, we can start to break the degeneracy between distance and mass baked into the formula for the Einstein angle. This characterization was used to discover an isolated stellar mass black hole in [35].

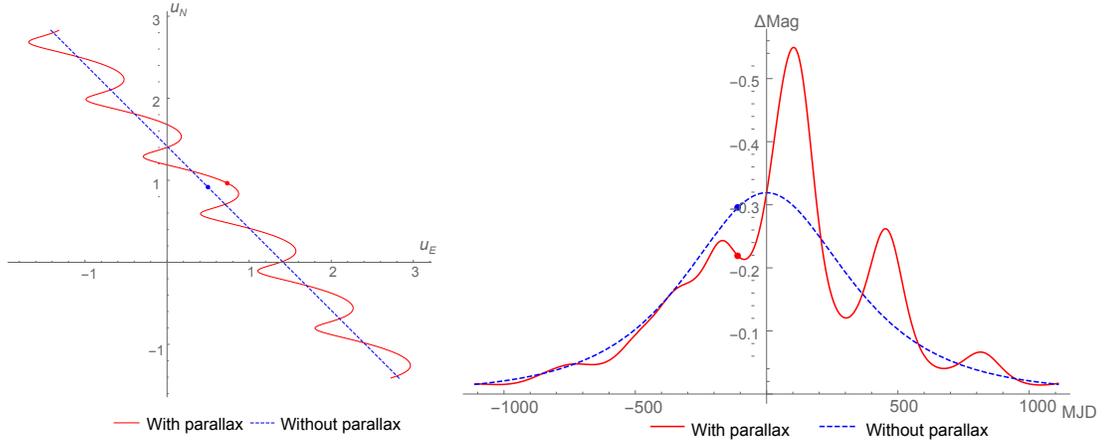


Figure 1.3: A point source point lens microlensing curve with parallax (red) and without parallax (blue), from a bulge source and a $0.1M_{\odot}$ lens halfway to the bulge from Earth, with an atypically ($\sim 4\sigma$) long crossing time of 370 days. The left shows the trajectory in impact parameter \mathbf{u} , and the right shows the corresponding change in magnitude over time, given in Modified Julian Days (MJD).

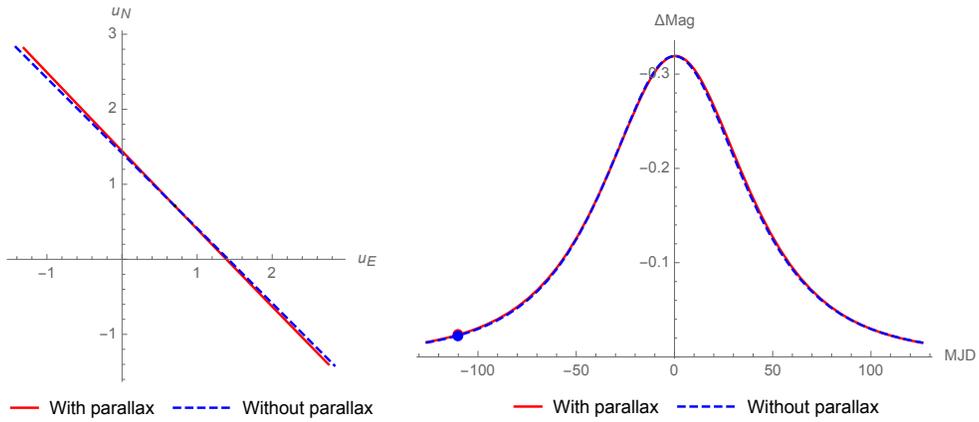


Figure 1.4: A median timescale event of ~ 6 days from a $0.1M_{\odot}$ PBH, showing the marginal effect of parallax on the typical event.

Chapter 2

Camera work

In this Chapter, I describe some of the key contributions I made to the science verification and testing of the LSST Camera. First I will provide some background information on the sensors which form the basis for its operation.

2.1 Charge Coupled Devices (CCDs)

Humans have been gathering data from the night sky for millennia, in one way or another. On that timescale, digital technologies and methods are still in their infancy. Yet, it would be hard to overstate the revolutionary impacts the field has seen as a result of their adoption in astronomy.

Astronomers' interest in the use of digital computers grew significantly during the 1960s, but even through the next decade, the principal methods for recording

images from telescopes relied on photographic emulsions [49]. Early adopters of digital methods measured the brightness of objects with machines that scanned the photographs and converted them to digital records. This interface between analog and digital methods was a major point of data friction in this transition period. The process of scanning and storing photographs required significant human intervention and storage of physical media. The logical progression seemed to require a natively electronic method of photon detection.

One of the first was the Charge Coupled Device (CCD), developed by Willard Boyle and George E. Smith at Bell Labs in 1969 as a data storage device [50]. J. Anthony Tyson, now Chief Scientist for the Rubin Observatory, pioneered its application to astronomy in the late 1970s by searching for faint galaxies [51]. Though there were several options in the context of rapid hardware and materials innovation of the era, by 1977 astronomers predicted that CCDs would eventually become the standard tools of the trade [52]. In the decades since then, the linearity, low noise, and high quantum efficiency of the technology would enable a new era of precision astronomy. The invention of the CCD would receive half the Nobel Prize in Physics in 2009 [53] for its impact on this field and many others.

2.1.1 Detection and integration of photons

CCDs are fundamentally composed of a semiconducting bulk, usually silicon based, and a set of control gates for each pixel, galvanically isolated from the bulk with a layer of insulation. The CCD's operation as a detector relies on the photoelectric effect [54] to create electron-hole pairs within a depletion region when an incident photon has enough energy to excite an electron across the gap between the valence band and the conduction band. The free electrons follow the electric field induced by the gate voltages, and collect in a potential well. The LSST Camera sensors use high-resistivity crystalline silicon, with a layer of n-doped silicon closer to the gates, and p-doped silicon on the back. This forms the depletion region, where electrons from the n-type region diffuse into the p-type region to recombine with holes, eliminating the available charge carriers. The depletion region is extended throughout the bulk by applying a high voltage substrate bias, nominally -70 V.

The redistribution of charges in the depletion region creates an electric field, which separates the conduction band electrons and holes quickly once they are produced by a photon, bringing the electrons toward the gates and the holes toward the back. Without this field, opposite charge carriers recombine before they could be collected and measured.

The depth of the depletion region is important for the absorption of long-

wavelength infrared photons. The key principle for detection is for a photon to supply an electron with enough energy to exceed the band gap of ~ 1.14 eV¹ and become conducting. However, for a photon to be absorbed with this minimum energy requires an interaction with an electron in the highest-energy valence band state. The density of states gets sparser near the band gap, as can be seen in Figure 1 of [55]. Since higher energy photons can excite electrons in lower energy valence states across the band gap, they have more opportunities to be absorbed, and therefore higher quantum efficiency. Additionally, interactions with phonons introduce a temperature dependence, which further reduces the absorption paths available to low energy photons compared to those with higher energies when operating at low temperatures [56]. Making the depletion region deeper gives the low-energy photons more opportunities to be absorbed and detected.

2.1.2 Readout operation

Once the photoelectrons have been integrated within the substrate and collected by the central gate of each pixel, they are read out in a process dubbed the “bucket brigade.” By controlling the voltage levels of the gates with a collection of clock signals, the charges from each pixel can be coherently shifted to adjacent pixels, and sequentially fed into the output electronics to be effectively counted and digitized. Figure 2.1 displays both parallel and serial shifts. The parallel

¹A photon with wavelength ~ 1090 nm

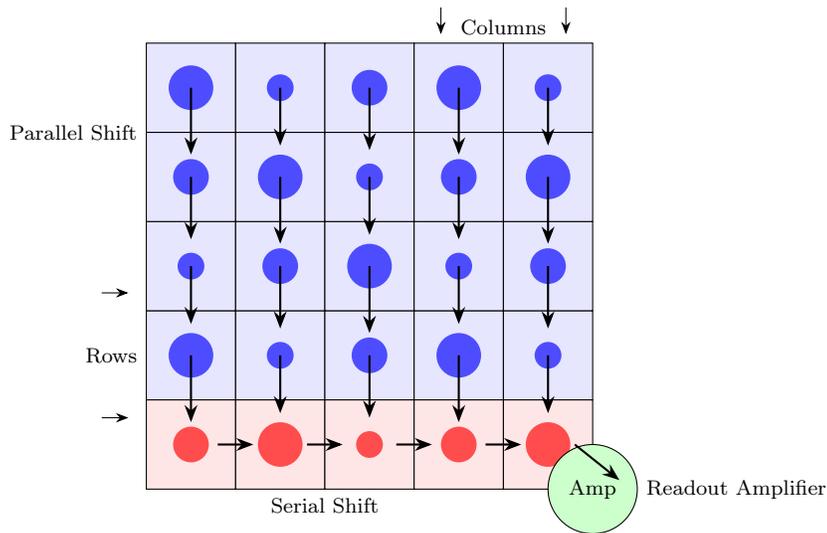


Figure 2.1: The bucket brigade operation of charge shifting in a CCD.

shifts move charges between rows, with charges from every column moving in sync. When a row of charges reach the final row, called the *serial register*, they are shifted in the orthogonal direction, as each charge packet reaches the readout chain at the end. Here, the *sense node* converts the charge packet via a capacitor to a voltage, which is then amplified by a MOSFET to the CCD video output channel.

The readout electronics make use of *correlated double sampling* to reduce the effects of variation of the reset level in the video output. After each readout, the sense node clears charges via a gate connecting it to the reset drain. The output electronics use a Sample-and-Hold Amplifier (SHA) to integrate, then hold this voltage at the inverting input of a differential amplifier, before the serial clocks

bring the next charge packet to the sense node. With the charges in place, the video level is integrated and held on the noninverting input. Now the differential amp’s output is proportional to the difference between the reset level and the output level measured with the sense node [57]. This voltage is then sent to an ADC for digitization. The measurement is encoded in Analog-Digital Units (ADU), and in the LSST Camera, our gain is $O(1)$ ADU/photon.

2.1.3 Bias correction

The bias level of a CCD readout chain is an offset level that persists through the ADC stage, and needs to be subtracted during Instrument Signature Removal (ISR). It would be nice if we could say “*constant* offset level,” but this is not the case, despite the readout electronics being designed with this in mind. The reality is that the bias level is a dynamic feature of the measurement, and requires careful consideration to avoid introducing systematic effects.

There are a few levels at which we can think about the variable nature of the bias:

1. Stochastics from Johnson noise in the circuits (each readout)
2. Profiles along the parallel shift direction (parallel bias profile)
3. Profiles along the serial shift direction (serial bias profile)

4. Non-factorizing profiles along the image readout (2D bias profile)
5. Variations across images (bias stability)

The per-pixel noise we can't do much about, besides stacking multiple images.

One standard tactic is to take *bias frames*, which are CCD readouts performed with no integration time on the sensor. These are meant to assess the response of the system when there are zero photon detections in a measurement. To correct the bias in an image A_{ij} , one can first take a bias frame B_{ij} , then simply subtract it to get a corrected image $C_{ij} = A_{ij} - B_{ij}$. This can be improved by considering the inherent read noise. If both A and B contain Gaussian read noise with width σ , their difference will contain noise with width $\sigma\sqrt{2}$. Averaging N bias frames together reduces the standard deviation by $1/\sqrt{N}$ and produces a *superbias* with arbitrarily small variance, leaving the corrected image with just the read noise of the original measurement, theoretically. For an extremely stable camera, this is a good strategy. One can take an arbitrary number of bias frames, average them together, and subtract the same superbias for every readout indefinitely.

But real cameras are dynamic. They respond to their environment. Changes in temperature, readout history, and power supply, to name a few, can have difficult-to-predict effects on the response of the system. In the LSST Camera, at least, relying on superbias alone is too optimistic — no single bias frame will be well-approximated by a superbias. To do the best ISR job possible, we need some

more data.

As far as bias correction, another powerful tool is *overscans*. To read out the charges integrated on a CCD, one shifts according to the number of rows and columns of pixels on the sensor (segment). After shifting all the columns to the readout chain, any further readouts will not correspond to pixels. Instead, they will read the level corresponding to no charge — the bias ². But now, instead of extrapolating the bias from a separate image, we have an almost instantaneous indication of the bias level, taken at the end of the readout of every row. We can use this data as a proxy for the serial bias profile.

Parallel overscans are measured by continuing to read *rows* after all the pixels have been read. Reading entire rows worth of pixels gives insight into how the bias changes within each row. There can be substantial left-to-right structure within each row, and subtracting an average parallel overscan profile can be effective for removing it.

If the bias in an image could be factorized as serial and parallel overscans, i.e. $B_{ij} = B_i^{\text{serial}} + B_j^{\text{parallel}}$, we would have a good solution here. And in fact, for many sensors this is perfectly suitable for correction - many sensors don't even need the parallel correction. But there are a handful of sensors for which the bias

²There is a caveat here: due to Charge Transfer Inefficiency (CTI), something like $1/10^5$ electrons are leftover in the previous pixel during each shift. This smearing of charge leaves a remnant of the last few pixels in the overscans, and by using hi-flux flat test images, we can leverage this fact to characterize the CTI.

does not factorize in this way. For these sensors we have to rely on superbiases, and be careful to assess their stability.

2.2 Multi-modal behavior of LSST Camera amplifiers

Through the course of Electro-Optical (EO) testing for the LSST Camera, it became clear that there was some state-dependent behavior in the readout electronics. This is exemplified by looking at the bias levels, which appear quasi-stable at the row-level (after removing parallel structure), until a transition to another internal state causes them to jump to a different baseline. Figure 2.2 shows the clustering of overall bias levels per image for the same CCD segment, compared across 16 different sequencer settings. The mean bias levels of the upper and lower segments, associated with one ASPIC³ each, are correlated depending on their state, making this clustering behavior around different baselines.

This statefulness can be seen not only at the per-image level, but also within images. This is known as the “bias-shift” issue described in detail in Section 2.2.1 below.

³Analog Signal Processing Integrated Circuit, a custom designed circuit with output electronics for the LSSTCam [58]

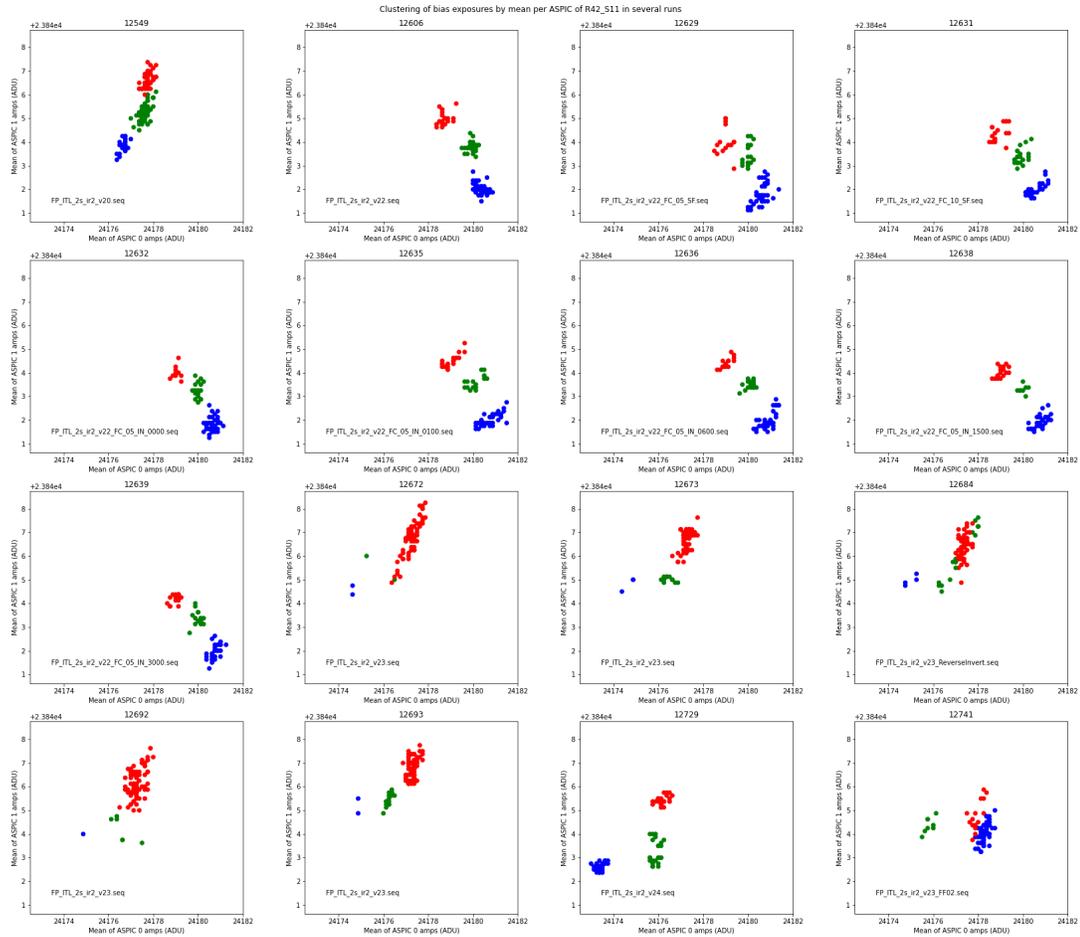


Figure 2.2: Visualizing the changes in bias level in response to different sequencer settings. Each datapoint is derived from one image, where the x and y positions are the mean of the bias level over all segments associated with one of the two ASPIC chips. Multimodal structure persists in different ways through all tested settings.

2.2.1 Bias shifts

One strong example of the serial overscans' utility is in the *bias shift* issue. For certain readout configurations of the camera, we observe a rapid change in the average bias level from row to row (e.g. Figure 2.3). The precise location of the shift is not consistent among different images for the same sensor, which makes this impossible to correct with the superbias method. This location variation can be seen at a significant level in Figure 2.9, but even at the smaller level seen in Figure 2.10, inconsistencies in the location of these shifts break all correction strategies except those that make direct use of the serial overscans. The serial overscans give us insight into the timing of the shift, though we don't know where it happens in the image between overscans. We can assess the shift in more detail by looking at a bias frame, as in Figure 2.4. We can see a trend climbing up from the middle of row 66 to the end of row 67, suggesting a time constant much slower than the pixel readout time. In a science image, we can only make use of the overscans to characterize the bias. Thus, we wouldn't have enough information to fully calibrate these two rows - subtracting the mean of the overscan region would leave this increasing gradient.

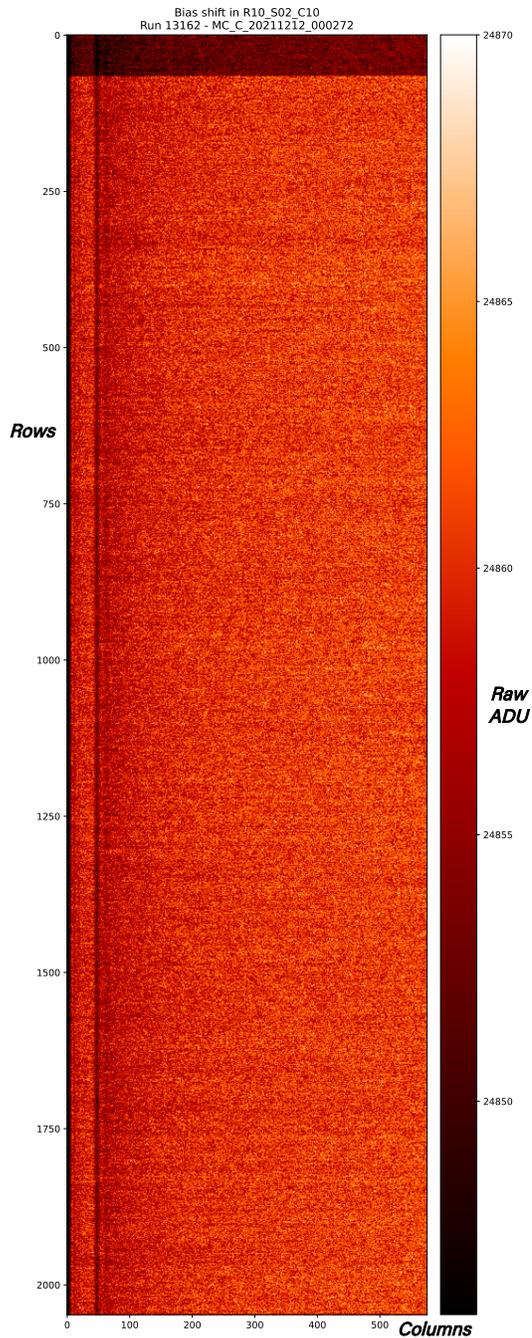
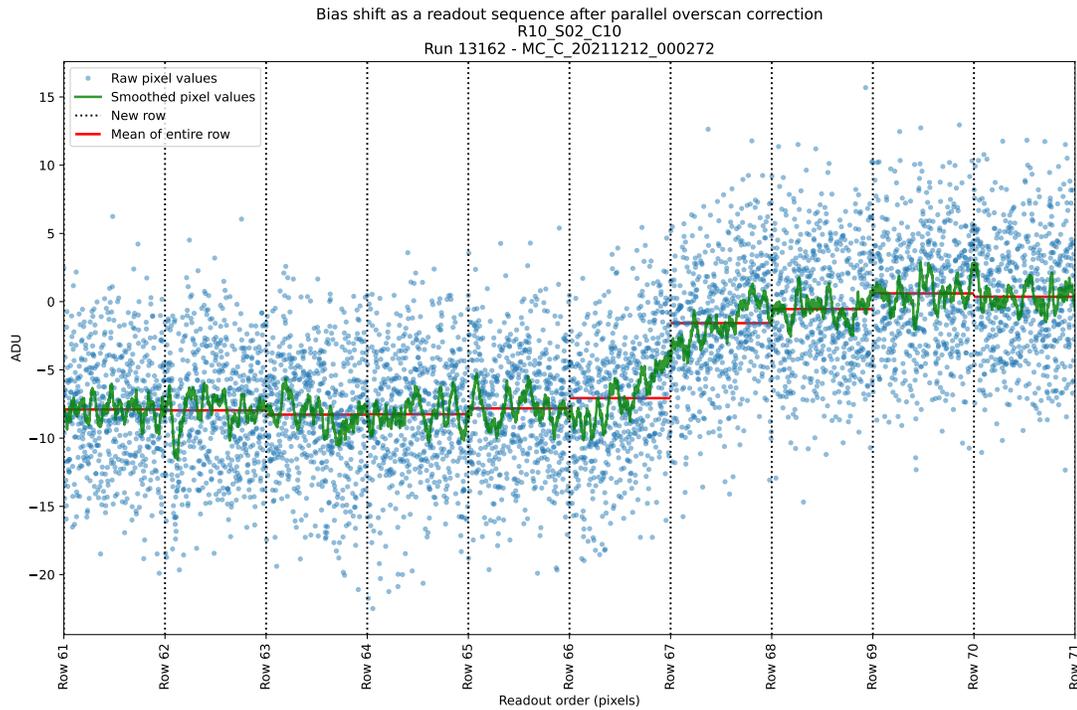


Figure 2.3: A sudden shift in bias from the amplifier in an LSST Camera CCD, seen as a brightness change in row 67. This example is not associated with any bright pixels. Although no photons are detected here, they would be measured with gain of $O(1)$ ADU/photon, making the bias level $\sim 25k$ times higher than the marginal contribution of a photon.

Figure 2.4: Figure 2.3 as a time series near the shift. This is different from the raw image as the parallel (left-right) structure has been removed. The blue dots show the individual pixel values, and the green lines are a moving average.



2.2.2 Bias shift detection algorithm

The shift detection algorithm is implemented in the Github repo `lsstcam-bias-shifts`, as well as in the main electro-optical test (EOTest) pipeline:

https://github.com/lsst-camera-dh/eo_pipe

1. Extract the serial overscans for each segment from an image, in readout order
2. Compute the mean M_i of each row's overscan, skipping the first `nskip` values, default 3 (to avoid CTI leftover charges)
3. Compute the high-frequency noise σ as the standard deviation of M_i after a 1st order Butterworth filter at 30 pixels.
4. Compute the convolution S_i of M_i with the `odd_local_kernel` k defined below, which detects rapid changes in average level before and after the center of the kernel.
5. Compute the shift detection metric $D_i = S_i/\sigma$.
6. Detect contiguous regions where D_i exceeds the `threshold` parameter, 3 by default. This corresponds to a shift in baseline above 3σ of the HF noise. Ignore any excursions in the first `skip_rows` rows, by default 30.

7. For each region, find the row with the highest value of D_i . Fit two lines in M_i for `window` pixels on either side of the peak. If the trend spans more than half the size of the shift in the same direction as the shift on either side, reject the shift. This is the “flatness” condition.

This kernel measures the difference in average level on either side of the center, with a stronger weighting near the center. It is normalized so that the result of the convolution approximates the average difference between the two sides, under that weighting scheme. The kernel is visualized in Figure 2.5.

```
def odd_local_kernel(window=30):  
  
    kernel = np.concatenate([np.arange(window), np.arange(-window+1,0)])  
  
    kernel = kernel/np.sum(kernel[:window])  
  
    return kernel
```

2.2.3 Association with bright pixels

The bias shifts were first recognized around 2019. They appeared as in Figure 2.6, with each shift appearing near a bright pixel defect. This was confirmed with an early version of the above algorithm, which found that there were no bias shifts in any of the test data without a nearby saturated pixel, above ~ 180 kADU. This was assumed to be causal. After a reduction in the gain of the readout amplifier, such that the bright defects stayed below ~ 180 kADU, another search showed

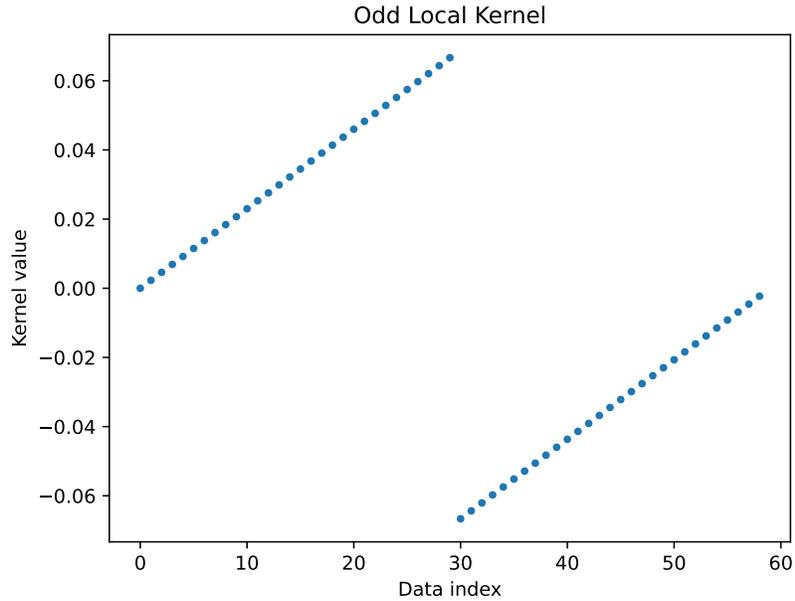


Figure 2.5: Visualizing the kernel used in bias shift detection.

there were no more bias shifts.

This effect was presumed to have been solved, until its recurrence in EO Testing Run 5 in late 2021. This time, there would be no such obvious causal connection. The algorithm ran over all test data, not just bias frames. There were a few particularly bad offenders — the “shiftest sensors” — and a broad distribution of shifts across the focal plane, with about half the segments displaying the behavior, and a quarter showing more than one shift (Figures 2.7,2.8).

Although an operational mitigation has not been found for this new manifestation of the effect, its implementation in the standard EO test pipeline means a plot is generated for every test sequence, allowing it to be easily monitored. As men-

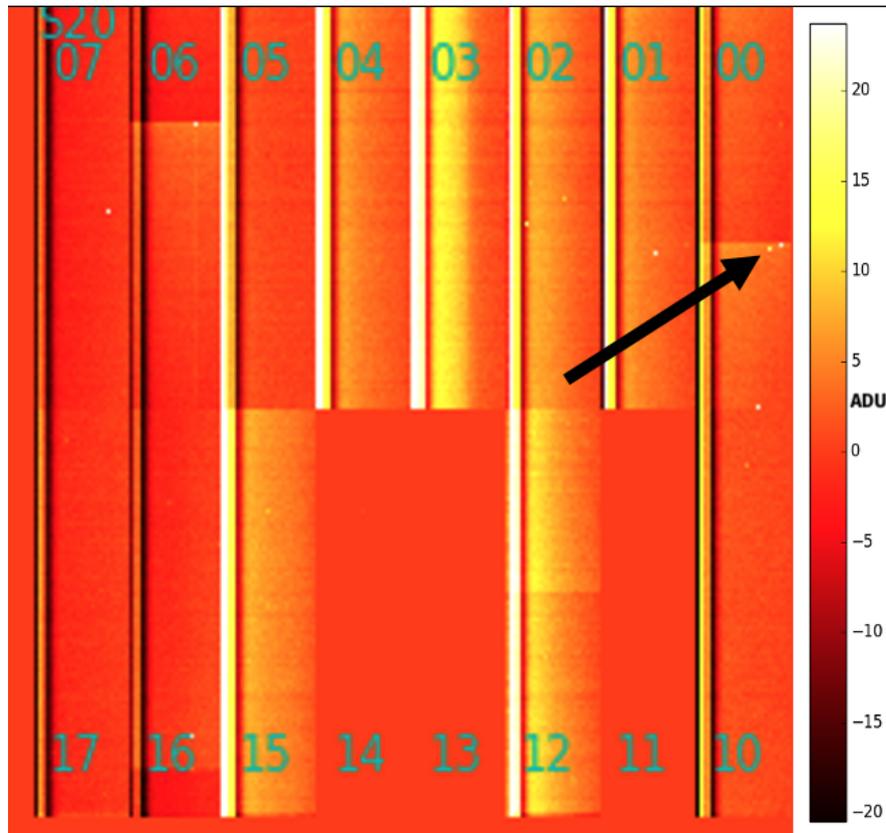


Figure 2.6: Bias shifts associated with a bright pixel. The vertical stripes are parallel structure.

tioned before, this issue demonstrates why serial overscan correction is necessary in certain cases. Since the locations of the shifts appear at random throughout the readout (Figure 2.9), we need the nearly real-time bias measurement provided by the serial overscans, taken after each row of data, to know what bias level to subtract.

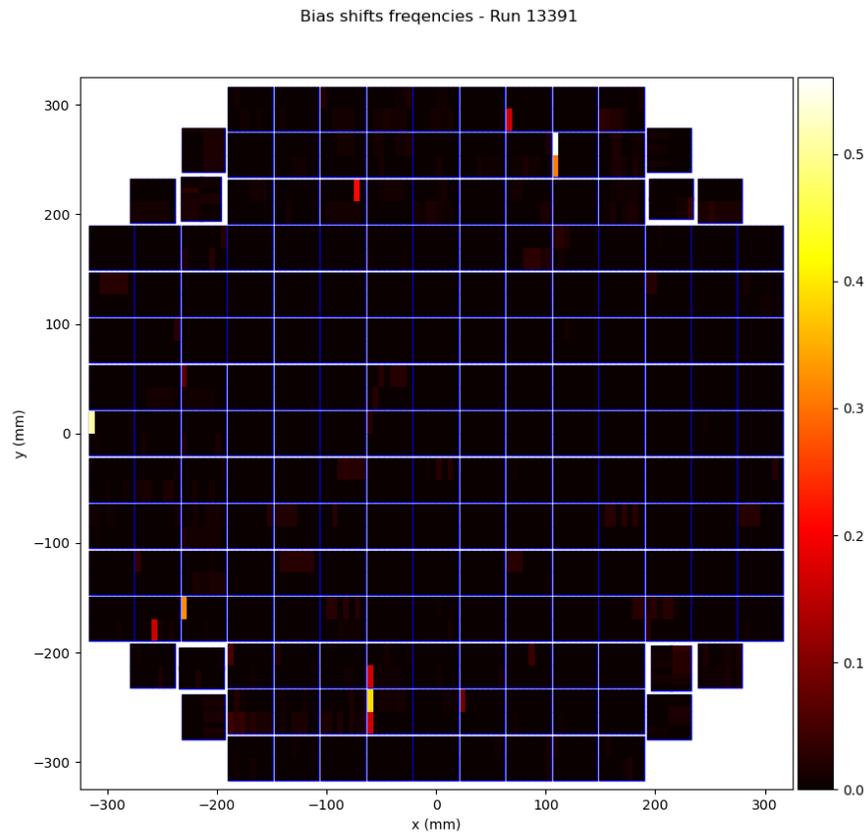


Figure 2.7: Distribution of the fraction of images with shifts for all amps in the focal plane. A few problematic sensors dominate the number of shifts.

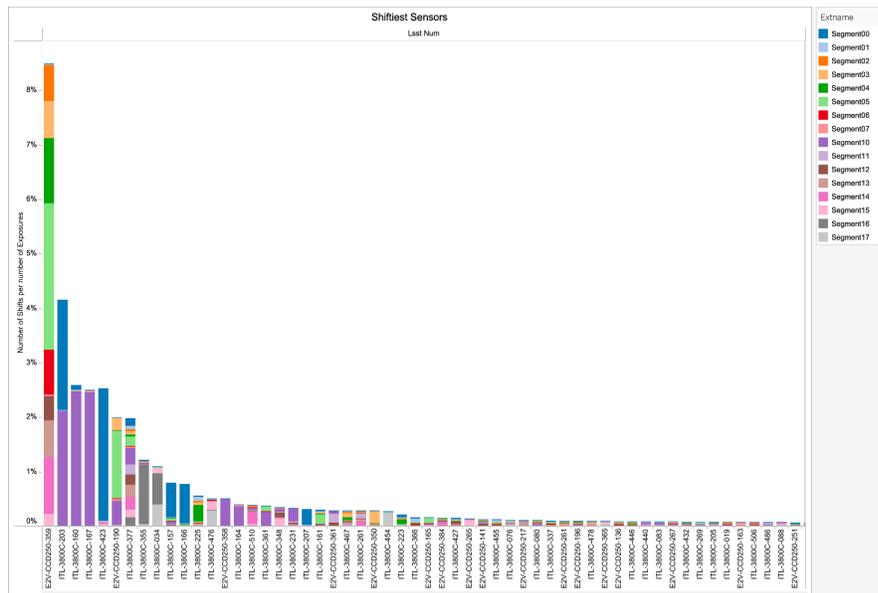


Figure 2.8: Shiftest sensors - distribution of the amplifiers (equivalently, CCD segments) exhibiting the most bias shifts, partitioned by sensors on the x-axis and segments in different colors.

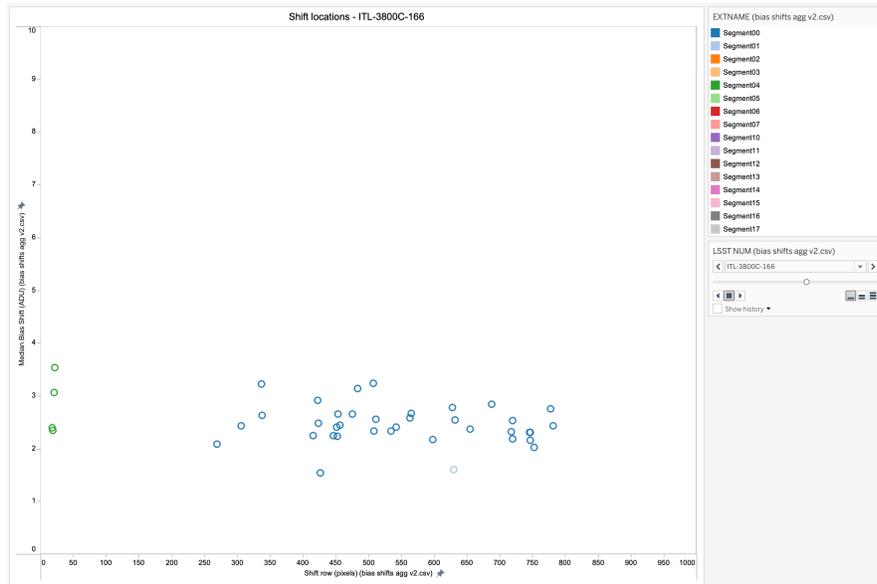


Figure 2.9: Distribution of shift locations and magnitudes for a sensor with shifts in many different rows in different images. Only two segments display shifts.

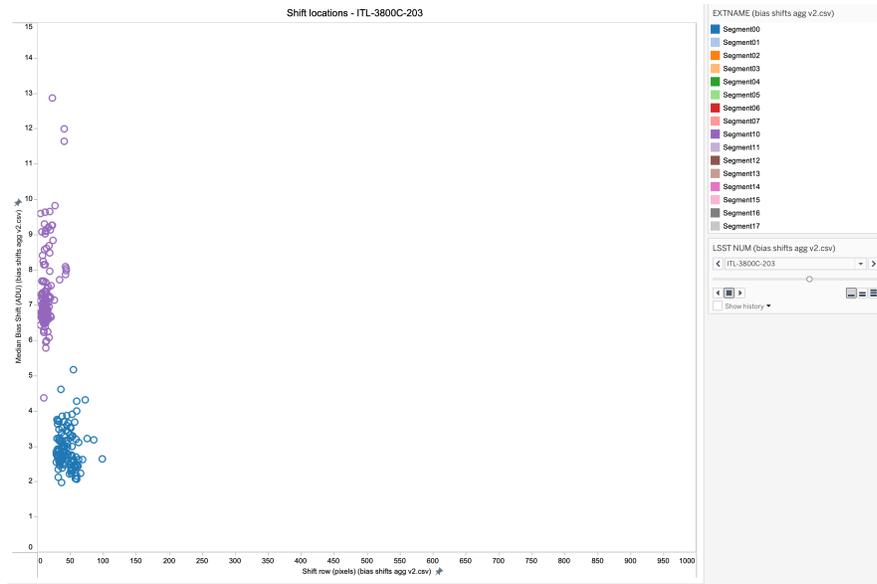


Figure 2.10: Distribution of shift locations and magnitudes for a sensor with shifts in nearby rows, although not perfectly consistent.

Chapter 3

Projected constraints on primordial black holes with the Legacy Survey of Space and Time

In this chapter, I use models for the dark matter profile and kinematics in the Milky Way, in conjunction with the simulated LSST SIM catalog [59], to predict total rates for microlensing events for LSST. I produce representative sets of such events with a Monte Carlo method. I then leverage the `rubinsim` [60] software to simulate LSST surveys, and its existing `MicrolensingMetric` to assess the detection efficiency of a set of simulated microlensing events.

3.1 Modeling the Milky Way

3.1.1 LSST SIM Catalog

The LSST SIM catalog was developed as a synthetic stellar catalog for galactic modeling in `rubinsim` [59], using the established TRILEGAL software [61]. This catalog simulates single stars based on evolutionary tracks from PARSEC v1.2S [62], COLIBRI PR16 [63], and additional post-AGB and white dwarf cooling tracks. Stellar sources are distributed according to a four-component Galactic model — thin disk, thick disk, halo, and bulge — each characterized by distinct star formation histories, metallicity distributions, and density profiles. These models were calibrated using photometric data from surveys such as the Deep Multicolor Survey [64], EIS data for the South Galactic Pole [65], and the Chandra Deep Field South [66].

TRILEGAL simulations rely on five primary input datasets: stellar evolutionary tracks, bolometric corrections for the photometric passbands, an initial mass function (IMF), parameterized star formation histories and age-metallicity relations for each Galactic component, and an analytic model of Galactic geometry, specifying the spatial distribution of stellar density and interstellar extinction. These inputs are combined to produce a realistic synthetic catalog of Milky Way sources.

3.1.2 Milky Way density profile of dark matter

In addition to the stellar contents of the Milky Way, one needs a model for the dark matter halo. In the case of PBHs, I take a density profile ρ and assume the compact objects are randomly distributed according to that profile. As implemented in `LensCalcPy` [67], I make use of an NFW profile for the halo:

$$\rho_{\text{NFW}}(r, \rho_s, R_s) = \frac{\rho_s}{(r/R_s)(1 + (r/R_s))^2} \quad (3.1)$$

with the scale density $\rho_s = 4.88 \times 10^6 M_\odot \text{ kpc}^{-3}$ and scale radius $R_s = 21.5 \text{ kpc}$, following [68]. This model is fit to a variety of observational constraints, including:

1. Satellite dynamics from observations of the Magellanic Clouds constraining the total mass within 100 kpc [69]
2. Outer rotation curve constrained by neutral hydrogen measurements [70] and classical cepheids [71]
3. Surface density of gas and stellar components at the solar radius [72]
4. Local vertical force K_Z at 1.1 kpc above the Galactic plane using kinematics of K dwarfs limit the total density of matter inside 1.1 kpc [72]
5. Oort's constants constraining the circular velocity curve at the solar radius [69]

6. Terminal velocities constrain the mass distribution and rotational velocities inside the solar radius [73]
7. Kinematics of OH/IR stars in the bulge constrain the mass within $\sim 100\text{pc}$ [74]
8. Total K-band luminosity $L_K = 8.9 \times 10^{10} L_\odot$ from [75]

3.2 Generating microlensing events

3.2.1 Simulation with PopSyCLE

The state-of-the-art tool for generating microlensing events is PopSyCLE [76].

The procedure is as follows:

1. Simulate a stellar catalog with Galaxia. The catalog has a circular footprint on the sky, or a cone in space. For each source, it generates: “position, velocity, age, mass, photometry in several filters, 3-D extinction, metallicity, surface gravity, and more”.
2. Generate compact objects. First, star clusters are simulated with PyPopStar [77]. Then, using an initial-final mass relation (IFMR), a catalog of white dwarfs, black holes, and neutron stars is generated. This is done such that the populations are consistent with the Galaxia simulated stellar catalog in kinematics, size, age, etc.

3. Simulate a microlensing survey, by linearly evolving all the objects according to their kinematics. Bin the objects by their on-sky positions, and at each sampling time, check for object pairs within some minimum impact parameter.

3.2.2 Monte-Carlo approach

The approach employed by PopSyCLE generates a complete list of sources and lenses (**Population Synthesis**), and evolves their positions through time. At first, it's an attractive idea to generate a catalog of objects that trace the density profile of interest. It gives concrete examples of lensing events and a realistic picture of the history of these persistent objects. However, this requires a massive amount of memory and computation time. It generates objects tracing the *entire* density profile, even though only a very small fraction of objects exhibit the right alignment event with a source. Furthermore, I would argue that in the case of microlensing events, the utility of the information gained from this procedure is minimal. Considering how extraordinarily infrequent the events are per source, the odds of the same unbound object lensing two different sources are vanishingly small. Therefore, with a density and velocity distribution for the lenses, it makes sense to simplify the problem to one where each source independently samples from a Poisson distribution with a rate matching the analytic event rate.

With a dark matter model for the Milky Way halo in hand, one can predict how many lenses pass through a particular pointing in a given amount of time. Instead of constructing one specific realization of the DM model, I leverage the fact that the rate of lensing Γ is a small parameter. The problem is factorized into one where each source has its own rate according to its position on the sky, distance from the observer, and the DM profile along the line of sight. This defines an event space of a few variables with a corresponding probability function, which can be sampled via standard Monte Carlo methods. MC methods are fast, and instead of following $\sim 10^{11}$ objects through the galaxy, I generate only the $\sim 10^6$ that generate microlensing events, plus $\sim 50\%$ of rejected samples in the MC procedure. This reduces a months-long computation with PopSyCLE on a computer cluster down to minutes on a laptop.

3.2.3 Estimating the density of events at one pointing

Supplementary Section 1.2 of Ref [41] contains a full derivation of the event rate. I provide a summary here.

Assume the PBHs comprising the DM halo follow an isotropic Maxwellian velocity distribution:

$$f(\mathbf{v}; r) d^3\mathbf{v} = \frac{1}{\pi^{3/2} v_c(r)^3} \exp\left[-\frac{|\mathbf{v}|^2}{v_c(r)^2}\right] d^3\mathbf{v} \quad (3.2)$$

and assume the velocity dispersion $v_c(r)$ to be:

$$v_c(r) = \sqrt{\frac{GM_{\text{NFW}}(< r)}{r}} \quad (3.3)$$

where $M_{\text{NFW}}(< r)$ is the mass within radius r of the center of the Milky Way halo. The differential rate of events is given by:

$$\frac{d\Gamma}{dM dd_L d\hat{t} du_{\min}} = \frac{2}{\sqrt{u_T^2 - u_{\min}^2}} \frac{v_T^4}{v_c^2} \exp\left[-\frac{v_T^2}{v_c^2}\right] n(d_L) f(M) \varepsilon(\hat{t}) \quad (3.4)$$

with $v_T \equiv \frac{2R_E}{\hat{t}} \sqrt{u_T^2 - u_{\min}^2}$, where \hat{t} is the Einstein crossing time, u_{\min} is the minimum impact parameter of the lens, u_T is the maximum value of u_{\min} to consider an event, d_L is the distance of the lens from Earth, $n(d_L)$ is the number density of lenses along the line of sight, $f(M)$ is the mass function of lenses, and $\varepsilon(\hat{t})$ is the detection efficiency. This study looks at delta function mass distributions, and does not assume a detection efficiency, rather evaluating it with simulated surveys.

`LensCalcPy` [67] provides tools to evaluate the resulting integral, including implementations of the NFW and velocity dispersion profiles. By integrating this equation for the desired bounds, I get a predicted rate of microlensing events.

3.2.3.1 Using differential rates as Monte Carlo probabilities

With `LensCalcPy` [67], one can evaluate integrals for the rate of microlensing events, with the derivative given by Equation 3.4. For a given domain of this

integral, it provides a number of events expected per unit time Γ for a source. Calculating this integral for a number of different sources, keeping in mind the relevant boundaries, will yield a total microlensing rate across those sources. Given a stellar catalog, one could compute the total number of events across the sky from that survey by evaluating Γ_i for each source, labeled by i . This would schematically look like

$$\Gamma_{\text{tot}} = \sum_i \int d\Gamma_i \quad (3.5)$$

This rate is static in time, so we can imagine an arbitrarily long survey with observing time T , and predict it could see $N = T\Gamma_{\text{tot}}$ events. If we were to discretize the above integral, we could take the constant factor T inside to get

$$N = \sum_{i,j} T \Delta\Gamma_i(\vec{\theta}_j) \quad (3.6)$$

where $\vec{\theta}_j$ is a coordinate in the domain of the integral. Taking T large enough, we start to see that a catalog making up all the events of the survey would contain a number of events at each point in parameter space $\vec{\theta}_j$ proportional to the differential rate at that point. Now we can think of the differential rate as a probability density in the same space as the total rate integral, including a discrete choice of which source we're looking at.

Computing the total rate of Equation 3.4 is expensive. Figure 3.1 demonstrates that the total is stable under randomly sampling a sufficiently sized subset of the

sources. But the differential rate $d\Gamma$ is fast enough (after some optimization) to compute many times in a Monte Carlo simulation. The Metropolis-Hastings algorithm [78] gives a standard, straightforward way to draw samples from the high-dimensional event space with density proportional to $d\Gamma$. I use a trivial (uniform) proposal function over the source index, lens distance, crossing time, and minimum impact parameter, which is fast enough for our needs.

3.2.3.2 Implementation in Python

When scanning the sky for microlensing events, one will essentially be looking for sources with consistent baselines which have some characteristic transient amplification. The LSST SIM catalog provides this list of baselines. Dimmer sources do not contribute here ¹. Then, based on a model of PBHs, one can compute the rate at which a lens passes in front of each source.

To generate microlensing events, I adopt a Monte Carlo procedure with discrete options for the pointings given by the catalog, and the probability densities given by the differential rate within the subspace of the remaining parameters: lens distance from Earth, crossing time, and impact parameter. This is implemented

¹Dimmer sources don't contribute as *sources*, but they certainly will contribute as *lenses*!

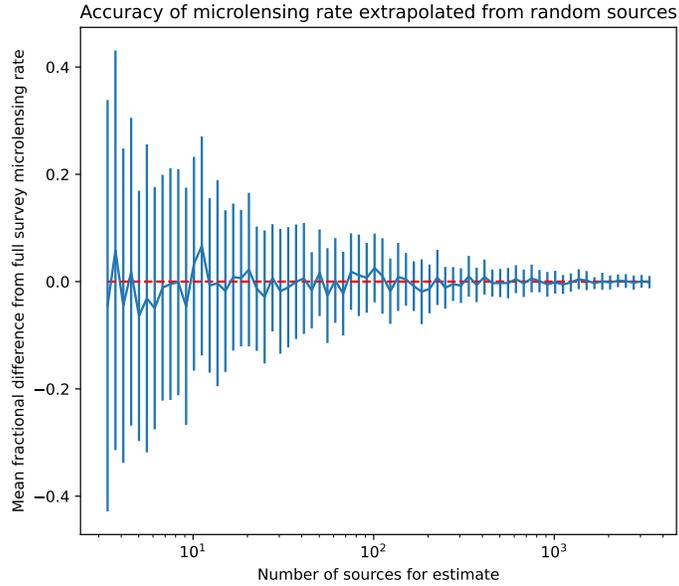


Figure 3.1: Demonstrating the stability of the total on-sky rate calculation estimated from using small subsets of the available sources. A “full” rate is calculated with 336265 sources, corresponding to the number of expected events for a particular simulation (though any number could have been used here). Subsets of these sources were taken with sizes specified by the x-axis, and 200 random subsets were taken for each size. The central line is the (signed) fractional difference between the average total rate of these 200 subsets and the full rate, and the error bars are the average fractional distance of each subset from the full rate. Since there are many other systematic and modeling uncertainties in this simulation, $\lesssim 10\%$ error would be acceptable. The errors from the subset sizes used here are considerably better.

via the following:

1. Uniformly randomly select sources from the LSST SIM catalog.
2. For each source, use its pointing and distance to calculate the microlensing event rate via the methods in `LensCalcPy`.
3. Sum the calculated rates and rescale by the fractional size of the subsamples compared to the full catalog to estimate the total rate across the survey.
4. MCMC sample from the sources using differential rates as probabilities as described in Section 3.2.3.1, to generate a representative set of microlensing events.
5. Use the original simulated magnitudes as baselines for the events within the `Microlensing Metric` in `rubinsim`, described in 3.2.4.2.
6. Use the detection efficiency calculated in `rubinsim`, combined with the predicted total rates, to find the total number of expected detections for the survey.

3.2.3.3 Resource usage/complexity comparison to simulation

Generating microlensing events via Monte Carlo is extremely fast. On an M1 Pro MacBook, depending on the parameter boundaries, typical speeds are 2000-5000 samples per second. The memory required is only as large as the baseline

source data, and the generated events. For the results shown, these are on the scale of a couple GB at most.

On the other hand, PopSyCLE is a job for a computing cluster. Depending on the sky pointing, generating hundreds of events (out of the hundreds of thousands predicted) took weeks on 64 nodes with hundreds of GB of memory, and produced hundreds of GB of data. Most of the data produced and moved around is not useful for PBH microlensing studies — age, metallicity, surface temperature.

3.2.4 Extension to generic cadences

3.2.4.1 rubinsim for simulating proposed LSST strategies

There is still much discussion about Rubin’s survey strategy. As described in Section 1.2.1, since there are many different goals for the survey, it takes careful thought to design a single strategy that balances the potential scientific results. In order to make predictions about the performance of the survey, members of the Rubin community developed `rubinsim` [60]. This set of software tools provides a framework to evaluate arbitrary metrics over different realizations of the LSST, including different survey strategies. It can simulate information about the dynamic observing conditions across the full survey, including seeing information, sky brightness, pointings, filter usage, timing and more. The Metric Analysis Framework (MAF) includes dozens of metrics to characterize the effects of differ-

ent survey strategies on the various studies one might perform on the resulting images. By comparing the improvement or decline of these metrics across simulations of proposed cadences, one can learn about the scientific trade offs inherent to choosing a strategy, and make more informed decisions about what to implement when the observatory comes online.

3.2.4.2 `MicroLensingMetric` for detection efficiency

One of the metrics implemented within `rubinsim` is the `MicroLensingMetric` [79]. It takes a list of sources with their magnitudes and positions, and a set of microlensing event parameters, and within the context of a simulated survey, calculates whether the lensing amplification of the event contains enough statistically significant measurements to constitute a detection. The options available for output of the metric include

1. `Npts` - returns the number of observations within 1 crossing time of the event peak
2. `detect` - returns a boolean, whether N observations are present before the peak, N defaulting to 2
3. `Fisher` - returns the fractional confidence on the crossing time, σ_{t_E}/t_E based on a Fisher matrix analysis

Whereas `detect` seems to have followup observations with other instruments in mind, the other two options can be used to estimate how detectable each event is.

I modified this metric to include Poisson noise in the signal-to-noise calculation, which is necessary to cut out undetectable high-impact-parameter events on dim sources.

3.3 Results

Figures 3.3 and 3.4 show the numbers of alignment events predicted for the 10-year LSST. These contain no considerations about the cadence of the survey, but do come with integration boundaries including: $u_{\min} \in [0, 5]$, $t_E \in [1 \text{ hour}, 20 \text{ years}]$. These should be thought of as a normalization for the integral of the density traced by the Monte Carlo procedure. For example, to simulate all the events in a survey, one would compute the `rate_total` N as in Equation 3.6, then generate N samples with the MC procedure.

To predict the number of “in-principle” detectable events, I make use of `rubinsim` for a particular cadence. I compute these numbers using the baseline survey strategy, but the efficiency only changes a few percent with different strategies. Figures 3.5 and 3.6 show the theoretical detection efficiency from the `MicroLensingMetric`. The normalized comparison indicates that alternative sur-

vey cadences can improve efficiency by 1–4%, especially in the lower-mass regime where events have shorter crossing times.

These efficiency estimates make use of the seeing of the observatory and the ideal microlensing lightcurves, but do not include information about the measured variations in brightness from stars. In the next chapter, I demonstrate detection on real lightcurves from an existing dataset. The efficiencies here are dependent on the parameter boundaries selected in the MC procedure. For example, one can generate events up to arbitrarily high impact parameters, but above $u_{\min} \sim 5$, these couldn't be detected. Setting a threshold of $u_{\min} = 10$ instead of $u_{\min} = 5$ would cut the efficiency roughly in half, because the events are distributed evenly in u_{\min} . However, the total number of detections would stay the same, because the total rate normalization would compensate for the drop in efficiency. In this sense, the raw efficiency numbers here are not the result of interest.

Figure 3.2 confirms the $t_E \propto M^{1/2}$ scaling expected from the lensing geometry, with color indicating the event density across crossing times. The normalization across masses in the figure isolates the effect of PBH mass on total event rate.

Comparing Figures 3.3 and 3.4 shows the impact of source selection. Inclusion of distant sources (e.g., from the Magellanic Clouds) boosts the total event rate and enhances sensitivity to lower mass PBHs. In the final results, I exclude these Magellanic Cloud sources, because a good prediction for the event rate would

require a model of the subhalo, which is not included here. The additional lenses in the subhalo would increase the rate, so this is erring on the conservative side. Exclusions would be even stronger after including this population.

In summary, when combined with realistic survey parameters via `rubinsim`, this Monte Carlo method suggests that LSST will be sensitive to PBH microlensing events across a range of masses, particularly in the few decades near $10^{-1}M_{\odot}$. These results provide the simulated events that will be applied to real lightcurve data in the next chapter.

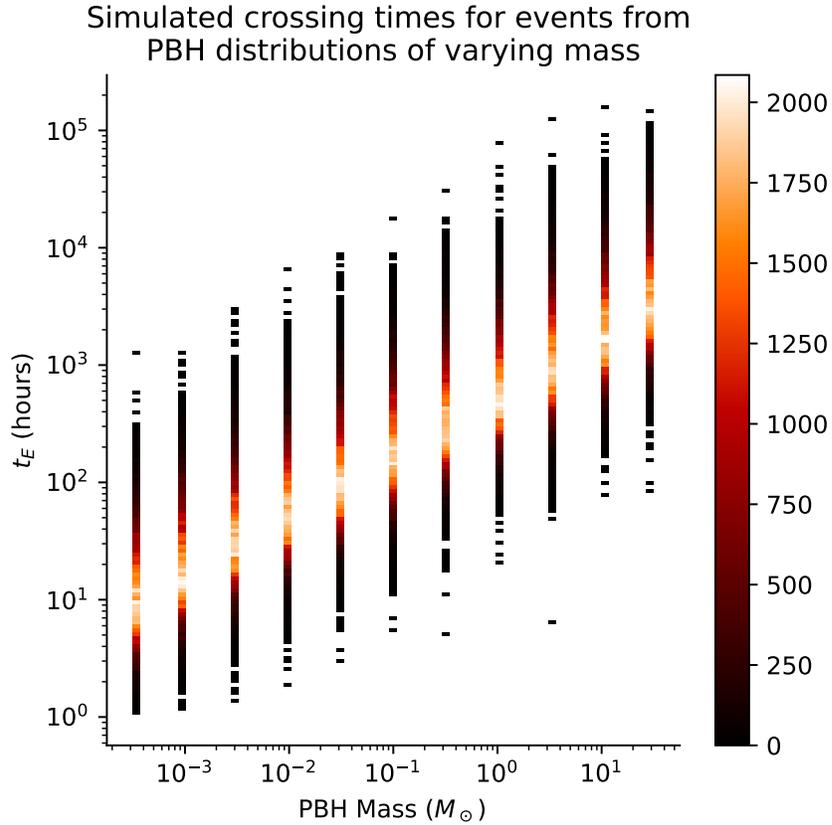


Figure 3.2: Showing the $t_E \propto M^{1/2}$ relationship in simulated events, without considering detection efficiency. Events were cutoff below 1 hour and above 20 years. For easier comparison, total number of events is normalized across different masses, though in reality the rate varies. Color corresponds to the density of events at a crossing time.

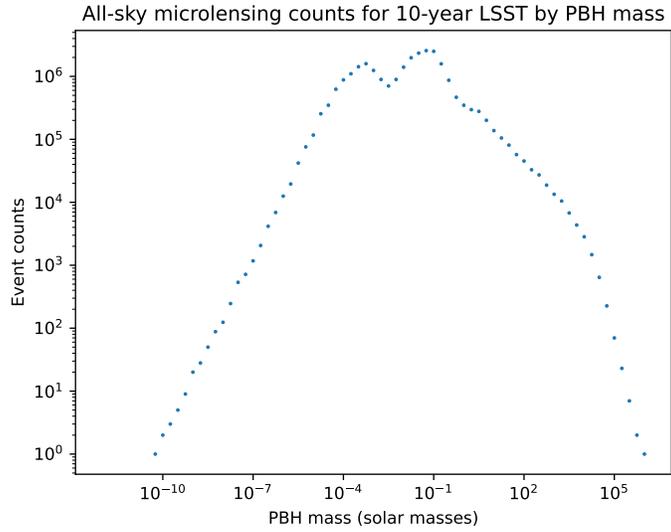


Figure 3.3: Total number of in-principle detectable microlensing events across the sky for populations of delta-function mass distribution PBHs comprising 100% of Milky Way dark matter. This includes sources from the Magellanic Clouds, but not PBH lenses. Including a subhalo with these lenses would increase the low-mass sensitivity, although these events are mostly too fast for Rubin’s cadence.

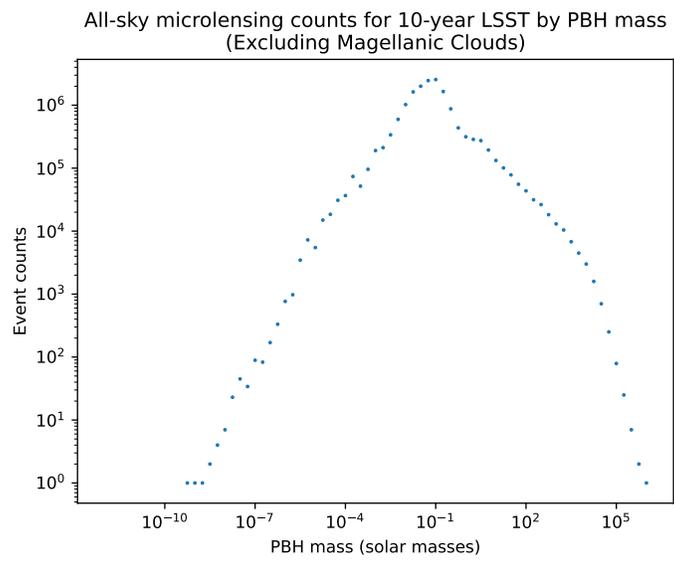


Figure 3.4: Same as 3.3, excluding the most distant sources corresponding to the Magellanic Clouds.

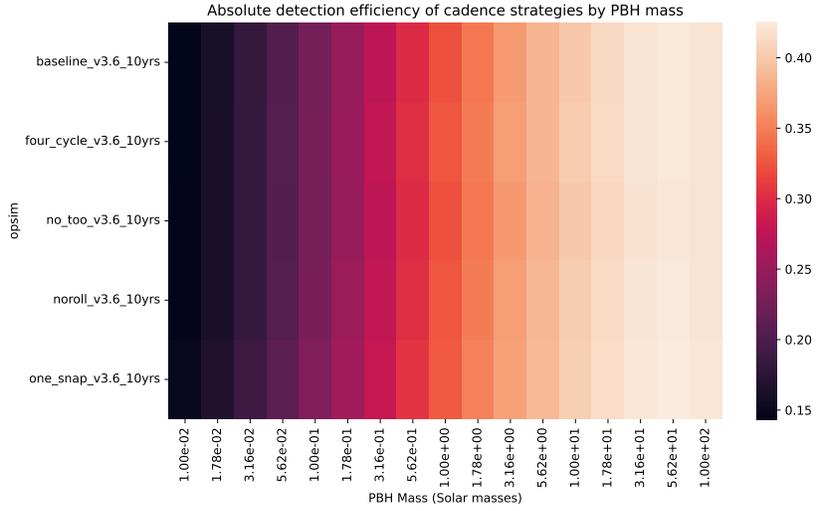


Figure 3.5: Absolute theoretical detection efficiency from the MicrolensingMetric in opsim.

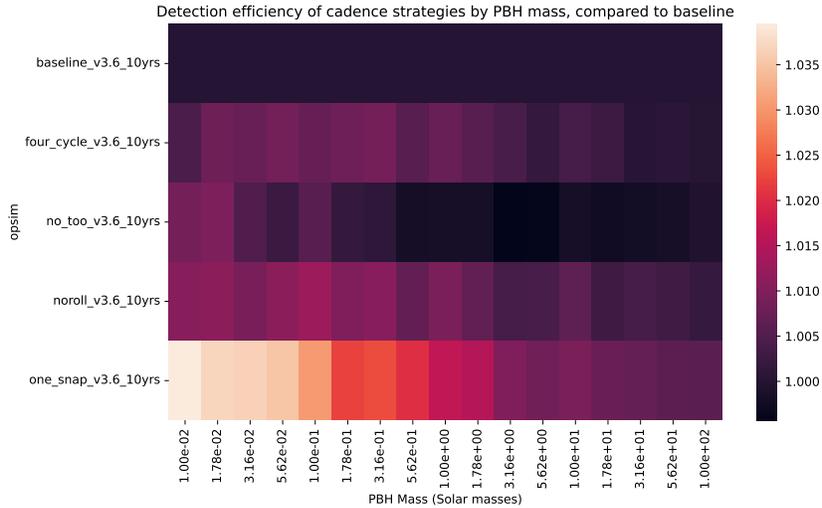


Figure 3.6: Comparison of theoretical detection efficiency for difference cadences, normalized to the `baseline_v3.6` run. Each alternative survey strategy has PBH masses with modest 1 – 4% increases in detection efficiency. This is particularly pronounced in the low-mass ($10^{-2} M_{\odot}$) range, which are centered at lower crossing times.

Chapter 4

A novel microlensing detection algorithm

4.1 Time series analysis

Much of physics is formulated in terms of the evolution of states through time. Developing techniques to analyze time-dependent data is central in science. This is broadly called *Time-Series Analysis*. It applies to any situation with measurements y_i taken at times t_i , and provides tools to understand the relationships between the two.

One of the most powerful and straightforward techniques is *convolution*. Although perhaps opaque at first glance, convolution is a wonderfully rich and ver-

satellite tool in signal processing, and serves as the basis for a wide swath of other techniques. In Section 2.2.2, I describe an algorithm that searches for a particular behavior in the bias level of a CCD amplifier. It is implemented as a convolution between the bias signal and a specially designed kernel. In this chapter, I describe a microlensing detection method implemented as an unevenly-sampled convolution over magnitude measurements using a different kernel. Since the technique is so central to these methods, I give a general introduction here.

The continuous convolution for a function $y(t)$ and a *kernel* $g(t)$ is defined as:

$$(y * g)(t) = \int_{-\infty}^{\infty} y(\tau)g(t - \tau) d\tau$$

Although I have distinguished the roles of y and g , it's easy to show (*u*-substitution) that the operation commutes: $y * g = g * y$. In signal processing, we pick a kernel to either perform a *transformation* to y or *detection* on y . This dissertation contains examples of both.

When working with real data, we need a discrete analog. The most common situation in practice has data which is evenly sampled, i.e. $t_{i+1} - t_i = \Delta t$, a constant sampling rate.

4.1.1 Evenly sampled

Let's consider samples $y_i = y(t_i)$ for some t_i , $i \in \{0, \dots, N - 1\}$ sampled evenly at intervals of Δt such that $t_{i+1} - t_i = \Delta t$, and a discrete kernel g_i with the same

timescale. Although g can have any length, the easiest to interpret has an odd number $2M + 1$ of entries, $j \in \{-M, \dots, M\}$. This gives a natural center of the kernel at g_0 . We can derive a discrete version of the above formula:

$$(y * g)(t_i) = \sum_{j=-M}^M y(t_j)g(t_i - t_j) \quad (4.1)$$

$$(y * g)_i = \sum_{j=-M}^M y_j g_{i-j} \quad (4.2)$$

This can be expressed as a matrix product by defining $\mathbf{g}_{ij} \equiv g_{i-j} = g(t_i - t_j)$:

$$\vec{y}_g \equiv \vec{y} * \vec{g} = \mathbf{g} \vec{y} \quad (4.3)$$

The above equation requires some further specification. Although there are different useful choices for the shape of the resulting convolution, relating to choices about boundary conditions, in the use cases presented here, the shape of \vec{y}_g will be the same as the input data \vec{y} . In any case, one dimension of \mathbf{g} does need to match \vec{y} . So here, we consider \mathbf{g} to be an $N \times N$ matrix. For the elements far from the diagonal \mathbf{g}_{ij} where $|i - j| > M$, we implicitly take the values to be 0. We also typically specify a kernel that's smaller than the input data. In the case of even sampling, the rows of this matrix look like the original kernel, transposed such that the center always lines up on the diagonal elements.

Even though modern computers are quite good at computing matrix products

like this, the convolution is generally much easier than a generic matrix product. This is because (a) it usually has a lot of entries that are 0 (away from the diagonal), and (b) in the evenly sampled case, every row contains the same few values transposed, allowing for easy cache locality optimizations. However, we don't even need to think about the matrix in this case. The simple form of (Equation 4.2) is much faster to compute than the naive general form of (Equation 4.3). It can be implemented as a cumulative sum along a sliding window, resulting in complexity of $O(NM)$, instead of $O(N^2)$. In the case of a kernel much smaller than the input data, this is a massive difference in performance.

4.1.2 Unevenly sampled

The situation is more complicated with unevenly sampled data. In the evenly sampled case, the kernel “scans” over the input data, and encodes something about the relative distances between datapoints. Since the distances of adjacent points are always the same, the same list of kernel values applies across the whole list of data. When the data has different distances in time between points, we are forced to go back to a continuous function for the kernel to deal with the continuum of possibilities for $t_i - t_j$. We can still compute the convolution as the matrix product Equation 4.3. However, the matrix no longer has the same simple form. It must be computed explicitly from its definition, $\mathbf{g}_{ij} \equiv g(t_i - t_j)$. The

one simplifying fact is that the diagonal of the matrix still contains the center of the kernel $g(0)$. Additionally, the matrix inherits symmetries from the kernel, such as an even kernel making a symmetric matrix and an odd kernel making an antisymmetric one.

The product has no fixed windowed sum implementation as described in Section 4.1.1, because each row depends on the distances between the unevenly sampled times. There is no way to avoid computing the distinct elements in each row. There are a couple optimizations to avoid computing large matrix products, though.

Consider a kernel $g(t)$ with support over a finite range, $t \in [-a, b]$, and some datapoints y_i ordered by their corresponding times t_i . We want to evaluate the convolution $g * \vec{y}$. A simple optimization here would be to implement a window analogous to the evenly sampled case. Although a fixed window is impossible, we can use dynamic windowing to achieve similar complexity. The idea is as follows:

1. Iterate over sample times t_i from $i = 0$ to N .
2. Initialize the window: begin with a window starting at $j = 0$ for the smallest t_i . This window will contain indices for which $g(t_i - t_j)$ is non-negligible.
3. Advance the window's lower bound: increment the window's starting index until the condition

$$t_i - t_j \leq a$$

is satisfied. Terms with $t_i - t_j > a$ can be excluded from the convolution sum.

4. Accumulate contributions: for each j within the window, compute the kernel value

$$g_{ij} = g(t_i - t_j)$$

and add the product $g_{ij} y_j$ to $y_g[i]$.

5. Update the window for subsequent iterations: adjust the lower window edge to ensure that for the next target time t_{i+1} , the window starts at the first index where the time difference satisfies the kernel's support threshold (e.g., $t_i - t_j \geq b$, if a lower support bound b is specified).

This procedure will work for any kernel with finite support, and the number

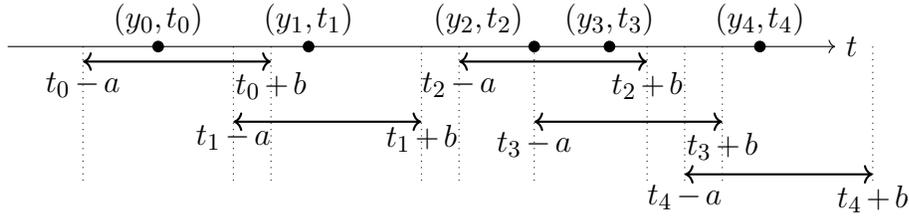


Figure 4.1: Visualizing the windowed approach to computing the convolution of an unevenly sampled dataset (y_i, t_i) .

of iterations is cut in half if the kernel has parity symmetry. Furthermore, most kernels for signal processing can be clipped to a finite domain. For the Weighted Moving Average in Section 4.4.1, the Gaussian kernel is symmetric and falls off such that I can, without consequence, clip it to 0 at 5 standard deviations. Appendix A.1 contains an explicit implementation in Python.

4.2 Precursor dataset: NOIRLab Source Catalog

By using a precursor dataset with stellar lightcurves, I can set an upper bound on the backgrounds to the signal of interest in a PBH search. In order to make exclusions (or detections) of compact dark matter, one needs to characterize our ability to find microlensing events in the data, considering both detection efficiency and background rejection. These events have a very particular shape, given by

the family of curves derived in Section 1.4.1. There are no known astrophysical phenomena which have exactly this shape. However, our instruments aren't perfectly precise, stellar sources aren't perfectly stable, and we're working with finite revisit times. Our ability to distinguish these events is limited to the properties of our dataset, and is fundamentally statistical. One *can* do quite a good job of it, as shown in this chapter, but it's important to understand just *how* well, and in what situations.

To this end, I apply the search algorithm to data from the NOIRLab Source Catalog (NSC) [4]. This is an aggregate dataset in seven filters ¹ from several ground-based telescopes taken over seven years. Even though it doesn't contain enough data to expect to find many microlensing lightcurves, by applying a microlensing search to this precursor data, one can characterize the real lightcurves that might issue false positive detections when scaling up to the massive volume of measurements expected for the LSST. These potential false positives are considered *backgrounds* to the microlensing signal. By showing there are no microlensing detections in this dataset, I derive an upper bound on the background rate for these events, and set a corresponding lower bound on the exclusion power of LSST, when combined with our models' predictions.

The idea is simple: I assume there are no microlensing events in this dataset. Based on the rates calculated in Section 3 with 100% compact DM, there is not

¹Similar, but not identical, filters to Rubin

enough volume of time \times sources to see one detection. So, if the algorithm flags any detections which can't be ruled out as microlensing, either there is a successful, however unlikely, detection, or the background to the real signal is too high to perform this analysis. If the algorithm manages to scan all lightcurves without a detection, then I've shown there is substantial power in the algorithm to see through the background, and there's an upper bound to the background rate.

Cutting out all backgrounds is useless if the same algorithm doesn't detect real events. One also needs to measure its *detection efficiency*. A real microlensing event should closely follow the ideal point source-point lens curve. This is a straightforward amplification of the flux in every color band. Barring nonlinear instrumental effects and source blending, if one of these events were to happen for any the sources in our dataset, one would see the same lightcurve, but with the ideal microlensing curve superimposed (in magnitudes). Therefore, a slick way to measure detection efficiency is to take exactly the same curves used to assess background rejection, add the ideal microlensing curve, and run exactly the same detection pipeline. This puts it on exactly the same footing for measurement cadence, photometric repeatability, source variations, and all the details that come with real data. By dividing the detection efficiency by (1 minus) the rejection rate, one has an analog to a signal-to-noise (S/N) ratio. Furthermore, the detection efficiency needs to be taken into account when computing model exclusions. If

the model predicts some number of microlensing events, one can't infer its fit to data without understanding if we expect to detect all of them, 1% of them, or somewhere in-between. Although one can't really know the efficacy until looking at real Rubin data, this analysis will demonstrate that the search algorithm really does pick out microlensing events and ignores the rest.

It's worth noting again that the detection efficiency here is different from that evaluated in Chapter 3. The `MicrolensingMetric` measures if there is enough statistical significance in the microlensing amplification during observation with some simulated seeing for a particular source. In this chapter, the question also involves background rejection, and with limited measurements on dynamic sources, there will certainly be some false negatives. These aren't possible in the prior assessment.

4.2.1 Compatibility with Rubin lightcurves, criteria on measurement frequency

The following conditions define the domain of data searched by the detection algorithm described in Section 4.4.

1. Cut 1. Filter out objects using provided columns in NSC `object` table.
 - (a) Requirement for the NSC star classifier: `class_star` > 0.9
 - (b) Require more than 20 detections: `ndet` > 20.

2. Cut 2. Filter on the cadence of measurements in NSC for each object to have datapoints on the right timescale for relevant events for Rubin. Called the “well-sampled” requirement. Requires object to have at least one sequence of measurements with the following properties:
 - (a) Maximum revisit time of `maxrevisit` ≤ 10 days.
 - (b) Number of detections `seqlen` ≥ 10 .
 - (c) Time spanning `interval` > 50 days.

These cuts are not a part of the detection algorithm. They are used to find a reasonable dataset that can give us an idea of the background rejection capabilities of a survey like LSST, which will have revisit times of ~ 3 days on average. A key difference is that LSST will maintain a cadence comparable to this for 10 years, for basically every persistent source like the stars targeted here. Because of this, we should have excellent baseline characterization for every source, except extraordinarily slow variable sources and microlensing events. In contrast, the NSC dataset does not contain many sources sampled so quickly and for so long. In order to target median events for $1M_{\odot}$ MACHOs, I am looking to detect events with crossing time $t_E \sim 40$ days. For this reason, I look for “well-sampled” (WS) regions of time that span at least 50 days, to get somewhat beyond the characteristic timescales of main interest. We will see that since the majority of WS regions in this subset span the lower range of this requirement, i.e. around 50

days, the detection efficiency drops off for longer timescale microlensing events. There is simply not enough information in the lightcurve to constrain the fit. However, these events do seem to make it past the first round of filtering, which doesn't rely on the quality of fit to any family of curves, rather only requiring a persistent excursion in all available filters.

4.2.2 Footprints of objects in consideration

The NSC dataset happens to contain enough coverage that the models would expect to see somewhat less than 1 microlensing event if the PBH mass is within a few decades of $10^{-1}M_{\odot}$, given the sources are randomly selected across the sky. However, the pointings measured in this dataset are clustered, with strong preference *against* the galactic bulge ² (Figure 4.2, compared to Figure 4.3 for Rubin). This means a far lower chance of finding a real microlensing event. But as far as assessing backgrounds, this is a feature, not a bug. One can measure just as many sources, which are unlikely to contain any microlensing events, but should contain similar levels of variability as the sources toward the bulge. Though there are differences in the stellar populations at different pointings in the Milky Way [80], microlensing rates roughly trace the density squared, and cover many orders of magnitude at different pointings.

²This is just a result of the selections from various studies included in the dataset. It wasn't designed for any coherent scientific goal since it's an aggregate from many instruments and many projects.

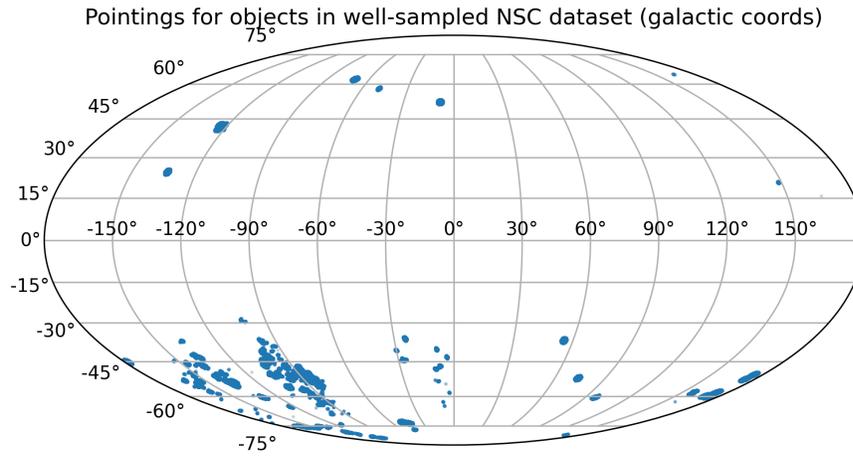


Figure 4.2: Spatial distribution of NSC sources on the sky, with galactic center at (0,0)

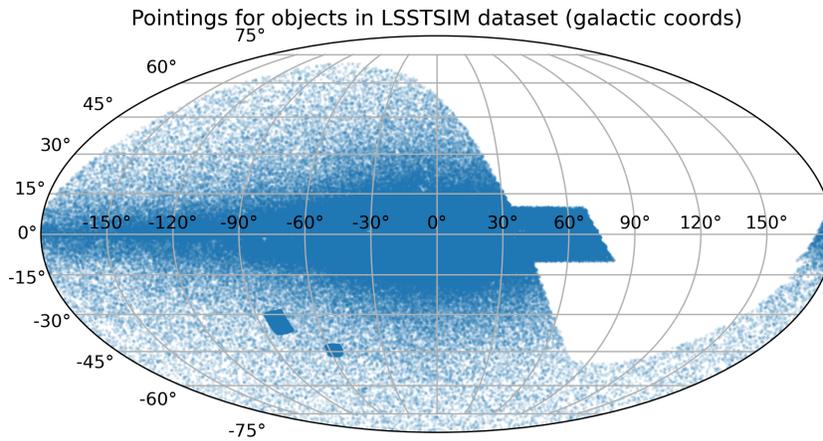


Figure 4.3: Spatial distribution of a random subset of LSST SIM sources in galactic coordinates. In contrast to the NSC data, there will be many galactic sources sampled in the densest regions. The small patches in the lower left are the Large and Small Magellanic Clouds.

4.3 Background Characterization for LSST

In section 3, I predict the amount of detectable microlensing events that would result from PBH dark matter. LSST will contain lightcurves from many billions of objects over 10 years, and even with 100% PBH dark matter, I predict a maximum of $\sim 10^6$ events. This means we’re looking for needles in a haystack. This is fine, but requires some understanding of how much hay can be confused for needles. In other words, how many lightcurves have transient events that might look like microlensing in LSST data? There are several known classes of objects that would exhibit this behavior, and I enumerate them in this section.

Stellar Origin Black Holes (SOBHs) microlensing

We have well-established models for the production of black holes as stellar remnants [81–83]. These allow us to predict the numbers of astrophysical black holes arising from a stellar population, based on age, metallicity, size, and thermodynamic factors. There will certainly be microlensing events due to these lenses during LSST, as we’ve already seen some [35,42]. These are “backgrounds” in the sense they are a known signal that we expect to find in this search. However, they are perfectly valid contributions to the present-day dark matter.

They are distinguished from PBHs in a few key ways: firstly, the Chandrasekhar limit [81] puts a lower bound on the mass of these objects, famously

near $1.4M_{\odot}$, whereas PBHs can, in principle, take any mass, since they aren't due to the collapse of a star.

Secondly, SOBH's should have different kinematics to those of PBHs. Whereas one would expect PBHs to be virialized, having existed since near the big bang, black holes produced in supernova appear to experience a "natal kick" of $\sigma > 260$ km/s [84]. Microlensing events involving these lenses would happen with somewhat faster crossing times.

Though production is well-understood, real abundances are not. In this work, I follow the results of [85] and assume the Milky Way contains $\sim 10^8$ SOBH's. I further assume they follow the mass spectrum reported in [86]. Both of these assumptions are poorly constrained at present.

Binary stars microlensing

Binary stars could be a common source of microlensing signals. According to [87, 88], "about half of all solar-type stars are members of binary systems". They have a wide spread of periods for their orbits, from a few days to 10^5 days. There have been catalogs generated in the above references which could help to calculate the number of expected alignments causing microlensing events.

In contrast to isolated SOBHs, which would be indistinguishable from a PBH of the same mass, lightcurves from binary star systems would have two effects to

distinguish them: occlusion when the stars align, and blending of their flux, since both objects are sources. Both of these imprint on the lightcurve. Additionally, the lightcurve will be periodic with the orbit of the system.

In the case of a source paired with a stellar remnant, like a black hole, neutron star, or white dwarf, it would be harder to distinguish, as the remnant may not be bright enough compared to the source star to exhibit blending, and a remnant could be small enough that the occlusion can't be detected. This system would still be periodic, and given enough time, these could be distinguished.

Variable stars

These are mostly chromatic events - as long as we have measurements in a few bands, and especially if we include followup observations from other instruments, it should be straightforward to discriminate these from microlensing events. The MACHO project describes the cuts required to rule out lightcurves that initially look like microlensing [89] . They report huge cuts in their dataset, from 17,440 sources with a reasonable lightcurve fit, down to 3 final candidates. Some of their cuts are based on statistics on the lightcurve data, while others are explicit cuts based on known nebulas with greater transient activity.

The two most relevant categories for microlensing backgrounds are eruptive/-cataclysmic variables and pulsating variables.

Eruptive/cataclysmic variables

These include several types of novae, plus symbiotic stars. For example, dwarf novae become elevated in brightness by 2 to 5 magnitudes over the course of a few days [90]. This is the right timescale and magnitude amplification to get confused with a microlensing event. For example, see Figure 1 of [91] for the lightcurve of a nova which looks a lot like microlensing if one only has a measurement in 1455 Å. Looking at multiple color bands shows that this event is chromatic and is not due to microlensing.

Pulsating variables

Pulsating variables can change their color both via surface temperature change, and due to Doppler shift as the surface expands and contracts [92]. Given measurements in different bands with sufficient S/N, this could be used to discriminate such sources from microlensing events.

4.4 A new microlensing discovery algorithm:

Weighted Overlapping Outlier Detection

The algorithm described in this section uses photometry measurements in any number of filters (including 1) to find microlensing events. It leverages the achro-

maticity of gravitational lensing to discriminate these events from the astrophysical variations of stars, which tend to change their color profile. This is implemented by considering only the change in baseline magnitude within each band.

The first step in looking for microlensing events is to find where there is a statistically significant positive deviation from the typical brightness of a source. One wants to find a contiguous period where the source’s brightness is measured with the following properties:

1. Brighter than normal
2. Achromatic: consistent brightening in all bands

It turns out these two properties alone are very powerful in removing astrophysical variations and preserving microlensing events, given a thoughtful implementation. Both of these descriptions require some interpretation. For “brighter than normal,” my condition is implemented as an effective z-score on a moving average of the deviations from baseline in each filter. The numerator of the z-score is the weighted moving average (WMA) with a tunable timescale, 2 days by default. The denominator is the sum in quadrature of the confidence on the WMA, the clipped standard deviation of variations from baseline for the object, and a metric I’ve called the *Weighted Moving Scatter* (WMS), visualized in Figure 4.4. The latter measures the proximity in magnitude of nearby measurements with the same envelope as the WMA. It can be thought of as a second moment of residuals

on the WMA, again weighted by the same kernel. It is designed to modulate the significance of bright deviations by the similarity between nearby measurements. This serves as an implicit check of achromaticity - if measurements in different filters diverge from each other, they will produce a larger WMS, and a bright excursion will require a larger deviation to trigger. It also reduces variations that happen much faster than the timescales of interest. For example, a lightcurve with a measurement close to baseline, quickly followed by a bright measurement, will have high WMS, and is less likely to trigger the threshold.

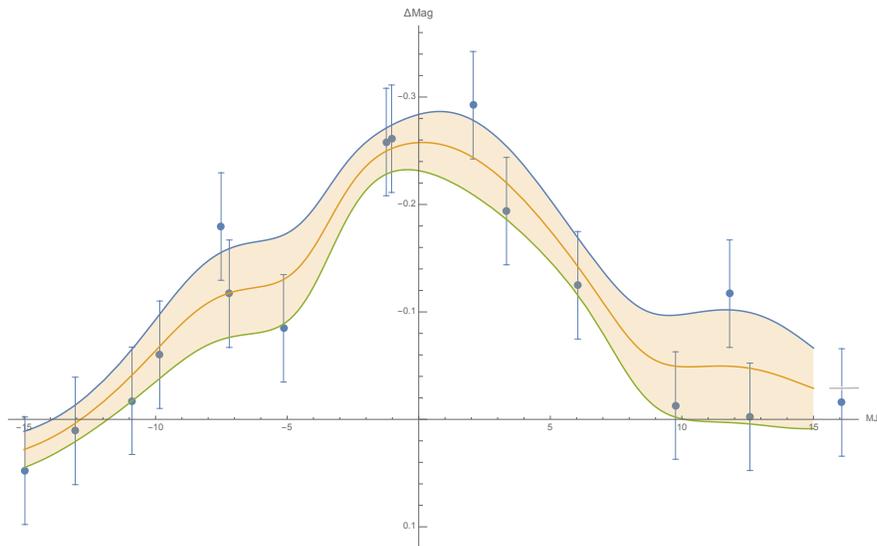


Figure 4.4: The Weighted Moving Average and Weighted Moving Scatter computed continuously for synthetic data. WMS is visualized as an envelope around the WMA. This continuous version is not computed during the analysis, but is a generalization of the discrete formula described below, found by substituting $\kappa_{ij} \rightarrow \kappa(t - t_j)$, giving functions $A(t)$ and $S(t)$. The discrete version is trivially equivalent to evaluating the functions at the same times as the datapoints, t_i . One must avoid interpreting the envelope as a confidence interval on the moving average. It is only a measure of the consistency between nearby measurements.

4.4.1 Algorithm specification

1. Round 1: detection. Look for persistent excursions in the average level across all filters.

(a) Calculate `deltamag` (δ_i) for each measurement, by subtracting the NSC calculated object baseline in the respective band from each measurement³.

(b) Calculate the outlier-excluded standard deviation `std` for the object.

Exclude outliers by:

i. Compute the median $\bar{\delta}$.

ii. Compute median of the absolute deviation of each point from the sample median: $\bar{\Delta} = \text{Med}(\text{Abs}(\vec{\delta} - \bar{\delta}))$.

iii. Exclude measurements deviating by a factor of this median deviation:

$$|\vec{\delta} - \bar{\delta}| < m_d \times \bar{\Delta}, \text{ with } m_d = 3 \text{ by default.}$$

iv. Calculate `std` = $\bar{\sigma}$ as the standard deviation of the remaining δ_i .

(c) Calculate the weighted moving average (WMA) of changes in baseline

³An outlier-excluded estimate of the baseline should be even more accurate for Rubin, given the expected number of measurements per source.

as

$$A_i(\vec{\delta}, \vec{\epsilon}) = \sum_j \kappa_{ij} w_j \delta_j \times \left(\sum_k \kappa_{ik} w_k \right)^{-1} \quad (4.4)$$

$$\text{with } \kappa(\Delta t) \equiv \exp\left(-(\Delta t/\tau)^2/2\right) \quad (4.5)$$

$$\kappa_{ij} \equiv \kappa(t_i - t_j) \quad (4.6)$$

$$w_j \equiv \left(\epsilon_j^2 + m_w \bar{\Delta}^2\right)^{-1} \quad (4.7)$$

where δ_i is the `deltamag` deviation calculated above in mags, ϵ_i is its measurement error in mags, t_i is its time of measurement in MJD, and $\tau = \text{timescale}$ is the smoothing time⁴. Notice the weights w_j are not simply the inverse squared errors, but are “tempered” by the clipped deviation $\bar{\Delta}$ with a factor m_w (parameter name `temper_errors`) set to 1 by default. This prevents low-error outlier measurements, whether underestimated or genuine, from pulling these metrics much more strongly than other measurements. These adjusted weights are used in place of the measurement errors throughout the algorithm.

(d) Calculate the confidence on the above WMA:

$$\sigma_i(\vec{\delta}, \vec{\epsilon}) = \left(\sum_j \kappa_{ij}^2 w_j \right)^{1/2} \times \left(\sum_k \kappa_{ik} w_k \right)^{-1} \quad (4.8)$$

⁴In practice, the windowing kernel function κ is clipped as described in A.1, because it is expensive to compute over all N^2 pairs of measurements for each object, when most of them evaluate effectively to 0.

(e) Calculate the *Weighted Moving Scatter* (WMS):

$$S_i(\vec{\delta}, \vec{\epsilon}) = \left(\sum_j \kappa_{ij} w_j (\delta_j - A_j)^2 \right)^{1/2} \times \left(\sum_k \kappa_{ik} w_k \right)^{-1/2} \quad (4.9)$$

This metric quantifies the variation in temporally nearby measurements.

(f) Calculate an effective z-score for each measurement as $Z_i = A_i / \sqrt{\sigma_i^2 + S_i^2 + \bar{\sigma}^2}$.

Flag “excursions” as measurements where $Z_i > \mathbf{z_threshold}$.

(g) Look for consecutive excursion points (*regions*) with the following properties:

- i. Region contains at least `n_measured=4` measurements.
- ii. First and last measurement are separated by at least `duration = 5` days.

2. Round 2: selection. Quality of fit to the ideal microlensing curve. Iterate over all excursion regions found in Round 1:

- (a) Extend the region to include all datapoints within `±context_size` of the original boundaries. By default, `context_size = 100` days.
- (b) Fit the ideal point source-point lens (PSPL) microlensing curve to the data in the extended region, yielding fit parameters `impact_parameter`, `crossing_time`, and `peak_time`, and a condition number of the covariance matrix. If the fitting procedure doesn’t converge, reject the event.

If two or more flagged regions converge to the same curve, consolidate them.

- (c) Compute the residuals of the fit throughout the extended region. Check whether the residuals are statistically consistent with the lightcurve outside the extended excursion region:
 - i. If there are at least `n_min_outside_fit` measurements for the object outside the fit region, perform a weighted two-sample Kolmogorov-Smirnov (KS) test between the residuals on the fit and the data-points outside the fit region, yielding a p-value p for each excursion.
 - ii. Otherwise, perform a weighted one-sample KS test on the residuals with a Gaussian with the same mean and standard deviation as the residuals, i.e. check that the residuals are Gaussian distributed.
- (d) Filter out events with condition number of the covariance matrix greater than `condnum` (default 10^5), optionally combined with the calculated p-values of residuals to the fit.

4.4.2 Table of default parameters

Parameter	Default value	Description
<code>outliers_cutoff</code>	3	Scale to remove outliers in <code>stdev</code> calculation
<code>cut_outliers</code>	False	Option to remove outliers entirely from consideration in excursions
<code>outliers_cutoff_data</code>	20	Scale to remove outliers with low measurement errors
<code>z_threshold</code>	3	Effective z-score threshold to flag excursion points
<code>timescale</code>	5 (days)	Timescale of smoothing for WMA and WMS
<code>n_measured</code>	4	Minimum number of measurements to flag an excursion
<code>duration</code>	5 (days)	Minimum duration to flag an excursion
<code>restrict_to_indices</code>	[Well-Sampled regions]	Only consider a subset of datapoints for excursions - used to exclusively parse regions with fast enough cadence in NSC data, while still using all the data for characterizing baseline
<code>usescatter</code>	True	Use Weighted Moving Scatter to modulate detection significance
<code>temper_errors</code>	1	Relative scale between clipped <code>stdev</code> and measurement errors in contributing to data weights
<code>context_size</code>	100 (days)	Number of days to extend the PSPL fit data on either side of the region detected in Round 1
<code>condnum</code>	10^5	Cut on condition number of the covariance matrix from the PSPL curve fit

Subset	# of objects
All NSC sources	3.9×10^9
NSC cut 1 (star classifier, <code>n_detected</code>)	2.9×10^8
NSC cut 2 (well-sampled)	264,295
Round 1 Detections	33
Round 2 Selections	0

Table 4.1: Counts of objects and regions before and after NSC data selection, and two rounds of selections on real lightcurves with default parameters applied. The NSC cuts are described in Section 4.2.1.

4.4.3 Examples of detections and near detections, descriptions of parameters

Round 1 flags contiguous, positive excursions from the baseline. Figure 4.5 shows an example lightcurve with some bright measurements above baseline. On the right, the WMS is not considered in the denominator of the z-score. This is visualized as the blue envelope around the WMA. Without the WMS, the algorithm assesses a high confidence on the WMA, because the error bars are relatively small on the datapoints. However, this deviation is not consistent between bands. This fact is seen on the left, when adding the WMS to the envelope. After this addition, this event is no longer flagged in Round 1, and the background event is

successfully rejected.

In Figure 4.6, we see an example where a couple bright, low-error measurements pull up the moving average substantially and flag an excursion. The PSPL fit at first seems plausible, but the Y-band measurement near the inferred peak is many sigma away from the fit, as is the earlier g-band point. Combined with another high outlier ~ 30 days later, this event comes with a high condition number and is removed in the final cuts.

This example also justifies the `temper_errors` strategy, where I weight the moving average not just by the individual measurement errors, but also by the typical variation of the source as a clipped `stdev`. By adding the two in quadrature (with an assignable weight factor defaulting to 1), one avoids the issues caused by underestimated photometric errors. While experimenting with different strategies, I found that this technique is more effective than explicitly removing outliers from consideration, letting me keep the `cut_outliers` option off by default. Note that outliers are always removed when calculating `stdev`.

The cut on condition number of the covariance matrix, `condnum`, has the default value of 10^5 . This came from comparing the values of the Round 2 fits from synthetic events, and real lightcurves which were flagged in Round 1. As can be seen in Figure 4.9, the synthetic events (histogrammed values) and background events (large red dots) are well partitioned by a cut at `condnum` = 10^5 . The con-

dition number’s scale is set by the scale of the original variables, and with the synthetic events mostly appearing near and above 10^4 , this seems to correspond to the scale of the MJD times, via the fit of the peak event time. It is possible to remove this scale by centering the data before the fit, but since classification has been successful without adjustment, I leave this to future investigation. An alternative strategy using the covariance matrix is to make explicit cuts on uncertainties and covariances, though one must still deal with the parameter scales.

The default parameters are somewhat tuned for finding events typical of $M_{\text{PBH}} \sim M_{\odot}$, with $t_E \sim 40$ days, and they reflect a good response on the NSC dataset in particular. Used on a telescope with faster cadence, for example, this algorithm would want to use lower values for `timescale`, `duration` and `context_size` to catch shorter events. For Rubin, with comparable revisit times to the precursor data, I would expect the default parameters to be appropriate. When Rubin lightcurves become available, one must repeat this procedure with synthetic microlensing events to assess the efficacy of the parameters as was done here.

4.4.4 Extended rejection techniques

The two filtering stages described above succeed in eliminating all background events in the NSC data while detecting a majority of synthetic ones. Without trying to simulate more lightcurves, we’ve reached the end of the line for background

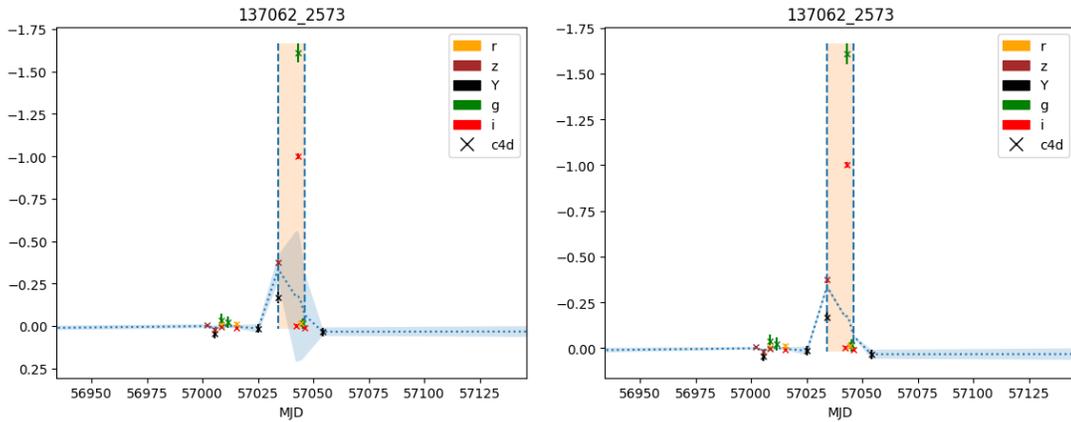


Figure 4.5: Loosely visualizing the effect of including (left) or not including (right) the weighted moving scatter as part of the confidence on the WMA. Without the WMS, high-confidence measurements in different bands can conflict to high sigma, and the WMA is reported at high confidence, even though the bands are very discrepant. This triggers a Round 1 detection when it could clearly be excluded due to high chromaticity.

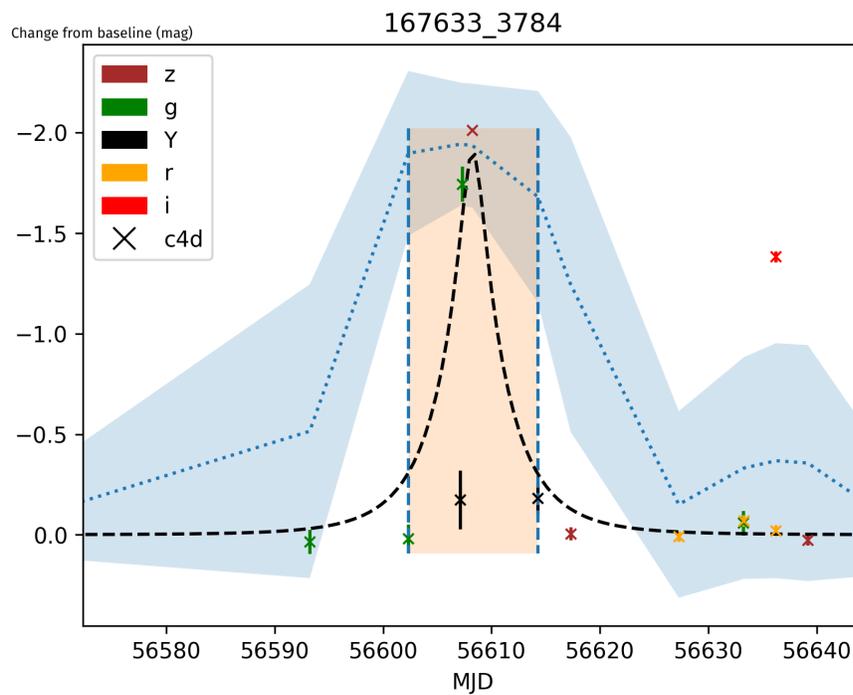


Figure 4.6: A background event, rejected for a poor fit (high condition number) to the PSPL curve in Round 2.

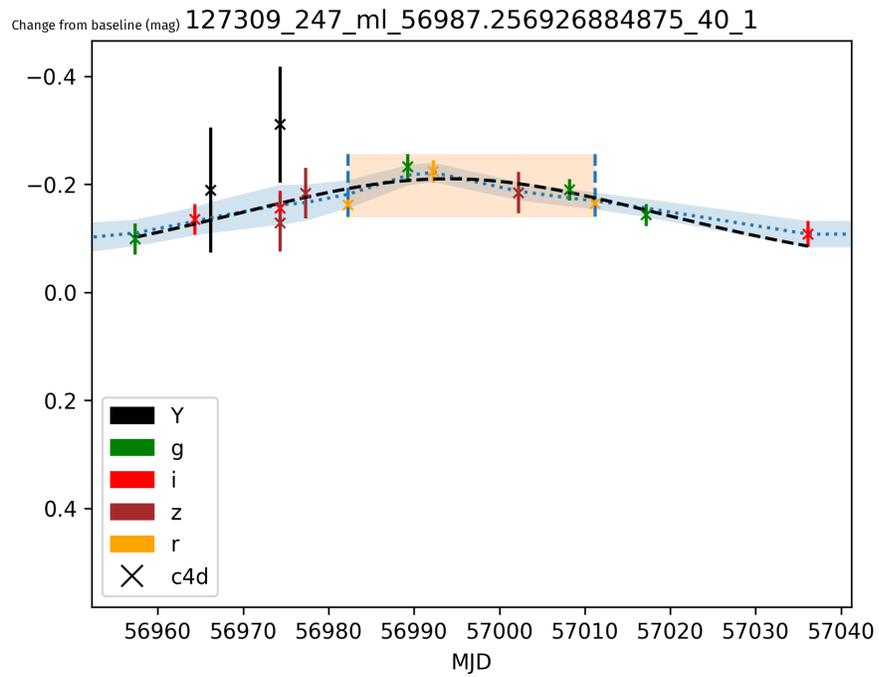


Figure 4.7: Even without particularly dense sampling, a good fit is inferred to this synthetic injected microlensing curve with $t_E = 40$ days and $u_{\min} = 1$.

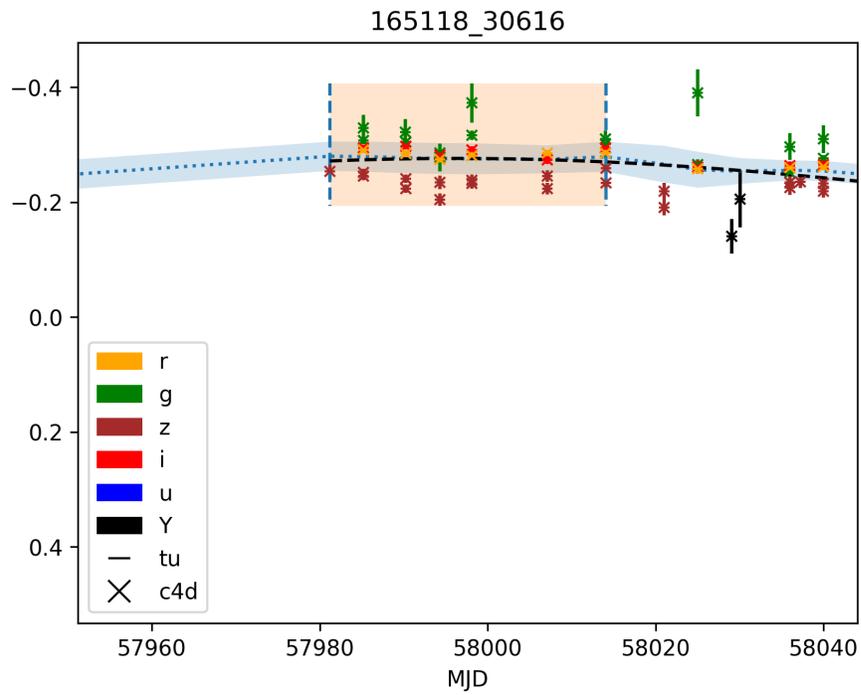


Figure 4.8: A lightcurve with a consistent bright excursion (away from 0), but not enough context to constrain the fitting procedure, resulting in a high condition number.

assessment in this data. However, there are more stages that could be used to extend this algorithm and improve the rejection power for the much higher volume of Rubin data.

4.4.4.1 More sophisticated cuts on fit parameters

I showed that a cut on condition number is sufficient to remove all backgrounds in the NSC data, and detect $\sim 50\%$ of synthetic events. This can be improved quite a lot by including more parameters in the cut. For example, Figure 4.9 achieves total background rejection with $\text{condnum} < 10^5$, but detection efficiency of only $\sim 50\%$ in the second stage. If I were to include a 2-dimensional cut involving the p-value of the KS test on residuals (y-axis), I could achieve a much higher efficiency, $> 90\%$. However, in this case with so few NSC background events in Round 2, this would feel like overfitting. With Rubin data, one will be able to use the same procedure with the much larger data volume to determine more robust cuts and achieve higher detection efficiency.

4.4.4.2 Moving baselines

Although stable sources are ideal for this type of search, every star is dynamic on some timescale. This algorithm makes use of fixed baseline estimates for each source, separately in each band. This could be replaced with a set of moving averages, with a much slower timescale than the one used for triggering excursions.

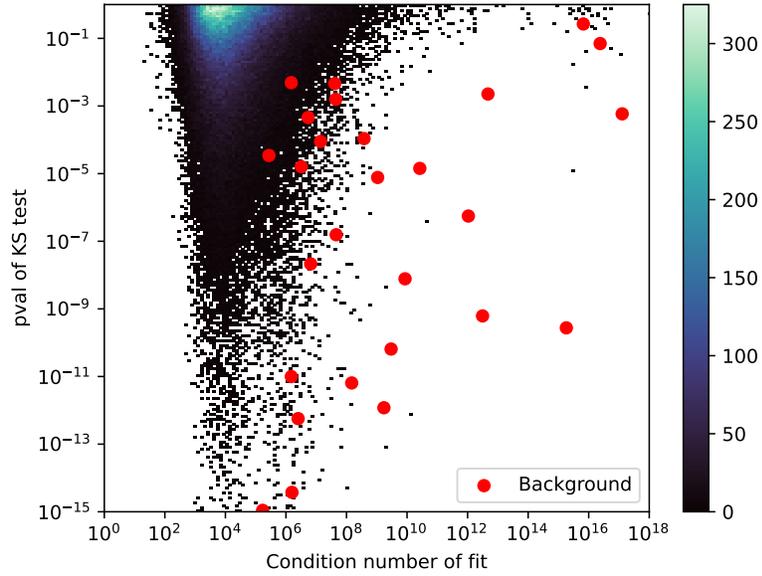


Figure 4.9: $1M_{\odot}$ synthetic events and real backgrounds plotted by the condition number and p-value of their fits to the PSPL lightcurve in Round 2. The color histogram shows the density of the large population of synthetic events. One could draw a curve delineating the two populations, while preserving $> 90\%$ of synthetic events, at the risk of overfitting and generalizing poorly. Note that this plot includes all events that pass Round 1, including those synthetic events with poorly inferred curve parameters.

This would improve detection efficiency on variable sources with long periods. A more involved system could flag variable sources, construct a model for their dynamics, and use that model to predict baselines. One could even jointly fit variable star models with the microlensing curves.

One reason to not implement this in the present search is that it introduces another hyperparameter which must be tuned depending on which timescale of events one looks for. Additionally, Round 1 is designed to be a fast algorithm for flagging excursions, and doing any intricate additional fitting would slow it down substantially. Round 2 is meant to contain the statistical heavy lifting after the domain has been reduced.

4.4.4.3 Repeat offenders

Microlensing events involving unbound objects are exceedingly rare. It would be extremely unlikely to see a source exhibit microlensing more than once. Other variability in stars, however, is commonly periodic. For this reason, many searches explicitly ignore events coming from sources that display multiple candidate events. My search has not implemented any such cuts, but they could easily be added.

4.4.4.4 Achromaticity

A key signature for microlensing in a multi-color survey is that light of all wavelengths is amplified the same amount. In other words, the lightcurves in all

color bands will move together. Round 1 of my algorithm makes *implicit* use of this fact when computing the WMS along the lightcurve, including measurements of every color. The “tightness” of the variation is used to modulate the threshold required for a detection, and a chromatic lightcurve would show large WMS.

An explicit cut could be implemented to further reject backgrounds. For example, one could require the residuals on the microlensing curve fit to be consistent between bands, perhaps with a KS test. If one band is pulling the average more than others, a test like this would show the difference.

4.4.4.5 Symmetry

Assuming linear trajectories for a source and lens, the microlensing amplification is symmetric in time when reflected across the peak (Eqn. 4.10). This is broken by the parallax from Earth’s orbit, but for sufficiently short timescales or large PBH masses, it could be used to discriminate events.

4.4.4.6 Dynamic context

The Round 2 fitting adds context to the detected excursion region, to give more information to the curve fitter, and ensure the lightcurve returns to baseline on the right timescale. These examples have hard-coded 100 days of context (if available) on either side of the excursions. This would likely be improved with a dynamic selection. For example, in Figure 4.10, which shows the closest background event

to acceptance, the fitting procedure did not include the datapoints shown on the rightmost edge. These points clearly don't return to baseline on the right timescale for a good fit, and their inclusion would have pushed both condition number and p-value further away from acceptance.

4.5 Detection efficiency with synthetic microlensing events

As described in Section 4.2, I evaluate the detection efficiency of the algorithm by superimposing the real data with ideal microlensing curves. The parameters of these curves can be specified arbitrarily, but to assess the efficacy for Rubin, I randomly sample the events generated with the methods in Section 3.2, and add the corresponding curve to a copy of the real lightcurve data. Explicitly, given original lightcurve magnitudes y_i taken at time t_i , I compute the simulated lightcurve data \tilde{y}_i as:

$$\tilde{y}_i = y_i - 2.5 \log_{10}(A(u(t_i))) \quad (4.10)$$

$$A(u) = \frac{u^2 + 2}{u\sqrt{u^2 + 4}} \quad (4.11)$$

$$u(t) = \sqrt{u_{\min}^2 + \left(\frac{t - t_0}{t_E}\right)^2} \quad (4.12)$$

For each Well-Sampled region in the dataset, potentially several per source, I make a copy of the entire original lightcurve, and inject one random event into

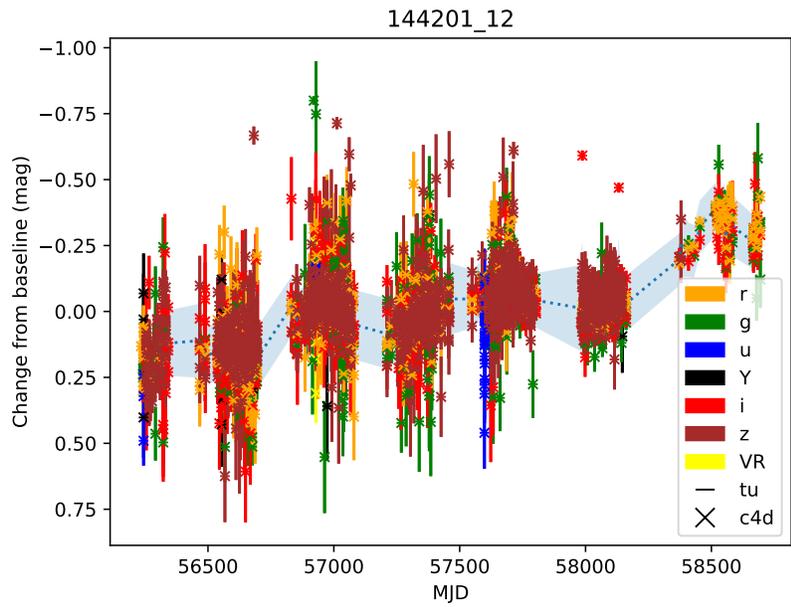
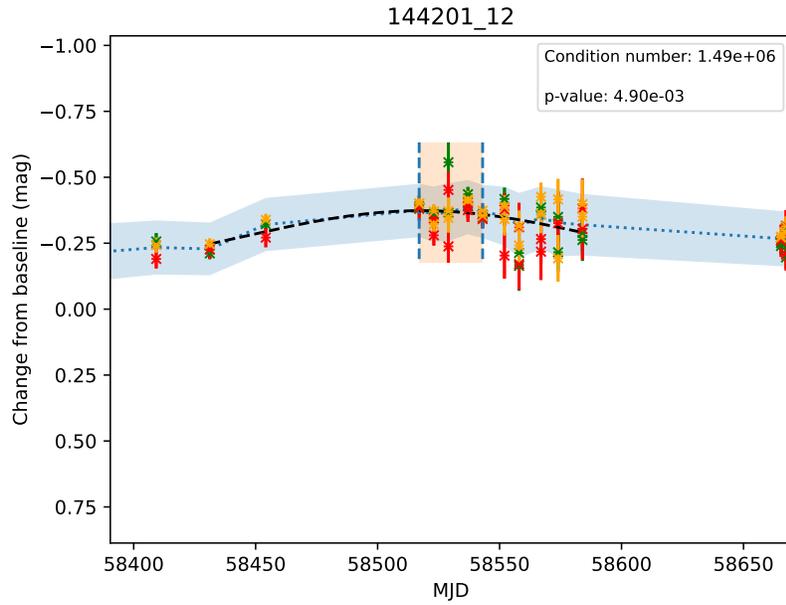


Figure 4.10: The closest background to acceptance, seen as the furthest top-left red point in Figure 4.9. Just the excursion is shown on top, and the full set of datapoints below.

the center of the WS region. While using this data as a precursor to Rubin, this is without loss of generality - every lightcurve coming from Rubin should be Well-Sampled for most of the survey.

4.5.1 Ensuring consistency between injected and detected events

In comparing performance of various options and parameters for the search algorithm, one must compare the background detection rate vs. the synthetic detection rate. For the backgrounds, it's as simple as counting every event that is flagged as microlensing. For the synthetic events, it's a little more complicated. It's not enough to trigger an event, though that is a good preliminary indicator that something is working if there isn't a corresponding event in the unmodified lightcurve. Really, one wants to know that the detected event infers the same characteristics that were injected. Since the detection of an event at datapoint i is dependent not just on the value of that measurement, but also of those in its vicinity, it's possible for the injection of an event to cause a nearby borderline detected event to peek over the threshold and become a detection. Additionally, by comparing the injected parameters and those inferred during the curve fitting procedure in Round 2, one can quantify the confidence on the inferred microlensing parameters.

To accomplish these comparisons, I save all the “ground-truth” information while generating the synthetic lensing events, within the function `nsctools.generate_synthetic_microlensing_events_from_population`. In addition to generating new copies of existing lightcurves with synthetic events superimposed, I save a `*-info.pickle` file which contains a tuple. The first element is a dictionary with meta-information about the original lightcurve files, event population to draw from, and output location. The second element is a `pandas.DataFrame` containing columns for the original `objectid`, the new synthetic `objectid`, index of the superimposed event within the sampled event `DataFrame`, and all the microlensing parameters of the event. With this information I can unambiguously reference the injected event’s information, given the synthetic object id.

Now, I can feed the synthetic lightcurves through the detection algorithm and compare the inferred parameters to the real parameters. Since every detection from Round 1 is fit during Round 2, one can look at distributions of the fit results even before making the final cuts. For example, in Figure 4.11, we see all the Round 1 detections for a sample of $1M_{\odot}$ PBH events. Closest to the correctly inferred peak time (0 days difference) is the largest population of events by 2.5 orders of magnitude. During Round 2, I make a cut on condition number, shown here as color in log scale. Notice the well-inferred bin has median condition number

over 2 orders of magnitude lower than adjacent bins, and over 10 orders lower than the wings. This plot demonstrates that the cut successfully removes detections with poorly-inferred peak times.

After the Round 2 cuts, one can assess performance of the entire procedure. Figure 4.12 shows the distance between real and inferred peak times as a fraction of their real crossing times. We can see the distances are tightly peaked near 0, with 95% of event peaks inferred within 9.5% of their crossing times from the true peaks. While the peaks are useful for verification, in practice the most important parameter to measure is the crossing time. This is the main dimension along which one can discriminate between different PBH populations. Figure 4.13 shows the distribution of inference performance, with 95% of events measured within 12.9% of the correct crossing time, and 99% measured within 27.7%.

4.5.2 Exclusions for different masses of PBH

Exclusions are calculated by considering monochromatic PBH models comprising a fraction f of the dark matter profile for the Milky Way. For a model with PBH mass M , I predict a total 10-year rate λ_0 for $f = 1$, giving the expected number of events for 100% PBH dark matter. This includes the detection efficiency estimates from the `MicroLensingMetric` (Section 3.2.4.2). I assume the number of real events detected follows a Poisson distribution with rate $\lambda = f\lambda_0$.

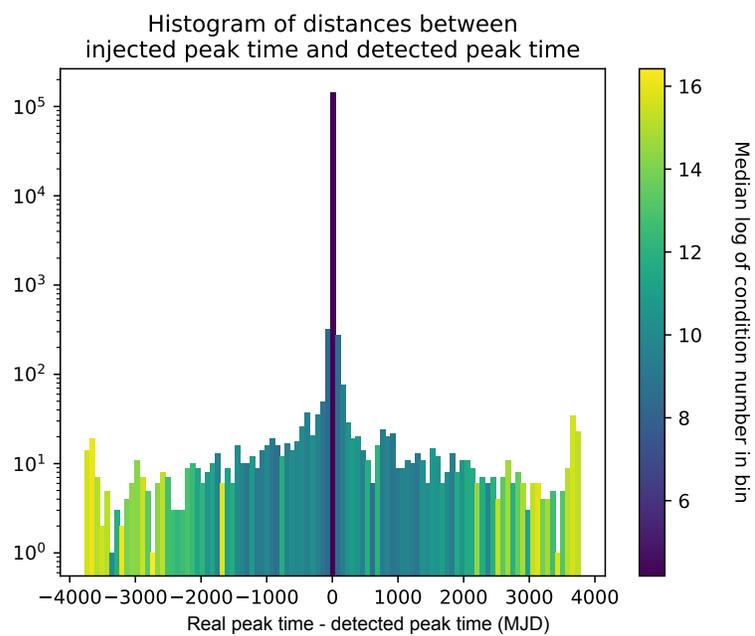


Figure 4.11: Before performing Round 2 cuts on fit quality, we can see a strong correlation between distance of the inferred peak time from actual peak time with the condition number on the fitting procedure.

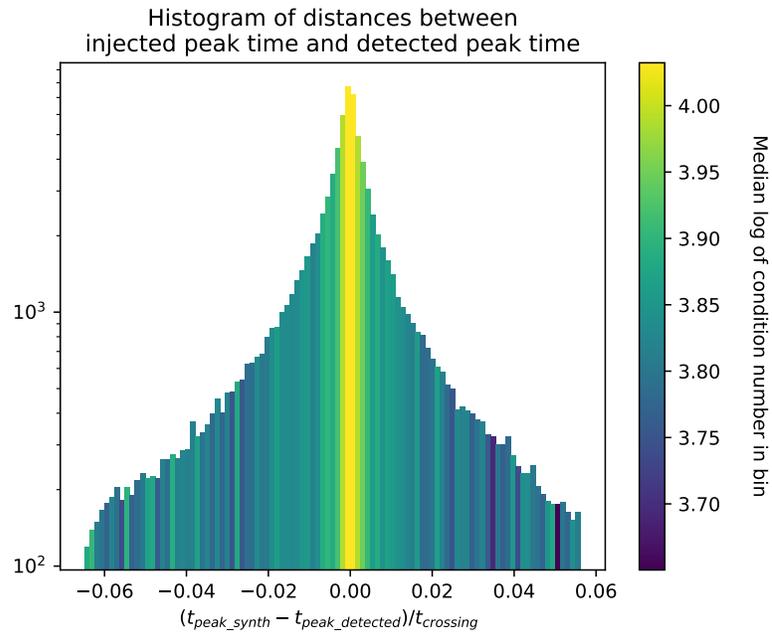


Figure 4.12: Validation of inferred microlensing peak time versus injected peak time after the Round 2 cut on condition number. Population standard deviation is a distance $\sim 5\%$ of the crossing time, FWHM of 3.6%.

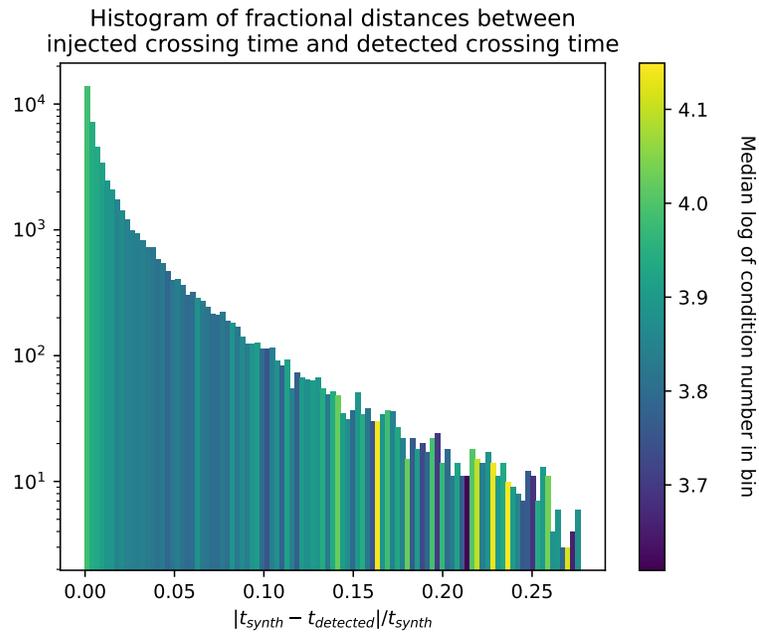


Figure 4.13: Validation of inferred microlensing crossing time versus injected crossing time after the Round 2 cut on condition number.

Given a number of detections, one can calculate the 95% confidence upper bounds on the value of f .

Since the NSC data only has a small fraction of the power of the full LSST, the best one can do is provide an upper bound on the background signal. Fortunately, we have reached nearly the best-case scenario, with a set of reasonable parameters that detects 0 events in the real data, and detects a meaningful fraction of the synthetic events. The only room for improvement at this stage would be to increase the detection efficiency while keeping the background detections at 0. Potential strategies for this are described in 4.4.4.

To make the exclusions in Figure 4.14, I take the upper bound assuming a rate of one background event in the NSC data. To extrapolate to the number of events in LSST, I divide the volume of $\text{time} \times \text{sources}$ for Well-Sampled objects, by the number of LSST SIM sources $\times 10$ years. This gives us a ratio between the amount of WS observations of galactic sources in the two datasets. The upper bound, with 1 background event in the NSC data, becomes 819,236 events as an upper bound in Rubin. I assume these events are log-evenly distributed in crossing time, and for each monochromatic PBH mass population, I consider backgrounds only relevant if their crossing time falls within the [5,95] percentile range of the simulated events. The same percentile clipping is applied to the simulated microlensing rates for each population.

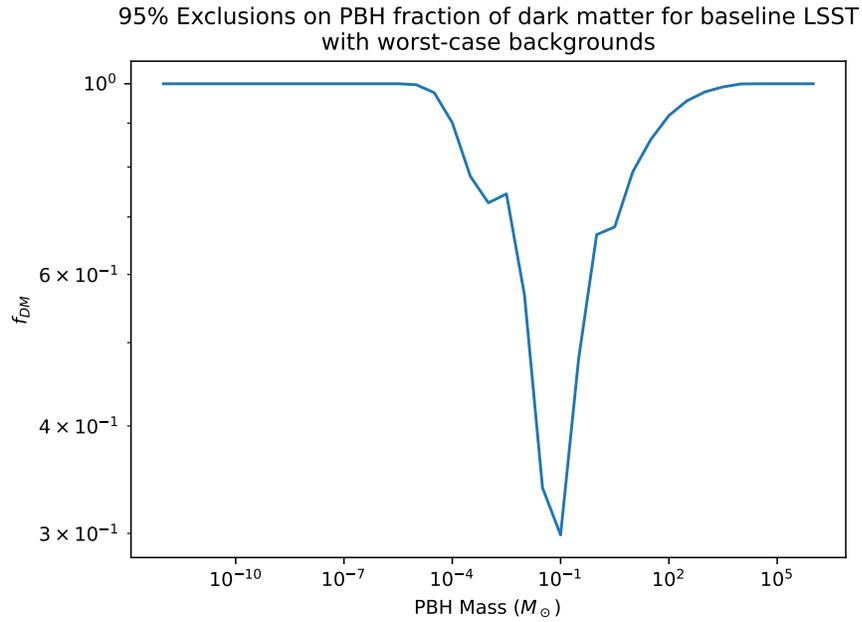


Figure 4.14: 95% confidence exclusions for LSST, given a number of detections corresponding to the upper bound of a background rate in NSC data, assuming they are log-evenly distributed in crossing time. There’s no reason to expect backgrounds this high, but it’s impossible to know what will be found in the novel LSST dataset. But even in the worst case, there will be some exclusion power in the slightly sub-solar mass range.

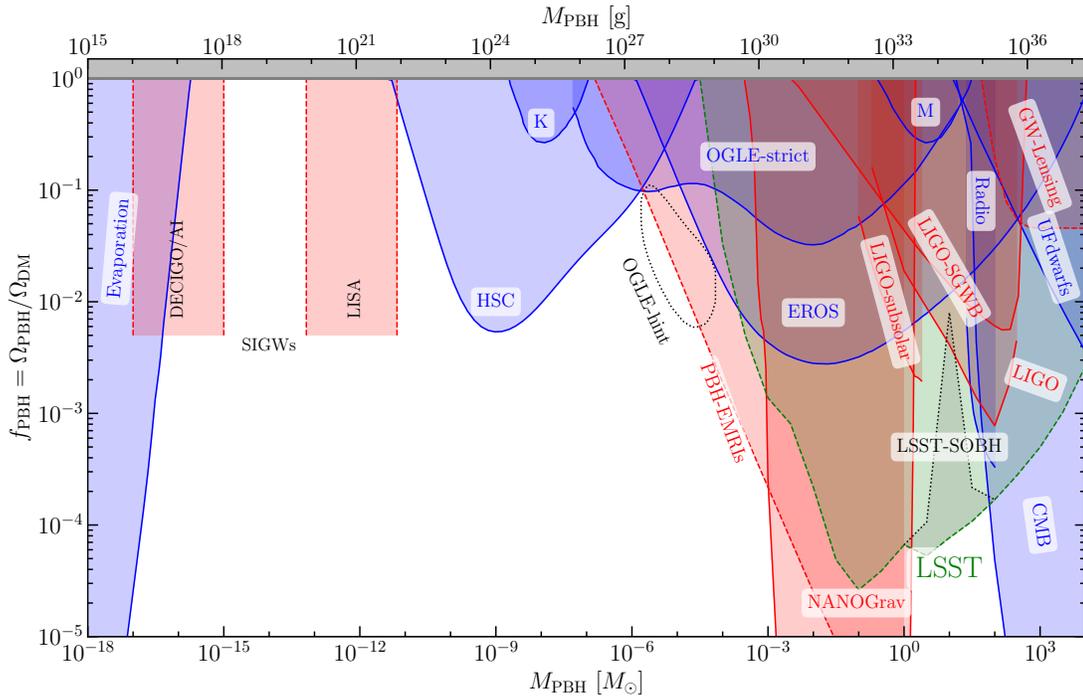


Figure 4.15: 95% confidence exclusions for LSST, given 0 detections over ten years, shown in green. Exclusions are superimposed on a collection of existing bounds from [1]. Despite the strongest exclusions coming from $\sim 10^{-1} M_{\odot}$, the strongest “new” exclusions are in the $1 - 100 M_{\odot}$ range. In any case, LSST provides independent coverage to areas which previously were only excluded by gravitational wave measurements. The dashed region around $10 M_{\odot}$ is due to the expected background of Stellar Origin Black Holes, which is described in Section 4.3. Although it will be difficult, if not impossible, to distinguish these from PBH’s of the same size with these methods, Rubin should provide novel insights into this currently poorly constrained population.

Chapter 5

Results and Discussion

5.1 Anticipated benefits and challenges with Rubin data

As discussed in Section 1.2.3, the Difference Imaging Analysis pipeline will enable nearly real-time reporting of transient phenomena in billions of sources across the sky. If an object makes a meaningful change from baseline, we will know about it quickly. Much of the power of a massive, centralized, accessible survey like LSST is not just contained in its own measurements, but in its ability to direct scientists where to look on short notice. If a significant deviation is detected on a stable star, for example, an alert will be generated, and followup measurements can be taken with another instrument. Without the large coverage

and fast revisit times of Rubin, a small telescope has little chance to catch a microlensing event. But by cutting down the domain of sources, we start to have a chance. Other instruments could provide rapid exposures in multiple colors, or even spectroscopy, to much more accurately characterize events as astrophysical variations or microlensing.

As with any ground-based telescope, the atmosphere will diffuse unresolved sources like most stars, changing them from points to a shape characterized by a Point Spread Function (PSF). The atmospheric PSF will change according to weather conditions, for example, and will have to be approximated in each image [31]. This diffusion of point sources makes so-called crowded-field photometry a difficult task, as flux from nearby sources will blend together in ways that can potentially be impossible to untangle with one image. Difference imaging has been used successfully to find microlensing events in crowded fields, for example the Subaru search of M31 [41]. One must be careful of blending effects while looking for microlensing events in crowded fields, since generally only one of the blended objects will be magnified. This means the change in *blended* flux will not necessarily be achromatic, and the PSPL curve will have a somewhat different shape [79]. A careful analysis of these effects should be undertaken with real Rubin crowded-field images and calibrated exposures when they become available.

5.2 Limitations of the NSC precursor data

5.2.1 Pointings

As discussed in Section 4.2.2, the preference against the galactic bulge means there was little chance of finding a genuine microlensing event in this NSC dataset. This is good in the sense that I get to assess lightcurves that are almost certainly 100% background events to real microlensing signals. However, it does come with the caveat that I assume the types of variations seen outside the bulge are comparable to the distribution within the denser pointings. The LSST should provide new insights into the populations of variable sources in different regions of the Milky Way, and will help us understand how good an assumption this is.

5.2.2 Timescales

The two principal limitations of the NSC data in this study are (a) the total volume of Well-Sampled data, which is $\sim 10^5$ times less than anticipated for LSST, and (b) the length of each of these regions which rarely exceeds 60 days.

The implications of (a) are simply that there are not enough sources observed over long enough time to expect to find any events, given the predicted numbers from Chapter 3. This affects events with any crossing time, and therefore any size PBH population. In any case, I was able to set an upper bound on the back-

grounds, and show that LSST will have some constraining power, even assuming the background events reach the upper bound set by the precursor data. This is shown in Figure 4.14.

As for (b), this is the primary distinguishing capability of LSST. Astronomers have never measured such a large number of sources so frequently for so long with the same instrument. With the NSC data, though there are some sources measured frequently for years, there is a trade-off with the number of objects that get this treatment. In contrast, the LSST cadence is designed to achieve fast revisits for most of the available sky for most of the survey. These long duration observations will be necessary to characterize events associated with larger PBH masses, since they have longer crossing times. Without covering a large portion of the lensing event time-frame, it would be difficult to properly infer crossing times, or even confirm that the event is from lensing instead of another astrophysical event.

It's worth noting that there could be more space for discovery in the full NSC dataset by modifying the Well-Sampled conditions described in Section 4.2.1. For example, for events with crossing times of a hundred days or more, a revisit condition of 5 days is far too strict. There would be a different subset of the NSC lightcurves that is ideal to study longer timescales, although I would also require a longer timespan for the observations. On the other side, for few-hour-long events,

there could be more data volume in the subset with a revisit timescale of an hour, and a timespan of several hours. The WS conditions in this study were designed to match the expected revisit times of LSST, and to have a long enough span of measurements to detect the median timescale of events from populations of PBH's near $1M_{\odot}$.

5.3 Rubin's sensitivity through the years

Figure 5.1 shows the expected 95% confidence exclusions of various PBH masses over the course of the survey, assuming zero microlensing detections in the corresponding range of crossing times for each bin. Efficiencies were generated with the `MicrolensingMetric` in `rubinsim` (Section 3.2.4.2), with a detection condition requiring at least 10 nights of measurement of the object, at least 10 total measurements during the microlensing event, and a characterization of the crossing time within 5% according to the Fisher matrix. This plot will be modified once real Rubin lightcurves are available to assess detection efficiency with synthetic events.

We can start to see exclusions around 100 days, at which point we quickly get to leading levels of upper bounds on the PBH fraction of dark matter in the near-solar mass range. By the end of the 10-year survey, we reach the green curve in Figure 4.15.

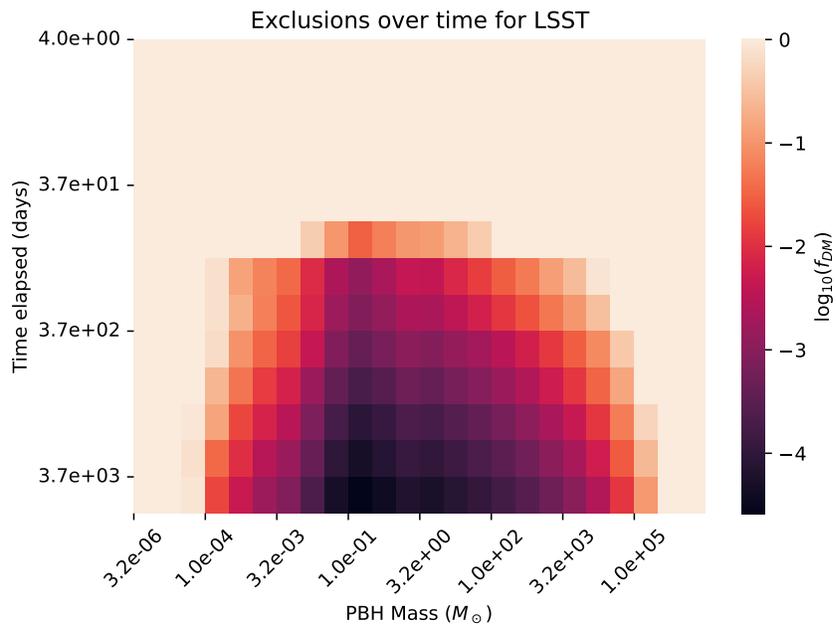


Figure 5.1: PBH exclusions over the course of the LSST, with time on a log scale and mass in units of solar masses.

5.4 Future work

5.4.1 Finding Stellar Origin Black Holes

Exactly the same methods described in chapter 4 are applicable to finding SOBHs as well as PBHs. As discussed in 4.3, I expect to find a large number of these stellar remnants with Rubin, but the total number in the Milky Way is currently not well bounded. In this work I have assumed there are 10^8 total SOBHs, following an exponential distribution starting at $6.3M_{\odot}$, from [86]. If we make the simplification that the SOBHs have the same distribution as the modeled PBHs, one would expect to detect around 400 microlensing events after the 10 years of LSST, in the range of 30-100 day crossing times. Fortunately there will be good detection efficiency in this range, catching $\sim 40\%$ of events with $u_{\min} < 5$. In the first data release, we should expect to tighten the constraints on total SOBH abundance in the Milky Way. It is beyond the scope of this work to make more concrete predictions, since precisely simulating SOBH crossing times would require modeling of the natal kicks, which will push the distribution toward faster events. PopSyCLE [76] has tools for this, but simulating a full-sky survey may prove computationally infeasible.

5.4.2 Search other datasets with the these methods

Prior microlensing searches have found convincing events, providing constraints described in Section 1.3.2. Each search uses a different set of methods. Another check of the efficacy of the method presented here would be searching the same surveys, and seeing whether one finds the same events, or perhaps new events.

One reason I selected the NSC dataset in this initial study is that it contains exposures in many colors. Whereas the existing microlensing datasets contain only one or two filters, working with the NSC data let me demonstrate the power of the implicit achromaticity tests in these methods. It was designed with the LSST in mind, and the multi-color data availability is a distinguishing element. It is not yet clear how it will fare on data with only one or two filters.

In any case, the version implemented in `nscml` is designed for the NSC data format employed by NOIRLab, and will require an interface to apply it to other datasets, including Rubin. This is hopefully as simple as changing the column names. A more general version may be implemented in a new package.

It may also prove valuable to make a specialized version that can catch potential events early, and filter or prioritize sources for follow-up with other instruments. There will be many variable sources to observe, and microlensing discovery could be greatly improved with a decent ordering scheme. The NSC dataset, or early Rubin data, could be used to assess the performance of such a metric by

injecting synthetic microlensing events, and cutting off datapoints such that only the beginning of the curve is represented. This could show the discovery efficiency and background rejection as a function of the number of transient measurements, or of the time measured before the peak of the event.

Appendix A

Search algorithm implementation notes

The full code is available in the `nsctools` package. Here, I will note some of the challenges and surprises made during its development.

A.1 Windowing for WMA and WMS

For my initial implementation, I used the very naive N^2 windowing tactic of computing temporal weights as a matrix of values between every pair of datapoints for each object:

```
windows = window_fn(t.reshape(-1,1)-t.reshape(1,-1), timescale)
```

where `t` is a `numpy` array of the MJD times of each measurement, and by default `window_fn` is a Gaussian centered at 0.

This method seems very wasteful, and indeed was a pretty expensive part of the code. The timescales are much smaller than the span of time measured, and most of this matrix contains values close to zero. Exploring the available options for the parameters in the algorithm, and running on many synthetic lightcurves for different PBH populations, makes it valuable to spend time optimizing such a bottleneck. I first tried to leverage the fact that most of these values are zero by designing and implementing a more clever algorithm to construct a *sparse* matrix which let me only store values above a certain threshold - something like 10^{-6} , for example, which would correspond to very conservative changes in weights considering the inherent scatter and errors in the data.

```
@njit
def sparse_gaussian_window_iter(t, timescale=2, nclip=10):
    rows = []
    cols = []
    vals = [np.float64(x) for x in range(0)]
    windowstart = 0
    for i, ti in enumerate(t):
        dt = ti - t[windowstart]
        while dt > timescale*nclip and dt >= 0:
            windowstart += 1
            dt = ti - t[windowstart]
            continue
        for j in range(windowstart, i):
            dt = ti - t[j]
            rows.append(i)
            rows.append(j)
            cols.append(j)
            cols.append(i)
```

```

        newval = np.exp(-((dt/timescale)**2)/2)
        vals += [newval]*2
    rows.append(i)
    cols.append(i)
    vals += [1]
    return (vals, (rows, cols))
def sparse_gaussian_window(t, timescale=2, nclip=10):
    sparse_matrix =
        scipy.sparse.csr_array(sparse_gaussian_window_iter(t,
            timescale, nclip), shape=(t.shape[0], t.shape[0]))

```

With Just-In-Time (JIT) compilation from numba and a reasonable clipping setting,

`sparse_gaussian_window_iter` is much faster than the naive implementation.

However, in a bit of a Catch-22, the initialization of the `scipy.sparse.csr_array` takes almost as much time as the JIT compiled naive code. The use of this `scipy` tool is incompatible with `numba`, and requires most of the WMA/WMS code to run without its optimization.

Instead of constructing a sparse matrix and using the same matrix products as before, I decided to calculate the entire WMA sum in one iteration.

```

@njit
def sparse_gaussian_wma(y, t, weights, timescale=2, nclip=10):
    wma = np.copy(weights*y)
    wme = weights.copy()
    windows_X_weights = weights.copy()
    windowstart = 0
    for i, ti in enumerate(t):
        dt = ti - t[windowstart]
        while dt > timescale*nclip and dt >= 0:
            windowstart += 1
            dt = ti - t[windowstart]
        continue

```

```

for j in range(windowstart, i):
    dt = ti - t[j]
    window = np.exp(-((dt/timescale)**2)/2)
    window_X_weight_i = window*weights[i]
    window_X_weight_j = window*weights[j]
    wma[i] += window_X_weight_j * y[j]
    wme[i] += window * window_X_weight_j
    windows_X_weights[i] += window_X_weight_j
    wma[j] += window_X_weight_i * y[i]
    wme[j] += window * window_X_weight_i
    windows_X_weights[j] += window_X_weight_i

wma = wma/windows_X_weights
wme = np.sqrt(wme)/windows_X_weights
return wma, wme, sparse_gaussian_wms(y, t, weights, wma,
    timescale=timescale, nclip=nclip)

```

`sparse_gaussian_wms` is constructed analogously, but it runs in a second iteration because it relies on the WMA which hasn't been computed yet during the first iteration. Fortunately, running a second iteration doesn't increase the complexity, and we've gone from $O(N^2)$ to $O(NM)$ where N is the number of datapoints and M is the typical number of datapoints in a clipped window. Generally, $M \ll N$. This has taken a typical computation of the WMA down from ~ 30 ms to ~ 1 ms, and the time to run Round 1 over the entire dataset from ~ 40 mins to ~ 4 mins. When searching the space of parameter options, this makes an enormous difference in agility.

Bibliography

- [1] Bradley J. Kavanagh. bradkav/pbhbounds: Release version, November 2019. <https://zenodo.org/record/3538999>.
- [2] Bernard Carr, Florian Kuhnel, and Marit Sandstad. Primordial black holes as dark matter. *Physical Review D*, 94(8):083504, October 2016. arXiv:1607.06077 [astro-ph].
- [3] LSST Collaboration. Lsst: from science drivers to reference design and anticipated data products. *arXiv:0805.2366 [astro-ph]*, May 2018. arXiv:0805.2366.
- [4] David L. Nidever, Arjun Dey, Katie Fasbender, Stéphanie Juneau, Aaron M. Meisner, Joseph Wishart, Adam Scott, Kyle Matt, Robert Nikutta, and Ragadeepika Pucha. Second data release of the all-sky NOIRLab Source Catalog. *The Astronomical Journal*, 161(4):192, mar 2021.

- [5] A. Einstein. Die grundlage der allgemeinen relativitätstheorie. *Annalen der Physik*, 354(7):769–822, 1916.
- [6] Planck Collaboration. Planck 2013 results. XVI. Cosmological parameters. *Astronomy & Astrophysics*, 571:A16, November 2014.
- [7] C. L. Bennett, M. Halpern, G. Hinshaw, N. Jarosik, A. Kogut, M. Limon, S. S. Meyer, L. Page, D. N. Spergel, G. S. Tucker, E. Wollack, E. L. Wright, C. Barnes, M. R. Greason, R. S. Hill, E. Komatsu, M. R. Nolta, N. Odegard, H. V. Peiris, L. Verde, and J. L. Weiland. First-Year *Wilkinson Microwave Anisotropy Probe* (*WMAP*) Observations: Preliminary Maps and Basic Results. *The Astrophysical Journal Supplement Series*, 148(1):1–27, September 2003.
- [8] P. de Bernardis, P. A. R. Ade, J. J. Bock, J. R. Bond, J. Borrill, A. Boscaleri, K. Coble, B. P. Crill, G. De Gasperis, G. De Troia, P. C. Farese, P. G. Ferreira, K. Ganga, M. Giacometti, E. Hivon, V. V. Hristov, A. Iacoangeli, A. H. Jaffe, A. E. Lange, L. Martinis, S. Masi, P. Mason, P. D. Mauskopf, A. Melchiorri, L. Miglio, T. Montroy, C. B. Netterfield, E. Pascale, F. Piacentini, D. Pogosyan, F. Pongetti, S. Prunet, S. Rao, G. Romeo, J. E. Ruhl, F. Scaramuzzi, D. Sforna, and N. Vittorio. First results from the BOOMERanG experiment. In *AIP Conference Proceedings*, volume 555, pages 85–94, 2001.

- [9] Small Scale Problems of the Λ CDM Model: A Short Review. <https://www.mdpi.com/2075-4434/5/1/17>.
- [10] Edwin Hubble. A relation between distance and radial velocity among extragalactic nebulae. *Proceedings of the National Academy of Sciences*, 15(3):168–173, March 1929.
- [11] LIGO Scientific Collaboration and Virgo Collaboration. Observation of Gravitational Waves from a Binary Black Hole Merger. *Physical Review Letters*, 116(6):061102, February 2016.
- [12] Mordehai Milgrom. MOND—a pedagogical review, December 2001.
- [13] Leandros Perivolaropoulos and Foteini Skara. Challenges for Λ CDM: An update, May 2021.
- [14] Georges Lemaître. L’Univers en expansion. *Annales de la Société Scientifique de Bruxelles*, 53:51, January 1933.
- [15] H. P. Robertson. Kinematics and World-Structure. *The Astrophysical Journal*, 82:284, November 1935.
- [16] A. Friedman. Über die Krümmung des Raumes. *Zeitschrift für Physik*, 10(1):377–386, December 1922.

- [17] Robert J. Nemiroff and Bijunath Patla. Adventures in Friedmann Cosmology: An Educationally Detailed Expansion of the Cosmological Friedmann Equations. *American Journal of Physics*, 76(3):265–276, March 2008.
- [18] S. Perlmutter, G. Aldering, B. J. Boyle, P. G. Castro, W. J. Couch, S. Deustua, S. Fabbro, R. S. Ellis, A. V. Filippenko, A. Fruchter, G. Goldhaber, A. Goobar, D. E. Groom, I. M. Hook, M. Irwin, A. G. Kim, M. Y. Kim, R. A. Knop, J. C. Lee, T. Matheson, R. G. McMahon, H. J. M. Newberg, C. Lidman, P. Nugent, N. J. Nunes, R. Pain, N. Panagia, C. R. Pennypacker, R. Quimby, P. Ruiz-Lapuente, B. Schaefer, N. Walton, and Supernova Cosmology Project. *Measurements of Omega and Lambda from 42 High-Redshift Supernovae*. December 1998.
- [19] Adam G. Riess, Alexei V. Filippenko, Peter Challis, Alejandro Clocchiatti, Alan Diercks, Peter M. Garnavich, Ron L. Gilliland, Craig J. Hogan, Saurabh Jha, Robert P. Kirshner, B. Leibundgut, M. M. Phillips, David Reiss, Brian P. Schmidt, Robert A. Schommer, R. Chris Smith, J. Spyromilio, Christopher Stubbs, Nicholas B. Suntzeff, and John Tonry. Observational evidence from supernovae for an accelerating universe and a cosmological constant. *Astronomical Journal*, 116(3):1009–1038, September 1998.
- [20] Michael S. Turner. Dark Matter and Dark Energy in the Universe, November 1998.

- [21] Gianfranco Bertone and Dan Hooper. A History of Dark Matter. *Reviews of Modern Physics*, 90(4):045002, October 2018.
- [22] William Thomson Kelvin. *Baltimore Lectures on Molecular Dynamics and the Wave Theory of Light*. London : C. J. Clay and sons; Baltimore, Publication agency of the Johns Hopkins university, 1904.
- [23] F. Zwicky. Die Rotverschiebung von extragalaktischen Nebeln. *Helvetica Physica Acta*, 6:110–127, January 1933.
- [24] Horace W. Babcock. The rotation of the Andromeda Nebula. *Lick Observatory Bulletin*, 498:41–51, January 1939.
- [25] Vera C Rubin and W Kent Ford Jr. Rotation of the Andromeda nebula from a spectroscopic survey of emission regions. *Astrophysical Journal*, vol. 159, p. 379, 159:379, 1970.
- [26] M. Davis, G. Efstathiou, C. S. Frenk, and S. D. M. White. The evolution of large-scale structure in a universe dominated by cold dark matter. *The Astrophysical Journal*, 292:371–394, May 1985.
- [27] Volker Springel, Simon D. M. White, Adrian Jenkins, Carlos S. Frenk, Naoki Yoshida, Liang Gao, Julio Navarro, Robert Thacker, Darren Croton, John Helly, John A. Peacock, Shaun Cole, Peter Thomas, Hugh Couchman, August Evrard, Jörg Colberg, and Frazer Pearce. Simulations of the formation,

- evolution and clustering of galaxies and quasars. *Nature*, 435(7042):629–636, June 2005.
- [28] G. F. Smoot, C. L. Bennett, A. Kogut, E. L. Wright, J. Aymon, N. W. Boggess, E. S. Cheng, G. de Amici, S. Gulkis, M. G. Hauser, G. Hinshaw, P. D. Jackson, M. Janssen, E. Kaita, T. Kelsall, P. Keegstra, C. Lineweaver, K. Loewenstein, P. Lubin, J. Mather, S. S. Meyer, S. H. Moseley, T. Murdock, L. Rokke, R. F. Silverberg, L. Tenorio, R. Weiss, and D. T. Wilkinson. Structure in the COBE Differential Microwave Radiometer First-Year Maps. *The Astrophysical Journal*, 396:L1, September 1992.
- [29] E. L. Wright, S. S. Meyer, C. L. Bennett, N. W. Boggess, E. S. Cheng, M. G. Hauser, A. Kogut, C. Lineweaver, J. C. Mather, G. F. Smoot, R. Weiss, S. Gulkis, G. Hinshaw, M. Janssen, T. Kelsall, P. M. Lubin, S. H. Moseley, Jr., T. L. Murdock, R. A. Shafer, R. F. Silverberg, and D. T. Wilkinson. Interpretation of the Cosmic Microwave Background Radiation Anisotropy Detected by the COBE Differential Microwave Radiometer. *The Astrophysical Journal*, 396:L13, September 1992.
- [30] David N Spergel, Licia Verde, Hiranya V Peiris, Eiichiro Komatsu, MR Nolta, Charles L Bennett, Mark Halpern, Gary Hinshaw, Norman Jarosik, Alan Kogut, et al. First-year wilkinson microwave anisotropy probe (wmap)*

- observations: determination of cosmological parameters. *The Astrophysical Journal Supplement Series*, 148(1):175, 2003.
- [31] LSST Science Collaboration. LSST science book, version 2.0. (November), 2009.
- [32] Rubin Observatory. LSST Camera About Page, March 2013. <https://www.lsst.org/about/camera>.
- [33] Large Synoptic Survey Telescope. Key Numbers, April 2013. <https://www.lsst.org/scientists/keynumbers>.
- [34] Zu-Cheng Chen, Chen Yuan, and Qing-Guo Huang. Confronting the primordial black hole scenario with the gravitational-wave events detected by LIGO-Virgo. *Physics Letters B*, 829:137040, June 2022.
- [35] Kailash C. Sahu et al. An isolated stellar-mass black hole detected through astrometric microlensing. *The Astrophysical Journal*, 933(1):83, July 2022. arXiv:2201.13296 [astro-ph].
- [36] The Event Horizon Telescope Collaboration. First Sagittarius A* Event Horizon Telescope Results. I. The Shadow of the Supermassive Black Hole in the Center of the Milky Way. *The Astrophysical Journal Letters*, 930(2):L12, May 2022.

- [37] S. W. Hawking. Black hole explosions? *Nature*, 248(5443):30–31, March 1974.
- [38] Jane H. MacGibbon and B. J. Carr. Cosmic rays from primordial black holes. *The Astrophysical Journal*, 371:447, April 1991.
- [39] Pedro De la Torre Luque, Jordan Koechler, and Shyam Balaji. Refining galactic primordial black hole evaporation constraints. (arXiv:2406.11949), June 2024. arXiv:2406.11949 [astro-ph].
- [40] Massimo Ricotti, Jeremiah P. Ostriker, and Katherine J. Mack. Effect of primordial black holes on the cosmic microwave background and cosmological parameter estimates. *The Astrophysical Journal*, 680(2):829–845, June 2008. arXiv:0709.0524 [astro-ph].
- [41] Hiroko Niikura, Masahiro Takada, Naoki Yasuda, Robert H. Lupton, Takahiro Sumi, Surhud More, Toshiki Kurita, Sunao Sugiyama, Anupreeta More, Masamune Oguri, and Masashi Chiba. Microlensing constraints on primordial black holes with the Subaru/HSC Andromeda observation. *Nature Astronomy*, 3(6):524–534, April 2019. arXiv:1701.02151 [astro-ph].
- [42] P. Mroz, A. Udalski, M. K. Szymanski, M. Kapusta, I. Soszynski, L. Wyrzykowski, P. Pietrukowicz, S. Kozłowski, R. Poleski, J. Skowron, D. Skowron, K. Ulaczyk, M. Gromadzki, K. Rybicki, P. Iwanek, M. Wrona,

and M. Ratajczak. Microlensing optical depth and event rate toward the large magellanic cloud based on 20 years of ogle observations. *The Astrophysical Journal Supplement Series*, 273(1):4, July 2024. arXiv:2403.02398 [astro-ph, physics:gr-qc, physics:hep-ph].

[43] P. Tisserand, L. Le Guillou, C. Afonso, J. N. Albert, J. Andersen, R. Ansari, É Aubourg, P. Bareyre, J. P. Beaulieu, X. Charlot, C. Coutures, R. Ferlet, P. Fouqué, J. F. Glicenstein, B. Goldman, A. Gould, D. Graff, M. Gros, J. Haissinski, C. Hamadache, J. de Kat, T. Lasserre, É Lesquoy, C. Loup, C. Magneville, J. B. Marquette, É Maurice, A. Maury, A. Milsztajn, M. Moniez, N. Palanque-Delabrouille, O. Perdureau, Y. R. Rahal, J. Rich, M. Spiro, A. Vidal-Madjar, and L. Vigroux. Limits on the Macho content of the Galactic Halo from the EROS-2 Survey of the Magellanic Clouds. *Astronomy & Astrophysics*, 469(2):387–404, July 2007.

[44] C. Alcock, R. A. Allsman, D. R. Alves, T. S. Axelrod, A. C. Becker, D. P. Bennett, K. H. Cook, A. J. Drake, K. C. Freeman, M. Geha, K. Griest, M. J. Lehner, S. L. Marshall, D. Minniti, C. A. Nelson, B. A. Peterson, P. Popowski, M. R. Pratt, P. J. Quinn, C. W. Stubbs, W. Sutherland, A. B. Tomaney, T. Vandehei, and D. L. Welch. The macho project: Microlensing optical depth towards the galactic bulge from difference image analysis. *The Astrophysical Journal*, 541(2):734–766, October 2000. arXiv:astro-ph/0002510.

- [45] K. Schwarzschild. On the gravitational field of a mass point according to einstein's theory. (arXiv:physics/9905030), May 1999. arXiv:physics/9905030.
- [46] E. B. Christoffel. Ueber die transformation der homogenen differentialausdrücke zweiten grades. *Journal für die reine und angewandte Mathematik*, 1869(70):46–70, 1869.
- [47] J. M. Bardeen. Timelike and null geodesics in the Kerr metric. *Proceedings, Ecole d'Eté de Physique Théorique: Les Astres Occlus : Les Houches, France, August, 1972, 215-240*, pages 215–240, 1973.
- [48] Peter Schneider, Jürgen Ehlers, and Emilio E. Falco. *Gravitational Lenses*. Astronomy and Astrophysics Library. Springer Berlin Heidelberg, Berlin, Heidelberg, 1992.
- [49] W. Patrick McCRAY. How Astronomers Digitized the Sky. *Technology and Culture*, 55(4):908–944, 2014.
- [50] W. S. Boyle and G. E. Smith. Charge coupled semiconductor devices. *The Bell System Technical Journal*, 49(4):587–593, April 1970.
- [51] J. A. Tyson and J. F. Jarvis. Evolution of galaxies: Automated faint object counts to 24th magnitude. *The Astrophysical Journal*, 230:L153–L156, June 1979.

- [52] R.W. Smith and J.N. Tatarewicz. Replacing a technology: The large space telescope and CCDs. *Proceedings of the IEEE*, 73(7):1221–1235, 1985.
- [53] The Royal Swedish Academy of Sciences. The Nobel Prize in Physics 2009. <https://www.nobelprize.org/prizes/physics/2009/press-release/>.
- [54] A. Einstein. Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt. *Annalen der Physik*, 322(6):132–148, 1905.
- [55] Nikita Medvedev and Baerbel Rethfeld. Dynamics Of Electronic Excitation Of Solids With Ultrashort Laser Pulse. *AIP Conference Proceedings*, 1278:250, April 2010.
- [56] K. Rajkanan, R. Singh, and J. Shewchun. Absorption coefficient of silicon for solar cell calculations. *Solid-State Electronics*, 22(9):793–795, September 1979.
- [57] Hank Zumbahlen. Sensors. In *Linear Circuit Design Handbook*, pages 193–243. Elsevier, 2008.
- [58] P. Antilogus, S. Bailey, Ph Bailly, H. Lebbolo, D. Martin, R. Sefri, C. de La Taille, J. Jeglot, M. Moniez, V. Tocut, and F. Wicek. ASPIC: LSST camera readout chip. Comparison between DSI and C&S. In *Topical Workshop on Electronics for Particle Physics (TWEPP-09)*, September 2009.

- [59] Piero Dal Tio, Giada Pastorelli, Alessandro Mazzi, Michele Trabucchi, Guglielmo Costa, Alice Jacques, Adriano Pieres, Léo Girardi, Yang Chen, Knut A. G. Olsen, Mario Juric, Željko Ivezić, Peter Yoachim, William I. Clarkson, Paola Marigo, Thaise S. Rodrigues, Simone Zaggia, Mauro Barbieri, Yazan Momany, Alessandro Bressan, Robert Nikutta, and Luiz Nicolaci da Costa. Simulating the legacy survey of space and time stellar content with trilegal. *The Astrophysical Journal Supplement Series*, 262(1):22, August 2022.
- [60] Peter Yoachim, Lynne Jones, Jr. Eric H. Neilsen, Tiago, John Parejko, Jeff Carlin, Matthew R. Becker, pgris, Loredana Prisinzano, Erik Dennihy, Eric Bellm, Jonathan Sick, lmptc, LI, nsabrams, Leanne Guy, Katja Bricman, Johan Bregeon, Kian-Tat Lim, Michael Kelley, and Igor Andreoni. lsst/rubin_sim: v1.3.2, October 2023. <https://doi.org/10.5281/zenodo.8409175>.
- [61] L. Girardi, M. A. T. Groenewegen, E. Hatziminaoglou, and L. Da Costa. Star counts in the galaxy: Simulating from very deep to very shallow photometric surveys with the trilegal code. *Astronomy & Astrophysics*, 436(3):895–915, June 2005.
- [62] Alessandro Bressan, Paola Marigo, Léo Girardi, Bernardo Salasnich, Claudia Dal Cero, Stefano Rubele, and Ambra Nanni. parsec : stellar tracks

- and isochrones with the padova and trieste stellar evolution code: parsec : tracks and isochrones. *Monthly Notices of the Royal Astronomical Society*, 427(1):127–145, November 2012.
- [63] Paola Marigo, Alessandro Bressan, Ambra Nanni, Léo Girardi, and Maria Letizia Pumo. Evolution of thermally pulsing asymptotic giant branch stars – I. The colibri code. *Monthly Notices of the Royal Astronomical Society*, 434(1):488–526, September 2013.
- [64] Patrick S. Osmer, Julia D. Kenefick, Patrick B. Hall, and Richard F. Green. A Deep Multicolor Survey. IV. The Electronic Stellar Catalog. *The Astrophysical Journal Supplement Series*, 119(2):189, December 1998.
- [65] S. Zaggia, I. Hook, R. Mendez, L. da Costa, L. F. Olsen, M. Nonino, A. Wicenec, C. Benoist, E. Deul, T. Erben, M. D. Guarnieri, R. Hook, I. Prandoni, M. Scodreggio, R. Slijkhuis, and R. Wichmann. ESO Imaging Survey - IV. Multicolor analysis of point-like objects toward the South Galactic Pole. *Astronomy and Astrophysics Supplement Series*, 137(1):75–81, May 1999.
- [66] S. Arnouts, B. Vandame, C. Benoist, M. A. T. Groenewegen, L. Da Costa, M. Schirmer, R. P. Mignani, R. Slijkhuis, E. Hatziminaoglou, R. Hook, R. Madejsky, C. Rit e, and A. Wicenec. ESO imaging survey: Deep public sur-

- vey: Multi-color optical data for the Chandra Deep Field South. *Astronomy & Astrophysics*, 379(2):740–754, November 2001.
- [67] Nolan Smyth and William DeRocco. Nolangsmyth/lenscalcpy: v0.0.4, March 2024. <https://doi.org/10.5281/zenodo.10734644>.
- [68] Anatoly Klypin, HongSheng Zhao, and Rachel S. Somerville. Λ cdm-based models for the milky way and m31. i. dynamical models. *The Astrophysical Journal*, 573(2):597, July 2002.
- [69] W. Dehnen and J. Binney. Mass models of the Milky Way. *Monthly Notices of the Royal Astronomical Society*, 294(3):429–438, March 1998.
- [70] J. Brand and L. Blitz. The velocity field of the outer galaxy. *Astronomy and Astrophysics*, 275:67–90, August 1993.
- [71] F. Pont, D. Queloz, P. Bratschi, and M. Mayor. Rotation of the outer disc from classical cepheids. *Astronomy and Astrophysics*, 318:416–428, February 1997.
- [72] K. Kuijken and G. Gilmore. The mass distribution in the galactic disc - I. A technique to determine the integral surface mass density of the disc near the Sun. *Monthly Notices of the Royal Astronomical Society*, 239(2):571–603, July 1989.

- [73] F. J. Kerr and D. Lynden-Bell. Review of galactic constants. *Monthly Notices of the Royal Astronomical Society*, 221:1023–1038, August 1986.
- [74] M. Lindqvist, H. J. Habing, and A. Winnberg. OH/IR stars close to the galactic centre. II. Their spatial and kinematics properties and the mass distribution within 5-100 PC from the galactic centre. *Astronomy and Astrophysics*, 259:118–127, June 1992.
- [75] R. Drimmel and D. N. Spergel. Three Dimensional Structure of the Milky Way Disk. *The Astrophysical Journal*, 556(1):181–202, July 2001.
- [76] Casey Y. Lam, Jessica R. Lu, Jr Hosek, William A. Dawson, and Nathan R. Golovich. Popsycle: A new population synthesis code for compact object microlensing events. *The Astrophysical Journal*, 889(1):31, January 2020. arXiv:1912.04510 [astro-ph].
- [77] Matthew W. Hosek, Jessica R. Lu, Jay Anderson, Francisco Najarro, Andrea M. Ghez, Mark R. Morris, William I. Clarkson, and Sandra M. Albers. The unusual initial mass function of the arches cluster. *The Astrophysical Journal*, 870(1):44, January 2019.
- [78] Nicholas Metropolis, Arianna W. Rosenbluth, Marshall N. Rosenbluth, Augusta H. Teller, and Edward Teller. Equation of state calculations by fast

- computing machines. *Journal of Chemical Physics*, 21:1087–1092, June 1953.
ADS Bibcode: 1953JChPh..21.1087M.
- [79] Natasha S. Abrams, Markus P. G. Hundertmark, Somayeh Khakpash, Rachel A. Street, R. Lynne Jones, Jessica R. Lu, Etienne Bachelet, Yian-nis Tsapras, Marc Moniez, Tristan Blaineau, Rosanne Di Stefano, Martin Makler, Anibal Varela, and Markus Rabus. Microlensing discovery and characterization efficiency in the vera c. rubin legacy survey of space and time. (arXiv:2309.15310), September 2023. arXiv:2309.15310 [astro-ph].
- [80] O. A. Gonzalez, M. Rejkuba, M. Zoccali, E. Valent, D. Minniti, and R. Tobar. Reddening and metallicity maps of the milky way bulge from vvv and 2mass: iii. the first global photometric metallicity map of the galactic bulge. *Astronomy & Astrophysics*, 552:A110, April 2013.
- [81] S. Chandrasekhar. The maximum mass of ideal white dwarfs. *The Astrophysical Journal*, 74:81, July 1931. ADS Bibcode: 1931ApJ....74...81C.
- [82] J. R. Oppenheimer and G. M. Volkoff. On massive neutron cores. *Physical Review*, 55(4):374–381, February 1939.
- [83] J. R. Oppenheimer and H. Snyder. On continued gravitational contraction. *Physical Review*, 56(5):455–459, September 1939.

- [84] Thomas A. Callister, Will M. Farr, and Mathieu Renzo. State of the field: Binary black hole natal kicks and prospects for isolated field formation after GWTC-2. *The Astrophysical Journal*, 920(2):157, October 2021.
- [85] Oliver D. Elbert, James S. Bullock, and Manoj Kaplinghat. Counting black holes: The cosmic stellar remnant population and implications for ligo. *Monthly Notices of the Royal Astronomical Society*, 473(1):1186–1194, January 2018.
- [86] Feryal Ozel, Dimitrios Psaltis, Ramesh Narayan, and Jeffrey E. McClintock. The Black Hole Mass Distribution in the Galaxy. *The Astrophysical Journal*, 725(2):1918–1927, December 2010.
- [87] Kareem El-Badry, Hans-Walter Rix, and Tyler M Heintz. A million binaries from Gaia eDR3: sample selection and validation of Gaia parallax uncertainties. *Monthly Notices of the Royal Astronomical Society*, 506(2):2269–2295, July 2021.
- [88] Maxwell Moe and Rosanne Di Stefano. Mind your ps and qs: The interrelation between period (p) and mass-ratio (q) distributions of binary stars. *The Astrophysical Journal Supplement Series*, 230(2):15, June 2017.
- [89] C. Alcock, R. A. Allsman, T. S. Axelrod, D. P. Bennett, K. H. Cook, K. C. Freeman, K. Griest, J. A. Guern, M. J. Lehner, S. L. Marshall, H. S. Park,

- S. Perlmutter, B. A. Peterson, M. R. Pratt, P. J. Quinn, A. W. Rodgers, C. W. Stubbs, and W. Sutherland. The macho project first-year large magellanic cloud results: The microlensing rate and the nature of the galactic dark halo. *The Astrophysical Journal*, 461:84, April 1996. ADS Bibcode: 1996ApJ...461...84A.
- [90] Eric W. Weisstein. Dwarf Nova – from Eric Weisstein’s World of Astronomy. <https://scienceworld.wolfram.com/astronomy/DwarfNova.html>.
- [91] Izumi Hachisu and Mariko Kato. A light curve analysis of 32 recent galactic novae — distances and white dwarf masses. *The Astrophysical Journal Supplement Series*, 242(2):18, June 2019. arXiv:1905.10655 [astro-ph].
- [92] M. Catelan and H.A. Smith. *Pulsating Stars*. Wiley, 2015.