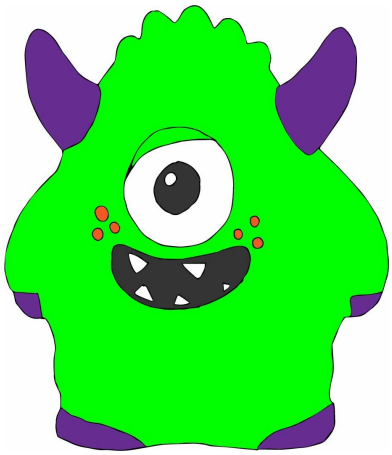
Discovery and Simulation of Solar System Bodies in the Age of Big Data and Artificial Intelligence

by
Kevin J. Napier

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Physics)
in The University of Michigan
2023

Doctoral Committee:

Professor David W. Gerdes, Co-Chair
Professor Fred C. Adams, Co-Chair
Professor Gary M. Bernstein, University of Pennsylvania
Dr. Matthew J. Holman, Harvard-Smithsonian Center for Astrophysics
Professor Michael R. Meyer



Create uninhibited.

Kevin J. Napier

kjnapier@umich.edu

ORCID iD: 0000–0003–4827–5049

The scientist does not study nature because it is useful. He studies it because he delights in it, and he delights in it because it is beautiful.

— Henri Poincaré

For all who delight in nature's beauty.

Acknowledgments

It's dangerous to go alone!

— Unnamed Old Man, *The Legend of Zelda*

I am so grateful that I had the opportunity to spend the past five years studying the most enticing secrets of our solar system. Although I am a fairly solitary creature, I recognize that I could never have completed this thesis on my own; it was only possible because of the help of dozens of people, a few of whom really stand out in my mind.

I was lucky to have not just one, but two incredible advisors, David Gerdes and Fred Adams. Both Dave and Fred allowed me absolute freedom to pursue new ideas, and gave me guidance whenever I needed it; they never led me astray. Dave's careful approach to pulling faint signals out of large noisy datasets, and Fred's incredible ability to distill any problem to its essence, left profound impressions on the way I think about science.

In addition to my advisors, I had a number of fantastic collaborators whose guidance was invaluable. First, although Ed Lin was not an official advisor, he may as well have been. For my first few years of graduate school, I would turn around and ask Ed whenever I had a question; he was like Google for planetary science. I also want to thank Pedro Bernardinelli, Matt Holman, and Konstantin Batygin for countless insightful discussions. I owe a great deal of gratitude to David Trilling, Mario Jurić, and the rest of the DEEP Collaboration for providing me with a singularly powerful dataset that enabled the second half of this thesis. I am grateful to NASA's

New Horizons team for letting me play with their spaceship, and for providing me with a seemingly endless list of difficult problems that have helped me to grow as a scientist. I am thankful to my committee for volunteering to take the time to read this tome. Finally, I want to thank John Moustakas for showing me the ropes of academia when I was an undergraduate.

Of course, I could never have even started graduate school if it weren't for my invariably supportive family. I can't possibly begin to mention all that you have done for me, so you'll each get your own bullet point. To avoid any kerfuffles, I will thank you each in the order that I met you.

Mom for the wit and the meticulous attention to detail that made all of this work possible.

Dad for the work ethic to see it through to completion.

Stevie for teaching me the gumption to voice my opinion and do what I know is right, regardless of opposition.

Brian for giving me the thick skin that academia seems to inexplicably require.

Grandma for your unconditional love. I wish you were here to see this.

Aunt Deb for your unwavering support of my academic endeavors.

Jeremy for showing me how to be unabashedly myself.

I would be remiss if I didn't recognize the contributions of my rubber duck debugger Jerry, who was always around to eagerly lend an ear and an eye.

Finally I have to thank Claire for constantly challenging me with insightful discussions that have made be a better scientist and person; for encouraging me to reach for new horizons; for sharing my delight in nature's beauty; and for so much more. Most importantly, thank you for showing me that there is more to life than work. You inspire me every day.

Table of Contents

Part I: Dynamics

Part II: Discovery

List of Figures

List of Tables

Abstract

As science enters an era defined by big data, high-performance computing, and artificial intelligence (AI), we have an opportunity to fundamentally change the way we study our solar system. In this thesis I leverage state-of-the-art computing technology and AI techniques to do analyses that would have been infeasible only a few years ago. The text is split into two parts, corresponding to *dynamics* and *discovery*.

In Part I, I investigate problems in solar system dynamics by combining pure analytic theory with highly detailed simulations. First I examine the claim that the orbits of the so-called extreme Trans-Neptunian Objects (ETNOs) are being aligned by the gravitational influence of the hypothesized Planet Nine. By carefully simulating the three most productive Kuiper Belt surveys of the past decade, I show that when the surveys' observational biases are fully accounted for, their ETNO detections are consistent with being drawn from an isotropic population, meaning Planet Nine is not necessary to explain the apparent clustering. Next, inspired by the recent detections of the first two interstellar objects 'Oumuamua and Borisov, I investigate the capture of interstellar objects by our solar system. Using a suite of numerical simulations, I calculate the capture cross section for interstellar objects as a function of hyperbolic excess velocity, as well as the characteristic dynamical lifetime of captured objects. Finally in Chapter 6, motivated by the continued absence of any known dynamically stable Earth Trojans, I propose a collisional mechanism by which such objects may have been destabilized, and show that an initially-stable population of primordial Earth Trojans would have been severely disrupted by asteroid impacts on Earth.

In Part II of this thesis I develop new techniques for discovering faint solar system bodies, beginning with a search for Kuiper Belt objects in data from the DECam Ecliptic Exploration Project (DEEP). By combining a shift-and-stack technique with convolutional neural networks, I reach an average depth of $m_r \sim 26.2$, accumulating ~ 2300 single-epoch detections. The main scientific result of this work is a measurement of the absolute magnitude distribution of the Cold Classical Kuiper Belt. Our data is well-fit by an exponentially-tapered power law, which is the functional form predicted by streaming instability simulations of planetesimal formation. In Chapter 8 I adapt the techniques from the DEEP search to do a proof-of-concept shift-and-stack search for Jupiter Trojans. The search yielded more than 100 new Jupiter Trojans, with an extremely low false positive rate prior to human inspection. In Chapter 9 I present a search for a flyby target for NASA's New Horizons spacecraft. Using data from four DECam fields, I used a novel combination of shift-and-stack, convolutional neural networks, and orbit linking to reject false positive detections. While the search did not turn up a flyby target, it did uncover some candidate objects that may be observable by the spacecraft. The technique is extremely promising, and warrants further study. Finally I introduce a novel approach to the minor planet detection problem, called `HelioStack`. By shifting perspective from the topocentric frame to the heliocentric frame, and then choosing an appropriate orbit parameterization, the algorithm can drastically simplify the description of the apparent motion of bodies on Keplerian orbits. It will be critical for enabling sub-threshold minor planet detection in future astronomical surveys such as LSST.

Chapter 1

Introduction

> It's a dangerous business, Frodo, going out of your door. You step into the road, and if you don't keep your feet, there is no telling where you might be swept off to.
>
> — Bilbo Baggins, *The Fellowship of the Ring*

While the study of our solar system is among humanity's oldest scientific endeavors, the basic underlying questions have remained the same for thousands of years:

— *What is out there?*

— *Why is our solar system the way it is?*

In the past 500 years, we have made significant progress on both fronts. Our ability to discover solar system bodies has largely been pushed by technological innovation, and our theoretical understanding of the solar system has been pushed by observations of solar system bodies, along with innovations in mathematics, physics, and most recently, computing technology.

1.1 Technological Innovation Drives Discovery

As I dig into the current state of our knowledge of the solar system, I would like to begin by demonstrating how our ability to discover solar system bodies has been driven by technological innovation. I encourage the reader to refer to Figures 1.1 and 1.2 while reading through this section. Figure 1.1 shows the visual magnitude at discovery vs epoch of discovery for all known

bodies with well-determined orbits.[1] Figure 1.2 shows the cumulative total of such bodies as a function of time, broken into in various dynamical classes. Together, these figures tell a story of the innovations that have enhanced our ability to detect minor planets.

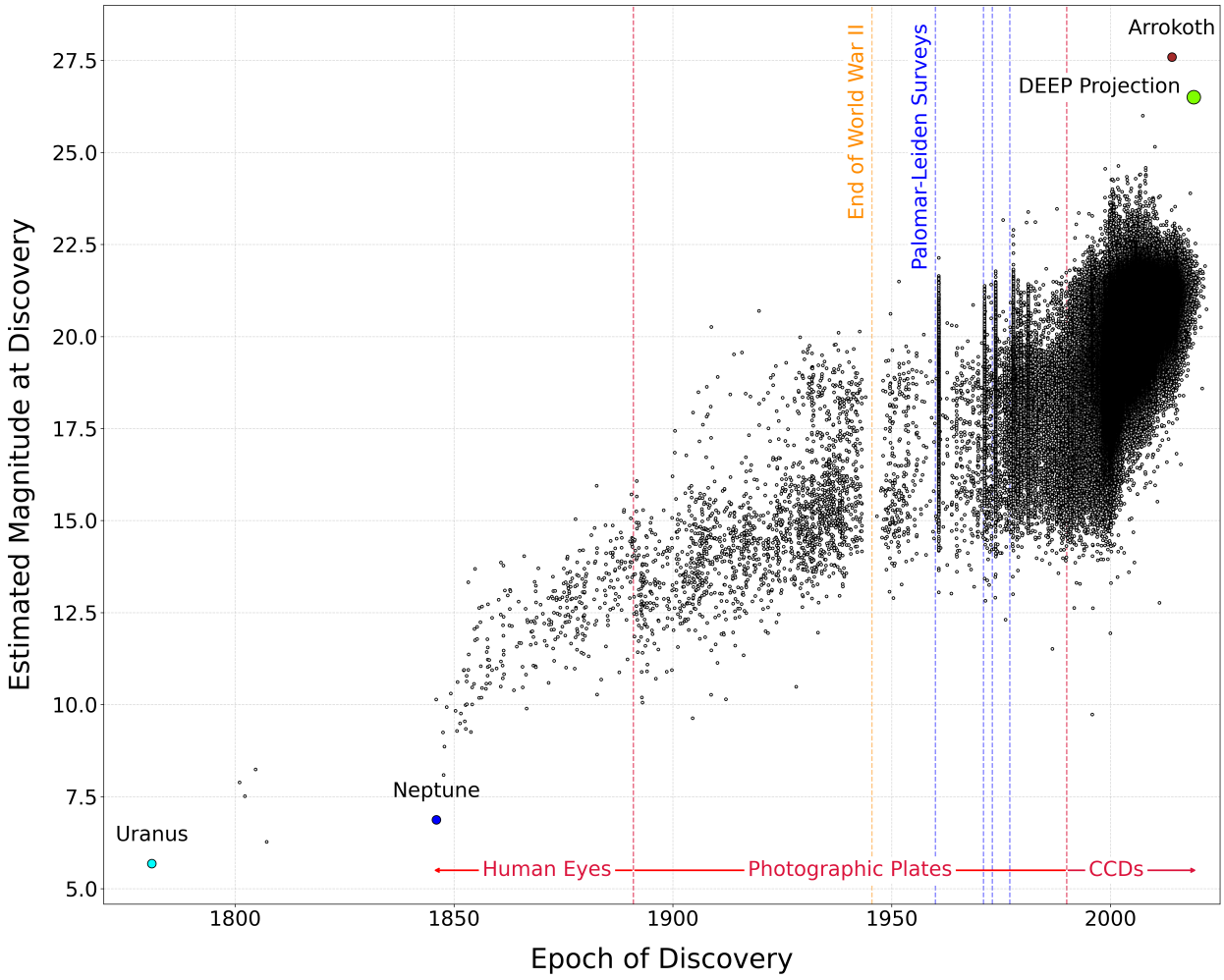


Figure 1.1: Visual magnitude of all numbered objects listed in the Minor Planet Center, as viewed from the Geocenter at the epoch of discovery.

The Telescope

The Sun, Mercury, Venus, Earth, the Moon, Mars, Jupiter, and Saturn are all visible to the naked eye, and have been known since prehistoric times. Aside from the occasional comet, these bodies were the only known residents of the solar system for the first several thousand years of recorded

[1]Specifically, I used all numbered bodies listed in the Minor Planet Center (MPC).

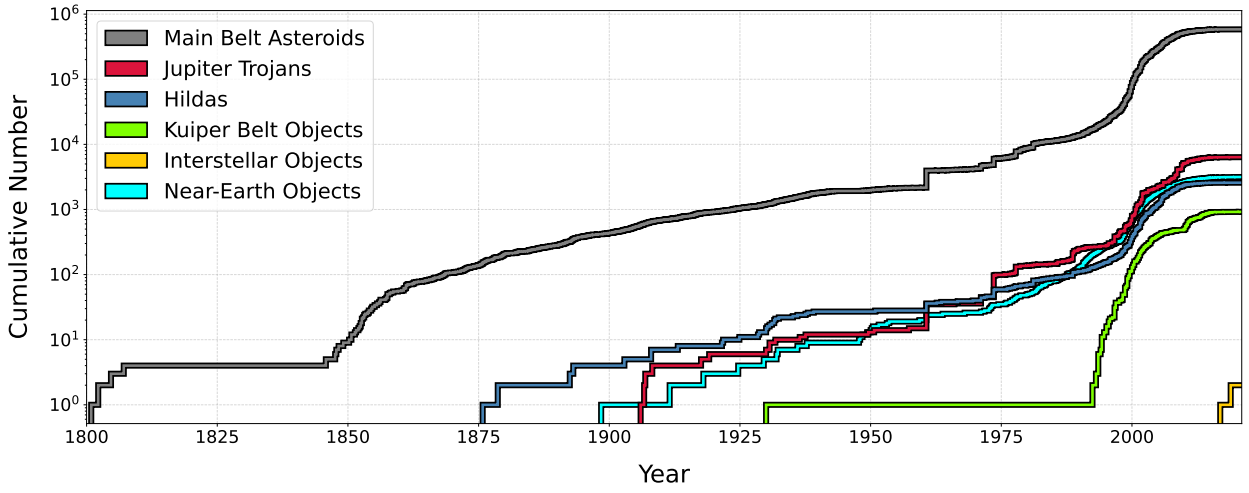


Figure 1.2: Cumulative total of numbered minor bodies in major dynamical classes as a function of time. Note that the ISOs are not numbered, but I have chosen to include them anyway for the purpose of demonstration.

human history. That changed in 1609 when Galileo first turned a telescope skyward, and almost immediately discovered Jupiter's four largest moons, *Io*, *Europa*, *Ganymede*, and *Callisto*. The so-called Galilean satellites were the first solar system bodies that can truly be said to have been *discovered* by an individual.

Over the next few decades astronomers made incremental improvements to the telescope, and discovered Saturn's five largest satellites. No new discoveries came for more than 100 years until 1781, when William Herschel reported his discovery of the planet Uranus (Herschel & Watson, 1781). The planet, which is just barely bright enough to be seen by the human eye under the darkest conditions, had eluded discovery until Herschel happened to spot it through his telescope.

In 1801 Giuseppe Piazzi discovered the first asteroid, *Ceres*. In subsequent years, astronomers searching for a planet between Mars and Jupiter discovered the asteroids *Juno*, *Pallas* and *Vesta*. In 1845 the discovery of a fifth asteroid, *Astrea*, prompted astronomers to realize that there is no large planet between Mars and Jupiter, but rather a large population of small bodies that we now call the asteroid belt. Shortly thereafter the rate of new discoveries ramped up, as astronomers began finding asteroids at a rate of about five per year.

Concurrently with the discovery of the asteroid belt, Le Verrier (1846) and Adams (1846)

independently predicted the existence of an eighth planet to explain anomalies in the orbit of Uranus, and within the year Neptune was discovered (Galle, 1846).[2]

The Photographic Plate

For the half-century following the discovery of the asteroid belt, astronomers continued to discover new asteroids at a rate of a handful per year, meticulously scanning the sky by eye through telescopes. The rate of asteroid discoveries saw its first boom in 1891 when Max Wolf introduced the use of photographic plates with the discovery the asteroid *Brucia*. With a record 20 discoveries in one year, Campbell (1892) claimed,

> "... at the present rate of [asteroid] discovery the resources of the computing staff of Berlin will soon be severely taxed to furnish satisfactory orbits for them."

By today's standards, determining orbits for 20 asteroids is rather simple. In the time before modern computers, however, the task would have been enormous. The photographic plate allowed astronomers to store their data to be examined at a later date by multiple people, rather than having a single person analyzing the sky through a telescope. The dramatic effect of the use of photographic plates is evident in Figure 1.1.

In the half-century following the introduction of the photographic plate, astronomers made incremental progress to the size and quality of telescopes, facilitating the detection of fainter solar system bodies. In the meantime, astronomers made two extremely important discoveries. In 1906, Wolf discovered the first known Jupiter Trojan *Achilles*. The discovery was the first observational confirmation of Lagrange's prediction that there exist two points of stable equilibrium in the restricted three-body problem (Lagrange, 1776). Then in 1930, Clyde Tombaugh discovered the (dwarf) planet Pluto (Tombaugh, 1946). At the time, Pluto was the most distant body ever discovered, pushing our understanding of the solar system to new horizons.

[2]For a nice review of the subject, see Batygin et al. 2019.

Large Telescopes and Astronomical Surveys

The next major advance came with the advent of large telescopes and the astronomical survey. In 1960, the Palomar-Leiden survey used a 48-inch Schmidt telescope with a 6.6 square degree field of view and superior image quality to record 130 images of the sky onto photographic plates, with the express purpose of finding solar system bodies. The effectiveness of the survey is clear in Figure 1.1. By leveraging a more sensitive observing instrument and systematically recording images onto photographic plates, the survey discovered of order 2000 new asteroids, with unprecedented sensitivity to faint objects (van Houten et al., 1970). The Palomar-Leiden survey had three additional observing campaigns in the 1970s, yielding a few thousand new asteroids and a few dozen new Jupiter Trojans (van Houten-Groeneveld et al., 1989; Van Houten et al., 1984). Although there were other fruitful surveys in the meantime, there were no more major leaps until the advent of the CCD.

The CCD

Although the CCD was invented in the 1970s, it took some time to be adopted by astronomers. It is likely that first *newly discovered* minor planet via CCD was the asteroid 1984 HW_1 (personal correspondence with Matt Holman and Gareth Williams). When the CCD was finally widely adopted by astronomers in the 1990s, it revolutionized astronomy. CCDs have nearly perfect quantum efficiency, enabling much greater sensitivity than photographic plates, which typically have quantum efficiency of order a few percent. Furthermore the ability to store data digitally (rather than on bulky photographic plates) enabled much more ambitious surveys. The dramatic effect of the CCD is apparent in Figures 1.1 and 1.2.

Notably, Jewitt & Luu (1993) used a CCD to discover the minor planet 1992 QB_1, only the second Kuiper Belt Object (KBO) ever discovered, and the first since the discovery of Pluto in 1930. The discovery provided the first observational confirmation of the long-theorized Kuiper Belt (see, e.g., Edgeworth 1949, Fernandez & Ip 1984, and Duncan et al. 1988), and kicked off a slew of subsequent discoveries.

The Space Telescope

Ground-based observatories are plagued by Earth's atmosphere. In the daytime the scattering of the Sun's light off of the atmosphere prevents optical observations altogether. In the nighttime when it is possible to see through the atmosphere, it still causes distortions to the incoming light from astronomical sources. Space-based telescopes, on the other hand, are not subject to atmospheric distortions, resulting in incredible image quality. While there have been several space telescopes launched over the years, the most important for solar system detections has been the Hubble Space Telescope (HST), launched by NASA and the European Space Agency (ESA) in 1990.

Bernstein et al. (2004) used a large campaign of HST images along with a software technique for combining images to find sub-threshold detections (see Chapter 7) to find what are still the faintest minor bodies ever observed. The objects, which have visual magnitudes greater than 28.0, have never been re-observed, so they do not appear in Figures 1.1 and 1.2.

In 2014 as NASA's New Horizons (NH) spacecraft was passing into the Kuiper Belt after observing Pluto and Charon, the NH science team was struggling to find a new flyby target using ground-based telescopes. They simply needed to be able to find fainter objects to have any real chance of finding a target. The team turned to an observing campaign using HST, resulting in the discovery of the KBO Arrokoth (Buie et al., 2020). NH flew by Arrokoth in 2019, returning images that have offered us some of our deepest insight about planetesimal formation (Spencer et al., 2020; Grundy et al., 2020; McKinnon et al., 2020). At the moment, Arrokoth is the faintest numbered object in the MPC (making it the faintest object in Figure 1.1), and is the only numbered object for which no ground-based observations have ever been reported.

1.2 The Current Inventory of our Solar System

Having now touched on the major technological innovations that have driven our ability to discover solar system bodies, I will give a brief overview of the current inventory of our solar system. Figure

1.3 shows the position of all solar system bodies that have been observed at multiple oppositions, projected into the ecliptic plane on 29 June 2023. In order to show all of the important details on the same set of axes, I have scaled the radial axis such that it is linear between 0 and 6 au, and logarithmic thereafter.

More than 1 million minor planet discoveries over the past four centuries have revealed rich dynamical structures in the solar system that have been sculpted over billions of years by the gravity of the major planets. It turns out that most of the bodies in the solar system belong to just a few unique groups. Broadly speaking, the most populous known dynamical classes are the asteroid belt, the Jupiter Trojans, the Hildas, the Near-Earth Objects (NEOs), and the Kuiper Belt.

The asteroid belt, existing in the space between Mars and Jupiter, is by far the most populous of the known solar system populations, with the current inventory exceeding 1 million bodies. The asteroid belt itself contains rich dynamical subclasses hewn by gravitational interactions with the planets, most notably mean motion resonances with Jupiter (Murray & Dermott, 2000). It also serves as a reservoir for objects to be kicked onto Earth-crossing orbits, becoming so-called NEOs. The NEOs pose a threat to life on Earth, and as such there have been a multitude of surveys working to detect them.

Further out are the Hildas and the Jupiter Trojans. Jupiter's gravity is so dominant in the inner solar system that it has trapped a massive herd of minor planets into mean-motion resonances; the Hildas are in a 2:3 internal mean-motion resonance with Jupiter, and the Trojans are in a 1:1 mean-motion resonance. The resonances between these bodies and Jupiter act as a dynamical protection mechanism, enabling the orbits to remain stable over Gyr-timescales. I will discuss the Jupiter Trojans in more detail in Chapters 2 and 8.

The final major population in Figure 1.3 is the Kuiper Belt, existing beyond the orbit of Neptune. The Kuiper Belt features an extremely varied set of dynamical classes created by Neptune's gravity. In no particular order, the primary dynamical classes are as follows.

— Cold Classicals—a quiescent population of KBOs that formed *in situ*, and have remained

dynamically unchanged since their formation 4.5 billion years ago. They are our best window into the primordial solar system, and are the primary focus of the analysis in Chapter 7.

— Hot Classicals, which have similar semi-major axes to the Cold Classicals, but higher eccentricities and inclinations.

— Bodies in mean-motion resonance with Neptune. The inventory stands at more than 1000 including two dozen Neptune Trojans and a few hundred Plutinos (objects in the outer 3:2 resonance), among more than 50 unique resonances.

— Scattering objects, which are currently undergoing rapid orbital evolution due to gravitational interactions with Neptune.

— Detached objects, which have achieved orbits that are largely decoupled from Neptune's gravitational influence, causing their orbits to evolve slowly.

— Centaurs, named as such because they spend time both interior and exterior to Neptune's orbit. Their orbits bring them inside of Neptune's orbit, into the realm of the other planets where they are subject to rapid and extreme dynamical evolution, sometimes acquiring cometary orbits.[3]

The above points are only a brief glimpse into the rich dynamical landscape of the Kuiper Belt; for full treatments, see Gladman et al. (2008), Batygin et al. (2019), Khain et al. (2020a) and Gladman & Volk (2021).

Although they are not visible in Figure 1.3, there are now two known instances of interloping interstellar comets, 'Oumuamua (Meech et al., 2017) and Borisov (Borisov, 2019; Jewitt & Luu, 2019).

Finally the rest of the bodies in Figure 1.3, represented as faint black points, do not fall into any of the above categories. For the most part they are on rapidly evolving planet-crossing orbits.

[3]Note that there are multiple definitions of centaurs, one of which refers to objects with orbits entirely between Jupiter and Saturn.

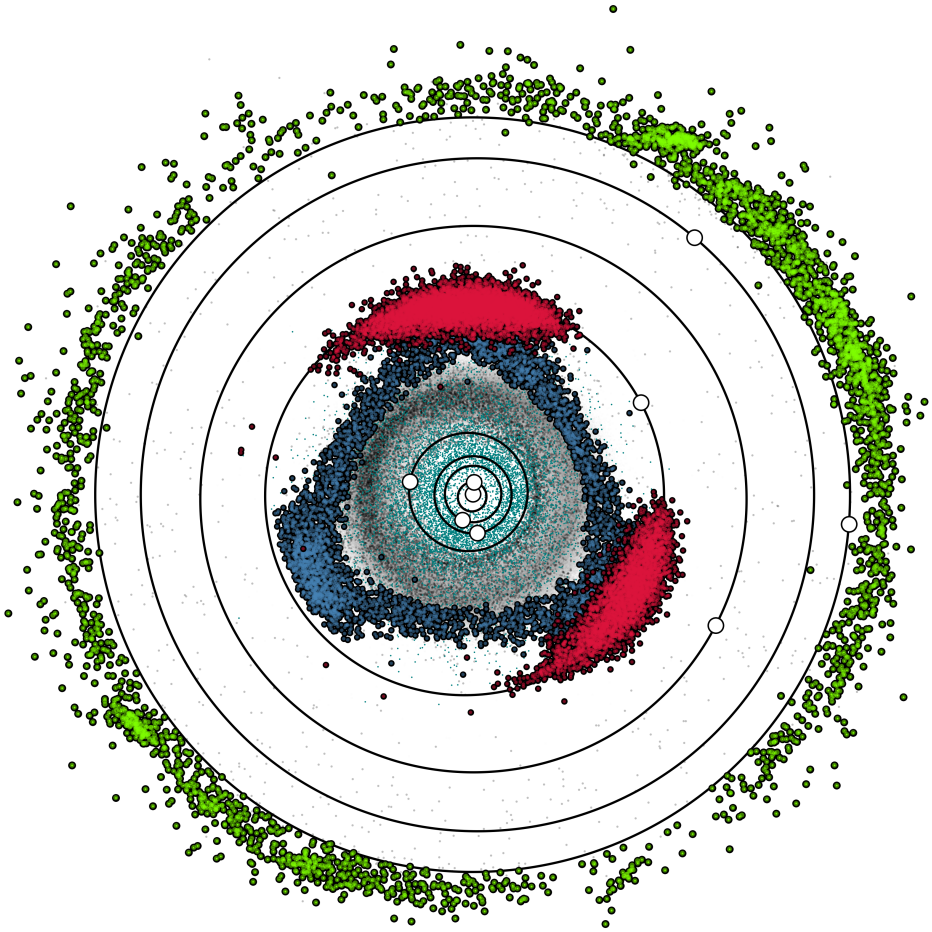


Figure 1.3: The current inventory of multi-opposition solar system bodies, projected into the ecliptic plane on 29 June 2023. The radial axis is scaled such that it is linear between 0 and 6 au, and logarithmic thereafter. The main visible structures are the asteroid belt (grey), the near-Earth objects (light blue), the Hildas (dark blue), the Jupiter Trojans (red), and the Kuiper Belt (green). The faint black points scattered throughout the plot do not belong to any of the aforementioned dynamical classes.

1.3 Our Theoretical Understanding the Solar System

Before I describe the new work in this thesis, I want to set the stage by briefly addressing the most relevant aspects of our theoretical understanding of the solar system.

1.3.1 The Laws of Planetary Motion

Early models of the solar system placed Earth at the center, and attempted to use complicated epicyclic motion to explain the apparent motions of the planets on the sky. In the mid 1500s, Nicolaus Copernicus offered a radical shift in perspective: perhaps the Sun is at the center of the solar system, and the planets (including Earth) orbit around it on *perfectly circular orbits* (Copernicus, 1543). Due to the imposition of circular orbits the idea was not perfect—it still required the use of epicycles to describe the motions of the planets. However, the model more naturally explained phenomena such as the occasional retrograde sky motion of the planets. Copernicus' idea illustrates the remarkable power of examining a problem from a different perspective.

In the early 1600s, Johannes Kepler used Copernicus' idea of a heliocentric solar system along with Tycho Brahe's meticulous planetary ephemeris tables to deduce three laws of planetary motion (Kepler, 1609, 1619). The laws are as follows:

1. Planets move on *elliptical* orbits with the Sun at one focus.

2. A string attached between the Sun and a planet sweeps out equal areas in equal time.

3. The square of a planet's period is proportional to the cube of its semi-major axis.

Remarkably with only a catalog of planetary sky positions, Kepler had made three profound realizations. The first law implied that the heliocentric model of the solar system was correct.[4] The second law was a consequence of the conservation of angular momentum, but at the time the principle was not yet known. Finally, it is possible to derive the third law using Newton's theory of gravity, but that theory would not come for another half-century.

[4] Ironically, Brahe was staunchly opposed to the idea of a heliocentric solar system, and pushed his own geocentric model.

In the 1660s Newton developed calculus and the theory of gravity, providing the mathematical machinery to not just *describe* planetary motion, but to truly understand it (Newton, 1687). While Einstein's eventual discovery of General Relativity succeeded in explaining some anomalies where Newtonian gravity breaks down (such as the rate of precession of Mercury's longitude of pericenter), Newtonian gravity is such a powerful theory that it provides all of the necessary physics for every dynamical study in this thesis.

1.3.2 Solar System Formation

In the past few decades, observations of protoplanetary disks (Lay et al., 1994; Andrews et al., 2018) have confirmed the long-standing theory (Laplace, 1796; Safronov, 1972) that planetary systems are formed when when molecular clouds undergo gravitational collapse into a protostar surrounded by a disk of gas and dust. Our own solar system formed 4.56 billion years ago as a circumstellar disk approximately 30–100 au in radius with the Sun at its center.[5] Shortly after the formation of the protoplanetary disk, the process of planetesimal formation began. While the exact process of planetesimal formation is not yet completely understood, it is likely to be driven by instabilities in the gas and dust of the circumstellar disk, causing dust to gravitationally collapse, and undergo hierarchical growth (Morbidelli et al., 2022). The current leading theory of planetesimal formation, called the streaming instability, uses numerical simulations to show that turbulent shear in protoplanetary disks can create filaments of overdensity, facilitating the rapid growth of planetesimals (Youdin & Goodman, 2005; Abod et al., 2019); it is noteworthy that some streaming instability models match observations in our own solar system rather well, though many caveats still apply (see Kavelaars et al. 2021 and Chapter 7 of this thesis). Once the planetesimals formed, they began to accrete more mass either by encountering other overdensity filaments, or by colliding with other planetesimals (Bottke et al., 2023). Gas giants such as Jupiter and Saturn (and to a somewhat lesser extent ice giants like Uranus and Neptune) must have accreted their $\sim 10 M_{\oplus}$ cores quickly—within the first few million years of the Sun's formation

[5] The size of our protoplanetary disk is rather typical, compared to the expected size of 10–100 au (Shu et al., 1987).

while there was still enough gaseous material around for them to acquire their gas envelopes. The terrestrial planets, on the other hand, did not have the same temporal restriction; they were able to slowly acquire mass via pebble accretion over the first $\sim$ 100 Myr of the solar system's lifetime (see Nesvorný 2018 and the references therein).

After the Sun, the planets, and the planetesimals formed, the solar system continued to evolve gravitationally. Tsiganis et al. (2005) used numerical simulations to show that it is not necessary for the planets to have formed *in situ*, but rather they can move around. The so-called Nice Model simulations began with the solar system in a more compact configuration than we see today with the outermost planet at about 20 au (rather than 30 au), amid a disk of planetesimals. Over millions of years in the simulations, Neptune migrated outward to its present orbit by scattering planetesimals inward. Further simulations have shown that Neptune's migration would have created a delicate dynamical structure in the Kuiper Belt, leading observers and theorists to use the observed Kuiper Belt populations to deduce details of Neptune's migration (Malhotra, 1993, 1995; Volk & Malhotra, 2019; Nesvorný, 2015; Nesvorný, 2020). In perhaps the most relevant result to this work from the original Nice Model simulations, Morbidelli et al. (2005) showed that the planetesimals that were scattered inward during planetary migration could have populated Jupiter's Trojan clouds. This result implies that if Neptune's migration really did happen, the KBOs and the Jupiter Trojans would have formed at similar places in the solar system. Notwithstanding the fact that they have experienced different environments for the past 4 billion years, the two populations should share some properties. Inspired by such predictions, NASA's Lucy mission is currently *en route* to investigate 7 Jupiter Trojans up close (Levison et al., 2021).

1.4 Observations Drive Theory (and vice-versa)

In the study of our solar system, observation and theory have always been intimately connected. In some ways, the interplay resembles detective work—as Fred Adams once told me, it's like playing "CSI Solar System". Observations of solar system bodies enable theorists to develop and

test theories, and in turn those theories make predictions that enable observers to discover new objects. The interplay between observation and theory is integral to the Part I of this thesis, in which I use a combination of analytic and numerical techniques to explore some of the following problems.

Unseen Perturbers

Perhaps the best example of theory enabling observation is the discovery of Neptune. In the 1800s, astronomers carefully tracking Uranus realized that its ephemeris was very slightly inconsistent with predictions calculated using the known solar system bodies. Independently, Le Verrier (1846) and Adams (1846) posited an eighth planet to explain the anomalies, and provided predictions of its orbit. With the predictions in hand, astronomers discovered Neptune within the year (Galle, 1846). Since the discovery of Neptune, there has been a long history of astronomers using apparent anomalies in the orbits of known bodies to predict the existence of an unseen perturber, but thus far all such efforts have been unsuccessful (for an excellent overview, see Batygin et al. (2019)).

In the past decade, observations of a group of KBOs on extremely long, elliptical orbits that are clustered in physical space have have sparked a new hypothesis for a yet-undiscovered planet called Planet Nine (Batygin & Brown, 2016; Batygin et al., 2019). I explore the problem in detail in Chapter 3.

Missing Populations

Sometimes non-detections of solar system bodies are just as interesting as detections. For example, in spite of dedicated searches, no long-term stable Earth Trojans have been discovered (Cambioni et al., 2018; Markwardt et al., 2020). Numerical simulations show that there is ample parameter space for dynamically stable Earth Trojans to exist, so their apparent absence poses a mystery. Inspired by this mystery, I present a novel mechanism for destabilizing Earth Trojans in Chapter 6.

Exotic Orbits

There is little that is more exciting to a theorist than an observer reporting the discovery of a body on a strange orbit. For example the object 2015 BP_{519} (Becker et al., 2018), a highly eccentric KBO with an inclination of 54°, poses a dynamical mystery as it is not easily produced with current solar system models. Interestingly, objects like 2015 BP_{519} are quite readily produced in simulations that include Planet Nine, but if it turns out that Planet Nine is a false alarm, the object will still require some other explanation. Another class of object that poses a dynamical mystery is the so-called Sednoids (Brown et al., 2004), with perihelion distances beyond 60 au. Again, the bodies are easily produced with distant sources of excess angular momentum such as Planet Nine or stellar flybys, but perhaps there is another explanation that has simply not yet been discovered or thought of.

Interstellar objects are among the most exotic orbits ever detected in our solar system. With the first two detections by Meech et al. (2017) and Jewitt & Luu (2019) (both of which were on unbound orbits, and passed through our solar system), interest around interstellar objects quickly grew and theorists worked to understand their place in the solar system. In Chapters 4 and 5 I present work combining analytic and numerical techniques to explore whether there may be any interstellar objects currently residing in our solar system.

1.5 A New Era of Discovery

It is clear that technological innovation has always driven our ability to discover solar system bodies. Now that we have entered an era defined by big data, high-performance computing, and artificial intelligence (AI), we have an opportunity to once again revolutionize the way we discover solar system bodies. Part II of this thesis focuses on the development of the techniques required to take full advantage of the technological leaps of this new era.

1.5.1 AI-Assisted Discovery of Minor Bodies

It is possible to find objects fainter than the single-image threshold by using a technique called shift-and-stack, in which one uses software to co-add images along the trajectory of a moving object. The general idea of shift-and-stack has been around for about 20 years (see Bernstein et al. 2004 and the references therein), but it has only recently become technologically possible at the scale of a survey. Previously there were two problems making large shift-and-stack surveys infeasible.

1. Searching a set of images for all possible orbits requires an enormous amount of computation.

2. When doing shift-and-stack searches in anything but the highest quality images, false positive detections quickly become unmanageable for humans to parse.

With today's computing clusters and powerful AI computer vision techniques such as convolutional neural networks, both of these hurdles can now be overcome.

The Dark Energy Camera Ecliptic Exploration Project (DEEP) is the first shift-and-stack survey to span multiple oppositions, covering more than 140 square degrees of sky over a 4-year period from 2019–2023. When DEEP is completed, it will demarcate another technological leap on Figure 1.1, and will double the number of known KBOs.

In Chapter 7 I use the DEEP data to develop techniques for AI-assisted discovery of minor planets, and use the resulting detections to measure the absolute magnitude distribution of the Cold Classical Kuiper Belt. In Chapter 8 I revisit the DEEP data to do a shift-and-stack search for Jupiter Trojans, necessitating 2 orders of magnitude more computation than a search for KBOs. In Chapter 9 I apply the techniques developed in Chapter 7 to search for a final flyby target for NASA's NH spacecraft, and develop a technique for discovering bodies that are too faint to be discovered in a single night of data. Finally in Chapter 10 I use everything I have learned while doing the work in this thesis to propose a reformulation of the problem of minor planet detection altogether.

1.5.2 A Word on the Limits of Computers

I have a particularly vivid memory of a conversation I had with Fred Adams in 2019 that shaped the way I think about science. While he and David Gerdes were teaching me for an independent study in celestial mechanics, we were discussing the disturbing function, which is an extremely powerful way of analytically predicting the motion of a body on a Keplerian orbit, and is infamous for necessitating the use of algebraic expressions that go on for pages. Baffled by the use of the seemingly antiquated technique, I asked why we bother learning such techniques at all anymore, when numerical simulations are capable of providing more accurate predictions of a body's motion, in less time and with less effort than analytic approaches. He explained to me that using a numerical approach to get a "correct" prediction for a body's motion is great when all you care about is knowing where the body will be at any given time. Numerical simulations can help to glean some insight into the dynamics of a body, or even a population of bodies. But often, the most profound insights into a body's dynamics are in the equations. Without equations, we can use numerical simulations to make some inferences and generalizations about the factors driving a body's dynamics. But when we can write down the equations, we can often pinpoint which terms are contributing to a body's dynamics, resulting in a better understanding of the dynamics of not just one object, but of all objects. The real power in dynamical analyses comes from combining analytic and numerical methods.

The idea stuck with me, and gave me a real appreciation for using a combination of numerical, analytic, and semi-analytic techniques in my work. I find this lesson to be especially pertinent as science enters the age of big data, high-performance computing, and AI.

In late 2022, AI-generated images exploded in popularity with free software such as `DALL·E 2` and `Stable Diffusion` (Rombach et al., 2021). Images produced by the software were showing up in all corners of the internet, and the results seemed like magic. You could tell the software what you wanted to see, and within seconds it would produce an image beyond the artistic ability of most people (myself included). I was intrigued, so I downloaded the software to run `Stable Diffusion` locally, and set out to explore the algorithm by mass-producing images with the goal

of creating an AI art-of-the-day calendar as a gift for my brother.

By the time I had generated a few hundred images, it became clear to me that while the software was very cool, it was not, in fact, magic. In Figure 1.4, I show a pair of images that contributed to my disillusionment. The image on the left resulted from the prompt, *painting of robots playing baseball*, and the image on the right resulted from the prompt, *painting of a seal balancing a ball on its nose*. It just so happened that the images shared a common keyword, *ball*, and a random seed. The image on the left looks pretty good—it is clearly a painting of robots holding baseball bats, standing on a baseball field. The image on the right, however, makes hardly any sense at all. There is some sort of a figure I could be convinced is seal (ignoring the fact that it has three eyes), but it is decidedly not balancing a ball on its nose. To make matters worse, there is a rogue baseball bat on the right side of the image. Embracing my inner astronomer, I examined the two images in a digital blink comparator, and noticed that the baseball bats are exactly the same shape, and in exactly the same location in the image.

Figure 1.4: Two AI-generated images showing similar features. (Left) AI-generated image from the prompt, *painting of robots playing baseball*. (Right) AI-generated image from the prompt, *painting of a seal balancing a ball on its nose*.

All of this is to say that AI algorithms have no capacity for imagination, or to think for themselves. **AI is not magic**. Instead, it is entirely up to the humans who develop and use AI

tools to supply the imagination. We must be extremely careful in our development of AI tools and in how we apply them to science. They are not, and hopefully never will be, a substitute for human thought and creativity.

Chapter 2

`spacerocks`: Software for Solar System Science

> Give me six hours to chop down a tree, and I will spend the first four sharpening the axe.
>
> — Abraham Lincoln

At the beginning of my PhD, I found myself unsatisfied with the general landscape of available general-purpose solar system software. The existing packages were all either slow, inaccurate, unintuitive to use, or difficult to install. It has always been vitally important for me to have my *mise en place* prior to any major undertaking, and I believe that stable, well-tested software is critical for the rapid advancement of science. So in early 2020 when I found myself with a plethora of free time, I jumped at the opportunity to write my own software package, `spacerocks`. Along the way I gained a deeper understanding of solar system dynamics and software development, and created software that is fast, accurate, and (hopefully) easy to use for undergraduates and tenured professors alike.

I describe a few use cases of the software in the remainder of this chapter, but it is really intended to be a Swiss Army Knife for developing ideas about solar system dynamics and observations. I have used it to prototype ideas and carry out calculations for *every* chapter in this thesis. The project is still under active development, and I am in the process of updating its documentation and developing "cheat sheets" to enhance the user experience (see, for example, https://matplotlib.org/cheatsheets/). Nevertheless, the project has reached a point of relative stability, and it has been such an indispensable tool for me that it deserves its own chapter

in this thesis.

2.1 Introduction

The precise calculation of solar system ephemerides is a sensitive problem, depending on details as minute as individual leap seconds and an observer's precise location on Earth. Notwithstanding their difficulties, such calculations are important for targeted followup, and are becoming increasingly common in efforts to model solar system surveys in order to understand survey biases. While services such as JPL Horizons and the Minor Planet Center (MPC) provide ephemerides calculations for known objects, it is not possible (or at least it is rather convoluted) for the general public to use these services for unknown solar system objects. This means these services cannot be used to study real objects until observations have been reported to the MPC, and they cannot be used at all to study survey efficiency with synthetic solar system models.

`spacerocks` is a Python package designed to fill this gap. It is a general tool that offers high-level abstractions for coordinate transformations, orbit propagation, and the computation of ephemerides for arbitrary solar system objects. It combines the practicality and development speed of Python with the performance of `C++` by using `ctypes` to wrap custom routines. Where applicable, the `C++` routines are multi-threaded using `OpenMP`. The package documentation and tutorials are permanently hosted at `https://github.com/kjnapier/spacerocks`. In Section 2.2 I give an overview of the code. In Section 2.3 I demonstrate its use with some example problems. I conclude in Section 2.4 with a discussion of the ongoing development of `spacerocks`, and how it fits into the current environment of solar system software.

2.2 Overview

`spacerocks` is written in the object oriented programming paradigm with a design philosophy that all operations should read like English. The primary class for describing a solar system body is called `SpaceRock`. It is used to instantiate objects, and it contains methods for propagating them to different epochs and computing their ephemerides. There are many ways to instantiate

SpaceRock objects, but at minimum a user must specify either a state vector or an appropriate set of 6 Keplerian elements, a coordinate frame (e.g. J2000, ecliptic J2000, etc.), and an origin (e.g., the Sun, the solar system barycenter, etc.). Alternatively, the user can invoke the instantiation method SpaceRock.from_horizons to pull the elements of known objects from the JPL Horizons database or SpaceRock.from_mpc to pull various minor planet center catalogs. Objects on bound ($e < 1$) and unbound ($e > 1$) orbits can be handled simultaneously.[1] If the user wishes to propagate the objects to past or future epochs, or to compute ephemerides, the epoch of the osculating elements must also be specified. One can give the objects a constant absolute magnitude (H), or specify an arbitrarily complicated rotation curve defined as a Python function.

When a SpaceRock object is fully defined by any 6 appropriate elements, a coordinate frame, an origin, and an epoch, any other orbital quantity may be computed as an attribute. Where applicable, attributes are lazily evaluated, meaning they are only computed when they are explicitly called for. After an attribute has been evaluated, the result is cached so that the computation is performed only once. For convenience, dates are stored as astropy (Astropy Collaboration et al., 2013) Time objects, distances are stored as astropy Distance objects, and angles are stored as astropy Angle objects. Most importantly, the SpaceRock object is fully vectorized, meaning that it can handle multiple solar system bodies simultaneously.

2.2.1 Units

Incorrect units are among the most common bugs in solar system calculations. spacerocks attempts to prevent such bugs and eliminate ambiguity by using a Units class to specify the units for input keyword arguments. The Units class currently allows users to specify units for angles, distance, speed, time scale, time format, physical radius, density, mass, and angular separation.

[1] Parabolic orbits are not currently supported, but will be in future versions.

2.2.2 Orbit Propagation

The SpaceRock object contains a method called propagate, which uses Rebound's IAS15 integrator (Rein & Spiegel, 2015) to numerically propagate bodies forward and backward in time. The bodies in the SpaceRock object can be instantiated with asynchronized epochs, as the method will automatically perform synchronization by integrating all objects to a common epoch before beginning the main integration. The objects can be propagated to any epoch, but they must be instantiated within the date range covered by the ephemeris. The perturbers' initial states are computed using spiceypy (Annex et al., 2020) to parse JPL ephemeris files (e.g., Park et al. 2021), which are automatically downloaded the first time the package is imported. However, I note that it is possible using the SpiceKernel class for a user to load in any ephemeris file they choose. The propagate method allows the user to specify a list of epochs to propagate the objects to, as well as a solar system model. In order of increasing complexity, the built-in models currently include the Sun, the Sun and the giant planets, the Sun and all planets, and the full JPL Horizons perturber model (including the Sun, all of the planets, the Moon, Pluto, and the 16 most massive asteroids). After propagating the objects, spacerocks returns a SpaceRock object containing the state vector of each test particle at each epoch, a SpaceRock object containing the state vector of each perturber at each epoch, and a snapshot of the rebound simulation.

The SpaceRock object also contains an analytic_propagate method that uses the time-linearity of the mean anomaly to propagate an object along its two-body ellipse. This method is much faster, albeit less accurate, than numerical propagation. It is useful for propagating Kuiper Belt Objects, as the ephemeris residuals between the numerical and analytic methods typically remain smaller than 1″ for at least a few years (as long as the orbits are propagated in the barycentric frame).

2.2.3 Ephemeris Calculation

The SpaceRock object contains an observe method for computing bodies' ephemerides as observed from an arbitrary location in the solar system defined as a different SpaceRock object.

For convenience, the method has a built-in shortcut allowing the user to use the keyword `obscode` to specify the observer as any observatory code available in the MPC[2], or the keyword `spiceid` to specify the observer as any body for which SPICE data is available (such as the Hubble Space Telescope, the New Horizons Spacecraft, or the James Webb Space Telescope).

The `observe` method computes the observed body's coordinates in the observer-centric J2000 reference frame at the epoch of observation including an iterative procedure to correct for light travel time, and returns an `Ephemerides` object. Using an `Ephemerides` object, one can compute a multitude of quantities such as equatorial coordinates, ecliptic coordinates, apparent rate of motion, and apparent magnitude. If applicable, the apparent magnitudes are corrected for rotation curves and light travel time.

2.2.4 Coordinate Transformations

It is often necessary to transform orbital elements between the barycentric and heliocentric reference frames (for instance, converting the MPC's heliocentric elements to barycentric). However, doing the transformation from scratch is rather involved because one needs to compute an object's state vector, and then either add or subtract the Sun's state vector at the epoch of the elements. The `SpaceRock` class contains the methods `to_helio` and `to_bary` so the user can seamlessly make these transformations at any epoch covered by the ephemeris. There is a more general method called `change_origin` which allows the user to change the origin of the coordinate system to any body for which SPICE data is available.

The `SpaceRock` class also provides a method called `change_frame` which enables the rotation of the coordinate axes between the J2000, ecliptic J2000, Galactic (Murray, 1989), and invariable (Souami & Souchay, 2012) frames.

[2]The elevations have been computed using the WGS84 model, but slightly more accurate elevations can be obtained by using values from a service such as JPL Horizons or Google Earth.

2.3 Worked Examples

In this section I highlight proof-of-fidelity examples with `spacerocks`, and show how to use it for some common applications.

2.3.1 Comparison to JPL Horizons

JPL Horizons is the gold standard for ephemeris calculation, and any alternative should be able to match it with reasonable accuracy over long time spans. Here I demonstrate the accuracy of `spacerocks` against JPL Horizons by computing the ephemerides of the main-belt asteroid, 24601 Valjean (1971 UW). I chose a main-belt asteroid for this demonstration because the closer an object is to the observer, the more difficult it is to accurately compute its ephemerides. I instantiated Valjean as a `SpaceRock` object using Keplerian elements obtained from JPL Horizons at the epoch of March 30, 2022 in barycentric dynamical time (TDB). I then propagated the object in the presence of the full set of JPL Horizons perturbers from January 1, 2000 to January 1, 2050 in universal time coordinated (UT) in intervals of 30 days, and computed the ephemerides as observed from Cerro Tololo Inter-American Observatory. Finally I pulled the predicted ephemerides from JPL Horizons for the same epochs. In Figure 2.1 I show the separation in sky position between the `spacerocks` predictions and the JPL Horizons predictions as a function of time. The predictions match with excellent accuracy over the 50 year time span tested, never differing by more than 1″.[3] We find even better agreement when accounting for General Relativity using `reboundx` (Tamayo et al., 2020); the residuals remain smaller than 0.2″. The increased accuracy from including General Relativistic effects comes at the expense of speed, as the propagation takes approximately 20 times longer than with pure Newtonian gravity. While `spacerocks` should provide sufficient accuracy for objects as close as the main belt asteroids, I note that objects that make close approaches to massive bodies, such as Apophis or planetary moons, require a great deal of care in their dynamical modeling, and as such its results will not

[3]The code used in this example is available in a tutorial notebook at https://github.com/kjnapier/spacerocks/blob/master/notebooks/Valjean.ipynb.

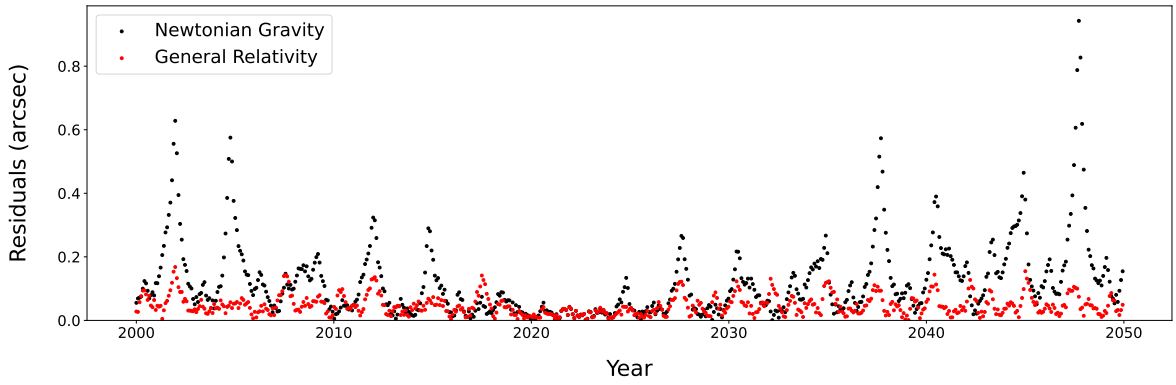


Figure 2.1: Difference in the predicted sky position of asteroid 24601 Valjean (1971 UW) between `spacerocks` and JPL Horizons over a 50-year time span from January 1, 2000 until January 1, 2050. The residuals never exceed (and are generally much smaller than) 1″.

agree as well with JPL Horizons.

2.3.2 Generation of Synthetic Resonant Bodies

Mean-motion resonances between planets and small bodies are among the most fundamental emergent phenomena in solar system dynamics. The phenomenon occurs when the ratio of the small body's period and the planet's period can be expressed as a rational number, which sometimes results in the geometric impossibility of a close encounter between the small body and the planet. The rationality of period ratios is a necessary, but insufficient condition for a small body to be in a mean motion resonance with a planet. To describe a complete sufficient condition for mean motion resonance, we must use the resonant angle, given by

$$\varphi = c_1\lambda_{outer} + c_2\lambda_{inner} + c_3\varpi_{outer} + c_4\varpi_{inner} + c_5\Omega_{outer} + c_6\Omega_{inner} \tag{2.1}$$

where λ refers to a body's mean longitude, ϖ refers to a body's longitude of pericenter, and Ω refers to a body's longitude of ascending node (Murray & Dermott, 2000). Variables with the subscript *outer* refer to the outer of the two bodies, and variables with the subscript *inner* refer to the inner of the two bodies. Finally c_i are integers satisfying the criterion $\sum_i c_i = 0$. When there exists some combination of prefactors c_i such that φ does not precess through the entire interval $[0, 2\pi]$, the small body and the planet are considered to be in a mean motion resonance.

Understanding bodies in mean motion resonance is critical to understanding the dynamics of our solar system. It is often useful to be able to generate a synthetic population of bodies in a given mean motion resonance. Here I show how to use `spacerocks` to generate a population of bodies in the most fundamental mean-motion resonance (one-to-one) with Jupiter—the so-called Jupiter Trojans.[4] In the case of one-to-one resonances, the resonant angle simplifies to $\varphi_{1:1} = \lambda_{\text{Jupiter}} - \lambda_{\text{trojan}}$.

I begin by generating a population of candidate Jupiter Trojans by sampling orbital elements from a set of distributions that I have chosen somewhat arbitrarily for the sake of demonstration. By definition the period of a Jupiter Trojan is approximately equal to the period of Jupiter, meaning that the bodies' semi-major axes are also approximately equal. To ensure that I have amply covered the parameter space, I generate a uniform distribution with $a \in [5.0, 5.4]$ au. I then randomly sample each of the remaining five parameters uniformly with $e \in [0, 0.4]$, $i \in [0°, 60°]$, $\varpi \in [0°, 360°]$, $\Omega \in [0°, 360°]$, $M \in [0°, 360°]$.[5]

After I have generated an initial population of candidate Jupiter Trojans, I use `spacerocks` to instantiate a simulation including the Sun, all eight planets, and the test particles. Since I do not need to worry about close encounters between the test particles and perturbers, I use a symplectic integrator with a step size of 10 days. I then integrate the system for 10^5 years, writing the state vector of each body in the simulation to memory every 10 years. Every time I save the state of the system to memory, I check each test particle's semi-major axis, and remove it from the simulation if it does not satisfy the condition 5.0 au $< a <$ 5.4 au. Then at the end of the integration, I check each of the surviving test particles for resonance with Jupiter by evaluating the boundedness of $\varphi_{1:1}$ using a method similar to that described by Khain et al. (2020a). I ran simulations with a total of 10^5 test particles in batches of 1000, yielding 5540 Jupiter Trojans

[4] Jupiter Trojans are among the most plentifully-known dynamical classes in the solar system, with an inventory approaching 10^4. Their physical properties may encode information about our solar system's formation and dynamical evolution, making the population particularly interesting to study. In fact, NASA's ongoing Lucy mission will send a spacecraft to the Jupiter Trojans to more closely study seven of the bodies Levison et al. (2021).

[5] In principle I could have used some physical priors on the orbits to shrink the parameter space, but the procedure is fast enough for the present purposes that such an optimization was unnecessary.

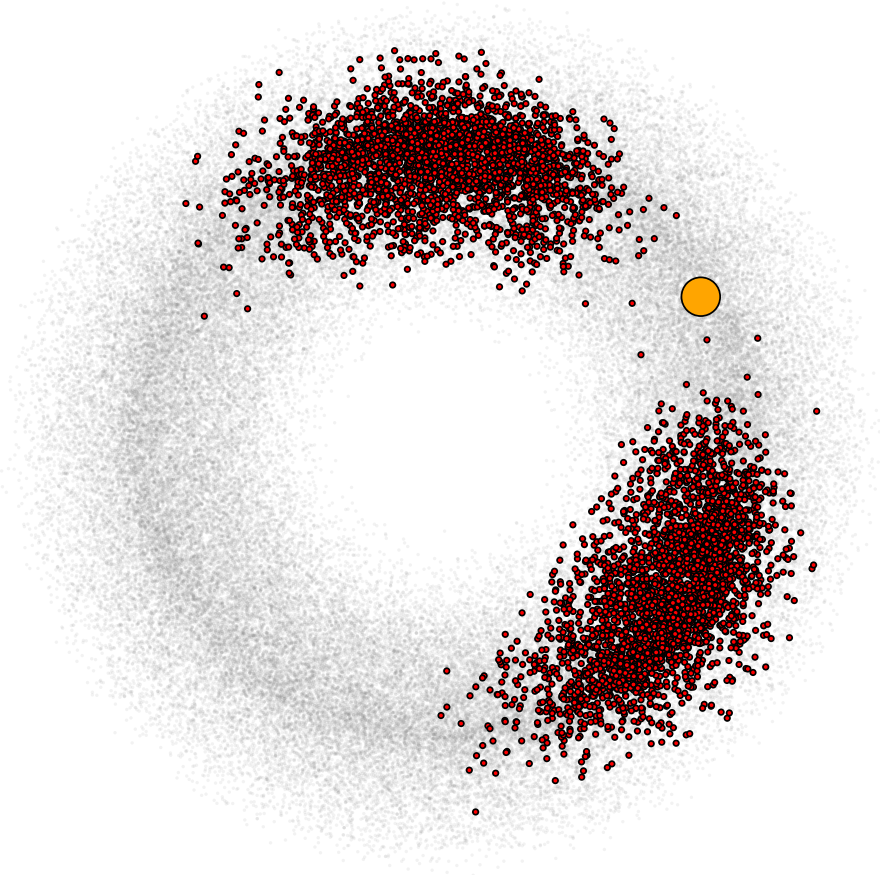


Figure 2.2: Initial positions of simulated Jupiter Trojans projected into the ecliptic plane. The bodies that remained in resonance over the 10^5-year integration are represented by the red dots, while the bodies that dropped out of resonance are represented by the faint grey dots. The large orange circle represents Jupiter.

stable for 10^5 years. Figure 2.2 shows the initial position of each test particle projected into the ecliptic plane. As expected, the resonant test particles occupy the stable Lagrange points leading and trailing Jupiter by 60°.

The code used for this example is available in a tutorial notebook at https://github.com/

kjnapier/spacerocks/blob/master/notebooks/generate-trojans.ipynb. It can easily be modified to generate bodies in arbitrary mean motion resonances with any planet.

2.3.3 Hamiltonian Computation

spacerocks is also useful for analytic work, and I demonstrate an instance here. Some Hamiltonians cannot be computed completely analytically, but rather must be integrated numerically over a test particle's orbit. One such example is the von Ziepel-Lidov-Kozai (vZLK) Hamiltonian (von Zeipel, 1910; Lidov, 1962; Kozai, 1962; Ito & Ohtsuka, 2019). This Hamiltonian can be used to predict the secular evolution of bodies that avoid close encounters with perturbers. I skip over the details as they are beyond the scope of this work, but I refer the reader to Thomas & Morbidelli (1996) for a complete formal treatment. The non-constant part of this doubly-averaged Hamiltonian can be written as

$$\mathcal{K} \sim \sum_i m_i \int_0^{2\pi} \frac{4}{\sqrt{r^2 + a_i^2}} \sqrt{1 - \frac{\xi}{2}} K(\xi)\, dM \tag{2.2}$$

where

$$\xi = \frac{4 a_i \sqrt{x^2 + y^2}}{r^2 + a_i^2 + 2 a_i \sqrt{x^2 + y^2}}. \tag{2.3}$$

In this expression the subscript i runs through the perturbing planets, the variables with no subscript refer to the test particle, m represents mass, x and y give the test particle's position in the plane of the planets, r is its heliocentric distance, and $K(\xi)$ is the complete elliptic integral of the first kind. Because the planets are assumed to be on co-planar circular orbits, the problem is rotationally invariant about the longitude of ascending node (Ω), and we can set constant $\Omega = 0$. The test particle's orbit is completely defined by its semi-major axis (a), eccentricity (e), inclination (i), argument of pericenter (ω), $\Omega = 0$, and its mean anomaly. Consider that one wants to compute the level curves of this Hamiltonian in the space of pericenter distance (q) vs. ω. In this case, assuming the test particle avoids close encounters and is not in a mean motion resonance, the dynamics can be well-approximated by holding fixed a and $H = \sqrt{1 - e^2} \cos i$. In

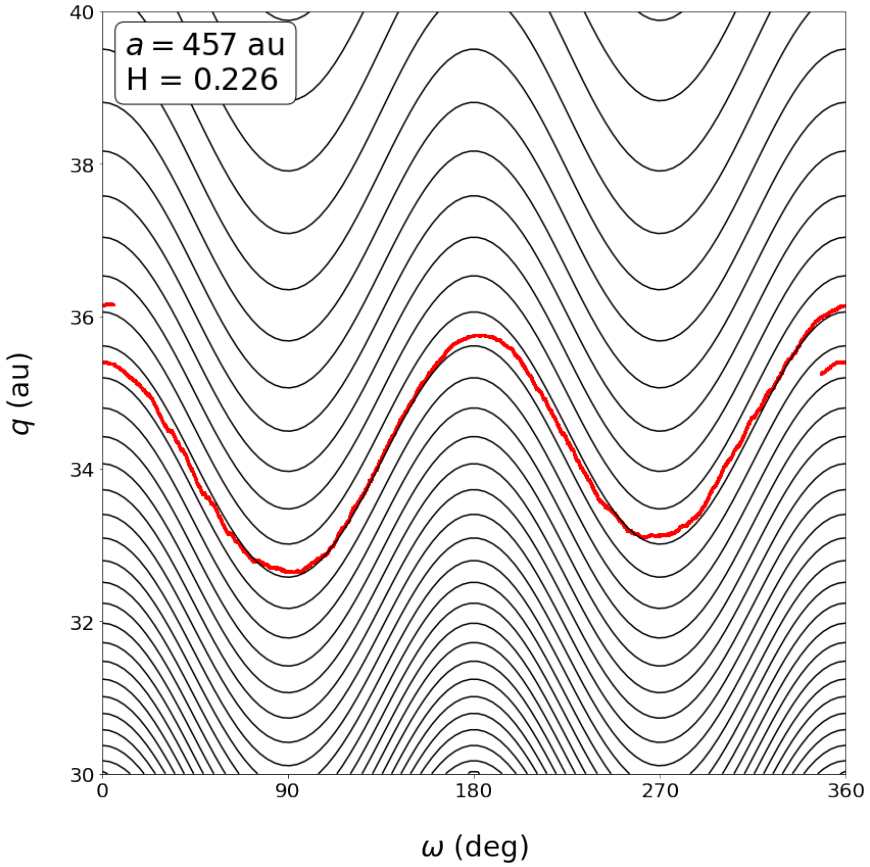


Figure 2.3: (Black) Hamiltonian level curves in the space of q vs ω for a fixed semi-major axis of 449 au and a fixed reduced Kozai action of 0.227. (Red) Trajectory of the TNO 2015 BP_{519} in this phase space during a 4 Gyr numerical integration with the Sun and the four giant planets.

this example, we compute the Hamiltonian for the trans-Neptunian object (TNO) 2015 BP_{519}, which has $a = 449$ au and $H = 0.227$.[6]

In Figure 2.3 we plot the level curves of the Hamiltonian for the TNO 2015 BP_{519}, along with its trajectory in the parameter space for a 4 Gyr integration with the Sun and the four giant planets. This figure shows that the secular evolution of 2015 BP_{519} is well-described by the vZLK Hamiltonian.

[6]The code used in this example is available in a tutorial notebook at https://github.com/kjnapier/spacerocks/blob/master/notebooks/kozai-hamiltonian.ipynb.

2.4 Discussion and Ongoing Development

In this chapter I have introduced `spacerocks`, a `Python` package for solar system dynamics and observations. `spacerocks` offers high-level abstractions for many operations on solar system objects such as coordinate transformations, N-body integration, and the calculation of ephemerides. While services such as JPL Horizons and the MPC can be useful for computing ephemerides of known solar system objects, they can not be readily used to study newly discovered or synthetic objects. Furthermore their APIs are not conducive to studying large populations of known objects.

`spacerocks` offers a solution to these problems by providing a more convenient API that can be used to study arbitrary solar system bodies. Its ephemerides predictions are consistent with those of JPL Horizons over time spans of hundreds of years for all solar system objects, with the exception of those on trajectories that make close approaches to massive bodies. `spacerocks` is useful both as a tool for general daily work, and in applications such as survey simulation and synthetic object generation. The code can be downloaded at https://github.com/kjnapier/spacerocks or installed via `pip`.

I am currently developing new features including an interface with the MPC, a suite of orbit fitting tools, and an implementation of the ASSIST N-body integrator, which uses a more sophisticated gravity model than my current implementation, and matches JPL Horizons to machine precision (Holman et al., 2023). Finally, I am developing a `spacerocks` package in `Rust` for performance-critical applications. The project, hosted at https://github.com/kjnapier/spacerocks-rust, is still in its infancy, but has already become a major component of my workflow.

Part I

Dynamics

Chapter 3

No Evidence for Orbital Clustering in the Extreme Trans-Neptunian Objects

It is a capital mistake to theorize before you have all the evidence. It biases the judgment.

— Sherlock Holmes, *A Study in Scarlet*

This chapter is published as Napier et al. (2021c).

3.1 Introduction

The apparent clustering in longitude of perihelion ϖ and ascending node Ω of solar system bodies known as extreme trans-Neptunian objects (ETNOs) motivated the hypothesis that the solar system contains a 5-10 Earth-mass planet (Planet X/Planet 9) at 400-800 times Earth's distance from the sun (Trujillo & Sheppard, 2014; Batygin & Brown, 2016; Batygin et al., 2019). Some have proposed even more exotic sources of the apparent clustering, such as gravitational perturbations from a primordial black hole captured into orbit around the sun (Scholtz & Unwin, 2020).

While there is no universally accepted definition for the ETNOs, recent literature has emphasized objects with semi-major axis $a \gtrsim 230$ au and perihelion $q > 30$ au. Because ETNOs follow highly elliptical orbits, and their brightness decreases like $1/r^4$, they are almost always discovered within a few decades of perihelion. Moreover, telescopic surveys observe a limited area of the sky,

at particular times of year, to a limited depth. These effects result in significant selection bias. The six ETNOs considered in the Batygin & Brown (2016) (BB16) analysis were discovered in an assortment of surveys with unknown or unpublished selection functions, making it difficult to establish that the observed angular clustering was indeed of physical origin.

More recent surveys have carefully characterized their selection functions and applied these tools to small samples of new ETNOs. The Outer Solar System Origins Survey (OSSOS, Bannister et al. (2016)) analyzed the bias present in the discovery of 8 objects they detected with $a > 150$ au and $q > 30$ au (Shankman et al., 2017). They found that their detected objects were consistent with a uniform underlying population in ϖ and Ω. Bernardinelli et al. (2020a) analyzed samples of 3-7 variously-defined ETNOs discovered by the Dark Energy Survey (DES, DES Collaboration (2005); Dark Energy Survey Collaboration et al. (2016)), and also found the data consistent with angular isotropy.

Brown & Batygin (2019) (BB19) attempted to reverse-engineer the survey bias in the entire then-known population of 14 ETNOs using a sampling method (Brown, 2017) on all TNOs known to the Minor Planet Center. In contrast to the individual survey-level analyses described above, BB19 concluded that the observed clustering is highly likely to be a physical effect, and they argued that the best explanation remains a massive distant planet.

While no single survey has discovered enough ETNOs to reach a statistically compelling conclusion, a stronger statement becomes possible when data from multiple surveys are combined. According to the criteria above, there are 14 ETNOs (Table 3.1) detected by three independent surveys with characterized selection functions, all published since BB16. Using the published pointing history, depth, and TNO tracking selections for DES (5 objects) (Khain et al., 2020b; Bernardinelli et al., 2020b), OSSOS (5 objects) (Bannister et al., 2018), and the survey of Sheppard & Trujillo (2016) (ST) (4 objects), we calculate the joint probability that these objects are consistent with the null hypothesis: an underlying population distributed uniformly in the longitudes ϖ and Ω. If the purported clustering is indeed a physical effect, we would expect it to remain consistent with the data in this larger, independent sample when selection functions are

Object	a (au)	e	i (°)	q (au)	ω(°)	Ω (°)	H (mag)	Survey
2015 BP_{519}	448.8	0.92	54.1	35.2	348.1	135.2	4.4	DES
2013 SL_{102}	314.3	0.88	6.5	38.1	265.5	94.7	7.1	DES
2013 RA_{109}	462.4	0.90	12.4	46.0	263.0	104.8	6.2	DES
2014 WB_{556}	289.1	0.85	24.2	42.5	234.6	114.9	7.3	DES
2016 SG_{58}	233.0	0.85	13.2	35.1	296.3	119.0	7.5	DES
2013 SY_{99}	733.1	0.93	4.2	50.1	32.2	29.5	6.7	OSSOS
2015 RX_{245}	426.4	0.89	12.1	45.7	65.1	8.6	6.2	OSSOS
2015 GT_{50}	311.4	0.88	8.8	38.5	129.0	46.1	8.5	OSSOS
2015 KG_{163}	679.7	0.94	14.0	40.5	32.1	219.1	8.2	OSSOS
uo5m93	283.0	0.86	6.8	39.5	43.3	165.9	8.8	OSSOS
2013 FT_{28}	295.4	0.85	17.4	43.4	40.7	217.7	6.7	ST
2014 SR_{349}	296.6	0.84	18.0	47.7	341.2	34.9	6.7	ST
2015 TG_{387}	1101.3	0.94	11.7	65.1	118.0	301.0	5.5	ST
2014 FE_{72}	1559.5	0.98	20.6	36.2	133.9	336.8	6.1	ST
2012 VP_{113}*	262.7	0.69	24.1	80.5	293.8	90.8	4.0	ST
2013 RF_{98}*	363.6	0.90	29.5	36.1	311.8	67.6	8.7	DES SN

Table 3.1: Barycentric orbital elements of the ETNOs used in our analysis. All reported values are at the epoch JD 2459000.5 (except for uo5m93 whose elements are for the epoch JD 2457163.82647). DES SN indicates discovery in the DES supernova fields. *Note that in order to maintain an independent sample from BB16 we do not include 2013 RF_{98} or 2012 VP_{113} in our main analysis. We discuss their effects separately.

modeled.

3.2 Methods

The three surveys we consider have very different designs and scientific goals, and consequently quite different ETNO selection functions. This is readily apparent from their survey footprints, shown in Figure 3.1. DES, which was on-sky between 2012 and 2019, used the Dark Energy Camera (DECam, Flaugher et al. (2015)) on the 4-meter Blanco telescope at CTIO to carry out an extragalactic survey designed to measure cosmological parameters. It consisted of two interwoven surveys. In the 30 sq. deg. supernova survey, ten separate fields were visited approximately weekly in the *griz* bands during the six months per year that DES was in operation. In the 5000 sq.

deg. wide survey, each field was imaged a total of 10 times at a sparse temporal cadence in each of the *grizY* bands over the duration of the survey. The wide survey reached a limiting *r*-band magnitude of ≈ 23.5. DES had limited near-ecliptic coverage centered near ecliptic longitude of 0, and a large off-ecliptic footprint that made it particularly sensitive to high-inclination objects. For our main analysis, we consider only the ETNOs detected in the DES wide survey, and treat the supernova fields separately. The OSSOS survey (2013-2017), by contrast, was optimized to detect and track TNOs in eight ~ 20 sq. deg. blocks distributed along the ecliptic. This survey used the 3.6-meter Canada-France-Hawaii Telescope and reached a limiting *r*-band magnitude of 24.1–25.2. Finally, the ST survey (2007-2015) used the Blanco, Subaru, Large Binocular, and Magellan telescopes to cover 1080 sq. deg. at an average distance of 13 deg. from the ecliptic to a depth of approximately $VR \sim 25$. This survey aimed to detect the most distant objects: ETNOs and Inner Oort-Cloud (IOC) objects such as Sedna. Therefore, only those candidates with an estimated heliocentric distance greater than 50 au were selected for followup and tracking.

The most complete way to account for survey bias in the discovery of the solar system objects is to use a survey simulator (Petit et al., 2011; Lawler et al., 2018). In essence, a survey simulator simulates detections of a model population of solar system bodies by using a survey's pointing history, depth, and tracking criteria. This allows for the computation of a survey's selection function for a given population, which enables us to account for bias, and therefore understand the true underlying populations. While it gives a reasonable approximation, the technique employed in BB19 cannot fully substitute for actually simulating each survey to calculate its selection function. Since the known ETNOs were discovered by a variety of surveys, the task of developing an appropriate simulator is nontrivial. Our simulator (`FastSSim`) is highly parametric, requiring only the few pieces of information common among all well-characterized surveys: pointing history, limiting magnitudes, and followup criteria.[1] The basic flow of the simulator is as follows:

[1]The tools for the `FastSSim` algorithm have now been compiled into an open-source Python package `SpaceRocks`. It is under active development at https://github.com/kjnapier/spacerocks.

1. Map a survey's published pointing history to a HEALPix[2] grid, as in Figure 3.1.

2. Generate a distribution of fake objects at a single epoch.

3. Calculate the objects' HEALPix pixels and apparent magnitudes.

4. Determine which fake objects fall in a survey's footprint.

5. Make cuts according to the survey's limiting magnitudes and followup criteria.

Note that this simulation method makes several approximations. We compute the sky coordinates of our objects at objects at a single epoch, we use a single color and limiting magnitude for each survey field, we do not consider CCD-level detections (so we do not account for complications such as chip gaps), and we employ a step-function detection criterion (so we do not model survey cadence or linking efficiency). We use a single HEALPix pixel for each survey pointing. We have chosen the pixel scales for each telescope as follows: Blanco uses NSIDE of 64 (except for the DES supernova fields, for which we use NSIDE of 1024), and the Magellan, Large Binocular, and Subaru telescopes use NSIDE of 128. These assumptions ignore the time history of the surveys, as well as apparent motion of the objects. `FastSSim` works well for this application because the objects move slowly, the telescopes have large fields of view, and the sensitivity does not have much spatial variation.

We acquired the non-DES survey pointings and limiting magnitudes from Sheppard & Trujillo (2016), Sheppard et al. (2019), and Bannister et al. (2018). We choose each HEALPix pixel size to most closely match the field of view of the telescope used. This does not allow for a perfect mapping between pointings and pixels, but it turns out to be sufficient for our needs. In fact, we find that `FastSSim` performs remarkably well in cross-checks against both our full chip-level DES simulator (Hamilton, 2019), and the OSSOS survey simulator described in Petit et al. (2011) (see Figure 3.2). While `FastSSim` misses some of the fine details of the selection functions, the small sample of ETNOs and the approximate nature of this analysis make such

[2]It is not important that we used a HEALPix mapping. We could have used any mapping onto the sphere.

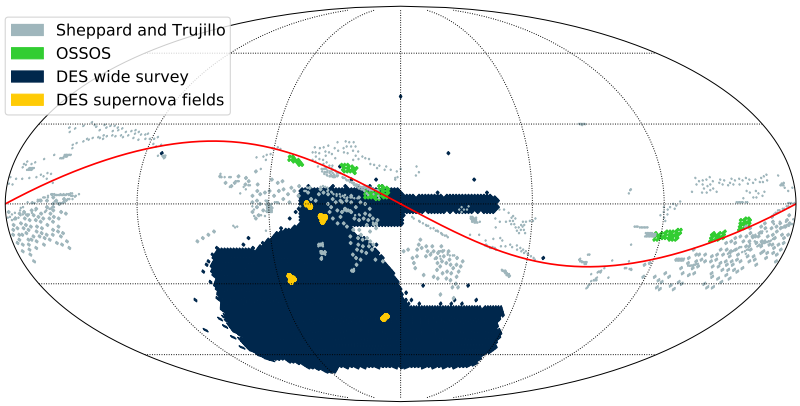


Figure 3.1: A HEALPix mapping of the currently-released sky coverage of the three major TNO surveys of this generation. The surveys by OSSOS and Sheppard and Trujillo hug the ecliptic plane (plotted in red), while DES, designed as a cosmological survey, has a much more expansive footprint.

fine details unimportant to our overall conclusion. Given the success of these cross-checks, we are confident in extending its use to characterize the survey of Sheppard and Trujillo.

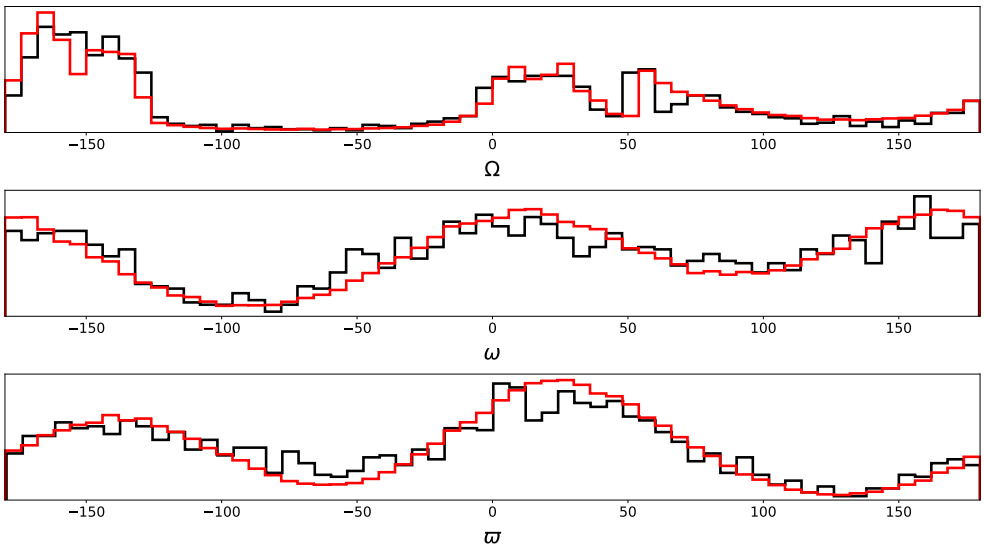


Figure 3.2: OSSOS selection functions for 1,615 detections from the nominal population described in Section 3 in the angles ω, Ω, and ϖ, as calculated by `FastSSim` (red) and the OSSOS/CFEPS survey simulator (black). The distributions have the same general shape.

To simulate the surveys we randomly generate ETNOs in accordance with a nominal scattered disk model (specified in §3.3) until each survey has accumulated 10^5 detections, to allow for

high-resolution characterization of a survey's sensitivity in ϖ and Ω. This typically requires the generation of approximately 10^{10} fakes, so our set of simulated objects spans the parameter space of the ETNOs. We consider an object to be detected if it is in one of the survey's HEALPix pixels, is brighter than the pixel's limiting magnitude, and has a perihelion distance $q \geq 30$ au. For the survey of Sheppard and Trujillo, we satisfy a tracking criterion specified in Sheppard et al. (2016) by requiring an object to have a heliocentric distance of at least 50 au at the time of detection.

As a quantitative example of the effectiveness of `FastSSim`, Figure 3.2 shows a comparison with the CFEPS/OSSOS simulator. For this test each simulator uses the population model defined in Section 3. Using Kuiper's test, we find that the distributions of the 1,615 data points calculated by `FastSSim` are statistically indistinguishable from those computed using the CFEPS/OSSOS survey simulator. Thus in order to distinguish the two simulators one would need $> 1,615$ ETNOs—well above the quantity discovered by OSSOS. We achieve similar results in quantitative comparisons against the distributions computed using the DES survey simulators (see Figure 5.1 in Hamilton (2019) and Figure 2 in Bernardinelli et al. (2020a)).

In Figure 3.3 we plot a Gaussian kernel density estimate of each simulator's detections in the (x, y) and (p, q) spaces (x, y, p, and q are defined in Section 4). Since the distributions appear to be in good agreement we believe that our selection functions, which are derived directly from the surveys' pointing histories, depths, and TNO tracking strategies, more faithfully model their respective surveys than than those inferred indirectly in BB19 (see their Figure 4). Furthermore, the p-values we calculate using the CFEPS/OSSOS survey simulator do not significantly differ from those we calculate using `FastSSim`.

3.3 Scattered Disk Model

To test the dependence of our analysis on the choice of scattered disk model, we simulated models with various distributions of semi-major axis (a), eccentricity (e), inclination (i), and absolute magnitude (H), while keeping the orbital angles Ω and ω (and thus ϖ) uniform from

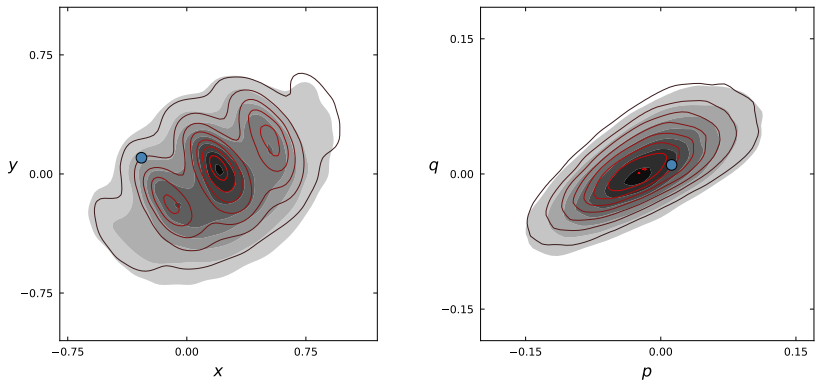


Figure 3.3: Comparison of simulated OSSOS selection functions in *xypq* space using `FastSSim` and the OSSOS survey simulator. The black contours (which follow a linear scale) are Gaussian kernel density estimates of 10^6 iterations of sampling 5 points from the OSSOS biases calculated using the CFEPS/OSSOS survey simulator (because OSSOS discovered 5 ETNOs), and plotting the mean (x, y) and (p, q) positions (see Section 4 for definitions of x, y, p, and q). The red contours represent the same statistic, calculated using `FastSSim`. The blue points are the mean positions of the 5 ETNOs discovered by OSSOS. Our simulator reproduces the results of the OSSOS simulator with better fidelity than the heuristic method of BB19 (see their Figure 4).

0° to 360°. We tested manifold permutations with the parameter distributions: $N(a) \propto a^{\zeta}$ with $a \in [230\,\text{au}, 1600\,\text{au}]$ and $\zeta \in [0.5, 1.0]$, uniform $i \in [0°, 60°]$, Brown distribution i (Brown (2001)) with a variety of widths ranging from 5° to 25°, and $N(H) \propto 10^{H\zeta}$ with $H \in [4, 10]$ and $\zeta \in [0.6, 0.9]$.

We found that our conclusions were not significantly affected by the variation of the model parameters. Our results are also robust to changes in pericenter distribution (see the Appendix for the distribution of orbital elements for populations with $q > 30$, 35, and 38 au). (Shankman et al., 2017) found similar resilience to changes in scattered disk model. Noting the weak dependence of the outcome of our simulations on the choice of model, we proceed using the following scattered disk model:

- a follows a single power-law distribution such that $N(a) \propto a^{0.7}$, where $a \in [230\,\text{au}, 1600\,\text{au}]$
- e is distributed uniformly $\in [0.69, 0.999]$
- i follows a Brown distribution such that $N(i) \propto \sin(i) \exp\left[-\frac{(i-\mu_i)^2}{2\sigma_i^2}\right]$ with $\mu_i = 0°$ and $\sigma_i = 15°$

– H follows a single power-law distribution such that $N(H) \propto 10^{0.8H}$, where $H \in [4, 10]$

– Perihelion distance $q > 30\,\mathrm{au}$

These model parameters produce posteriors in a, e, q, i, and H that appear to be in reasonable agreement with the real ETNO detections by each survey. See the Appendix for histograms of the posteriors in each of these variables, overlaid with a rug plot of each survey's real detections.

3.4 Analysis and Results

Performing a clustering analysis in the variables ϖ and Ω is complicated, as the two are strongly correlated. We proceed by working in the orthogonal $\{x, y, p, q\}$ basis discussed in BB19 (importantly, ϖ and Ω are linearly independent in this basis). Note that these vectors are not normalized, but instead have their lengths modulated by eccentricity and inclination. The coordinates are defined as follows.

$$\Gamma = 1 - \sqrt{1 - e^2} \qquad Z = \sqrt{1 - e^2}\,[1 - \cos(i)] \tag{3.1}$$

$$x = \sqrt{2\Gamma}\cos(\varpi) \qquad y = \sqrt{2\Gamma}\sin(\varpi) \tag{3.2}$$

$$p = \sqrt{2Z}\cos(\Omega) \qquad q = \sqrt{2Z}\sin(\Omega) \tag{3.3}$$

Note that Γ and Z have been scaled by a factor of $\sqrt{GM_\odot a}$ from their traditional forms, since the semi-major axis is not relevant to this argument. Figure 3.4 shows our calculated selection functions in the xy and pq-planes.

For the sake of comparison, we used the method presented in BB19 to test the consistency of each survey's detected ETNOs with its selection function. We first perform 10^6 iterations sampling from our simulated detections a set of objects whose cardinality is equal to that of the set of real ETNOs detected by the given survey. We then take the average $\{x, y, p, q\}$ position of each sample, and use these values to construct a 4-dimensional histogram. We display a Gaussian kernel density estimation of these data in the xy and pq-planes in Figure 3.5.

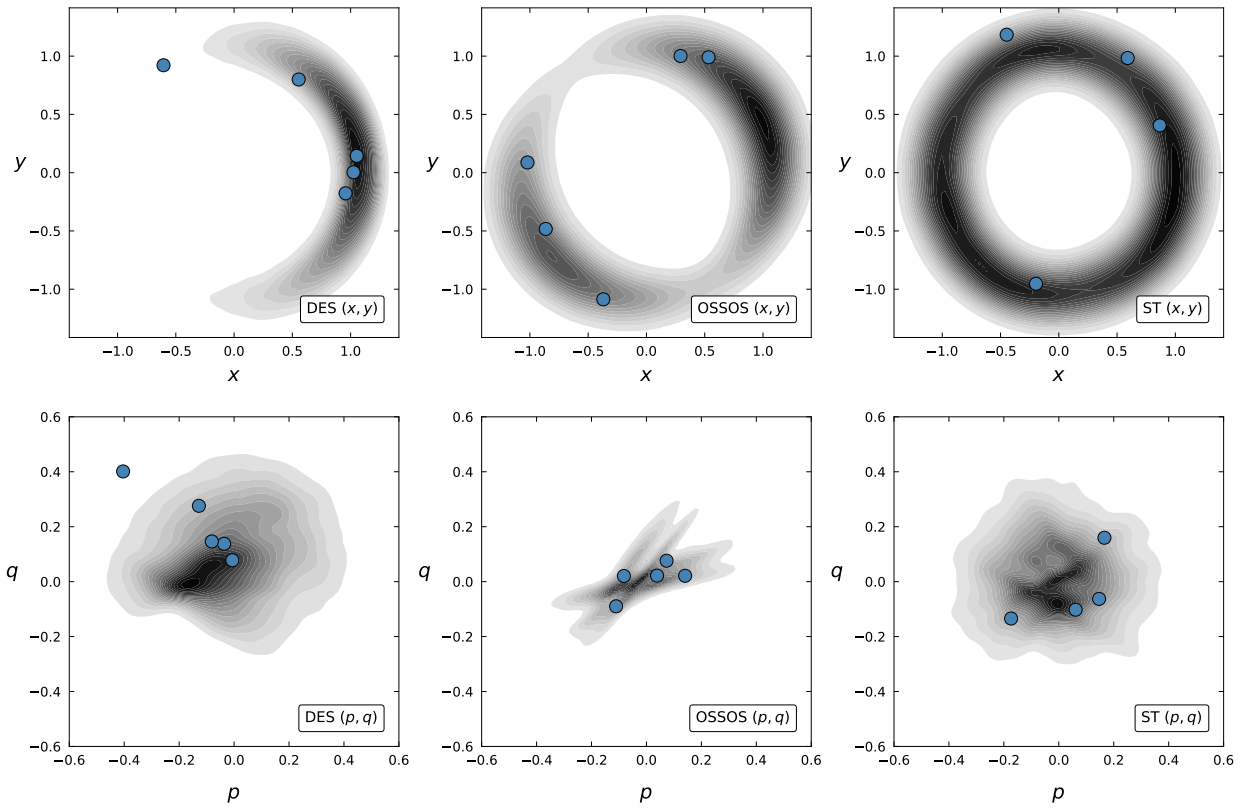


Figure 3.4: Kernel density estimates of each survey's selection function in the canonical xy-space (top row) and pq-space (bottom row). The contours represent simulated detections (the contours scale linearly, and darker contours are more densely populated), while the blue dots represent the ETNOs detected by each survey. The outlier in both DES panels is the object 2015 BP_{519} (Becker et al., 2018).

We perform a Gaussian kernel density estimation on our mean-sampled histograms to obtain a probability distribution function (PDF). Next we draw N samples from our simulated data (where N is the number of ETNOs actually detected by the survey), find the mean $\{x, y, p, q\}$ position, and evaluate our PDF at that position. We repeat this 10^5 times to construct a likelihood function. Next we compute this value for the ETNOs actually discovered by the survey. To calculate the probability of a survey detecting the ETNOs it actually detected (as opposed to some other set of ETNOs), we find the fraction of the 10^5 sample likelihood values that the survey's actual likelihood value exceeds. Rounded to the nearest 1%, this probability for each survey is as follows: $\mathcal{P}_{DES} \sim 0.06$, $\mathcal{P}_{OSSOS} \sim 0.53$, and $\mathcal{P}_{ST} \sim 0.59$.

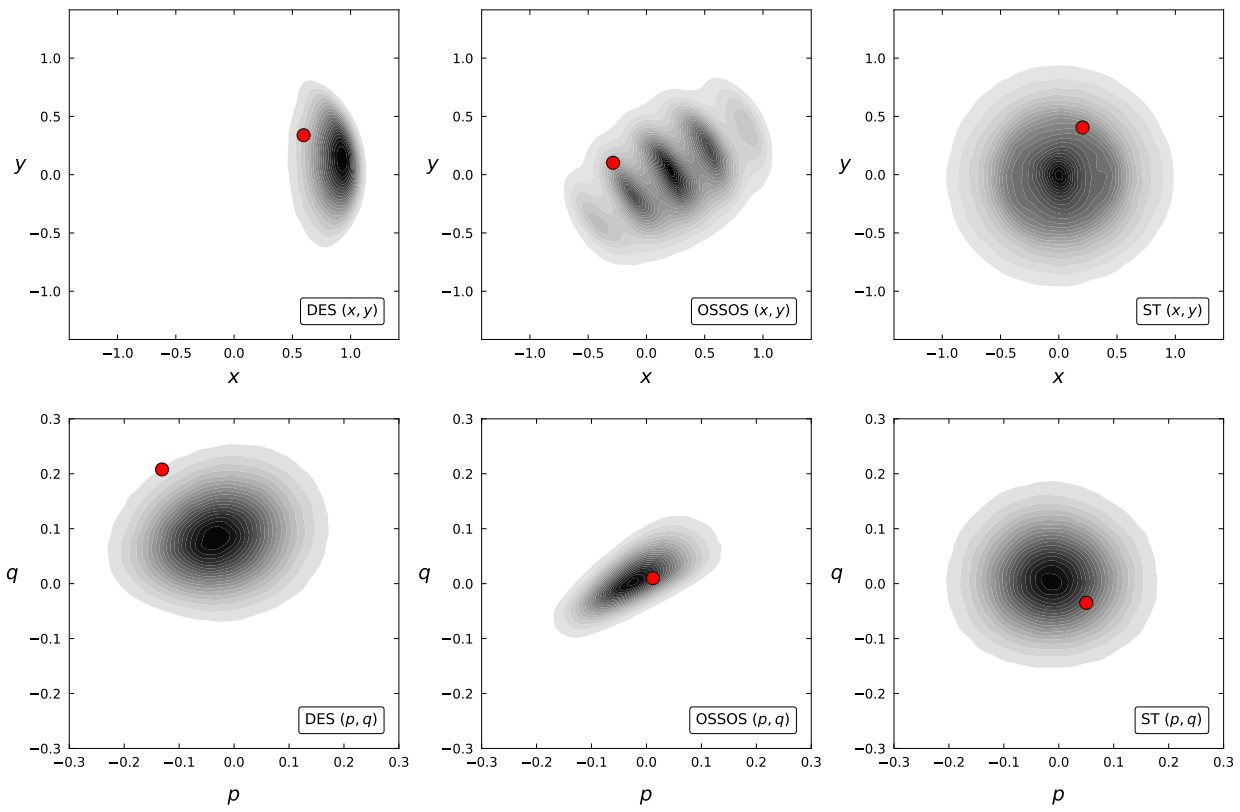


Figure 3.5: Kernel density estimates of the mean (x, y) and (p, q) position of 10^6 samples of ETNOs drawn from the PDFs shown in Figure 3.4. The number of objects in each sample corresponds to the number of ETNOs detected by the given survey. The contours represent the samples (the contours scale linearly, and darker contours are more densely populated), while the red dots represent the mean position of the ETNOs detected by each survey.

The joint probability of N surveys detecting objects with given probabilities (or some less likely set of values) can be calculated as the volume under the surface of constant product of probabilities in the domain of the N-dimensional unit hypercube, given by

$$\mathcal{P}_{joint} = P \sum_{k=0}^{N} (-1)^k \frac{\log(P)^k}{k!} \tag{3.4}$$

where $P \equiv \prod_k \mathcal{P}_k$. In our case, $k \in \{\text{DES}, \text{OSSOS}, \text{ST}\}$. Using Equation 3.4, we calculate the joint probability to be 24%.

With such a small sample size, this work is sensitive to outliers and the definition of the ETNOs itself. The high-inclination object 2015 BP_{519} is among the most dynamically anomalous objects in the solar system (Becker et al., 2018), and we cannot discount the possibility that it is

of a different dynamical origin than the other ETNOs. If we redo our analysis without 2015 BP_{519}, $\mathcal{P}_{DES}$ increases to 84%, and thus $\mathcal{P}_{joint}$ increases to $\sim$ 85%. 2014 FE_{72} has an extremely large semi-major axis—roughly four standard deviations above the mean of the ETNOs considered in this work. Its large semi-major axis carries it deep into the IOC region, where interactions with galactic tides make its secular relationship with a putative Planet X/Planet 9 less certain. If we exclude 2014 FE_{72}, $\mathcal{P}_{ST}$ increases to 88% and thus $\mathcal{P}_{\text{joint}}$ increases to 31%. If we include 2012 VP_{113}, $\mathcal{P}_{ST}$ increases to 60%, and $\mathcal{P}_{\text{joint}}$ remains 24%. We also address the fact that the clustering by a putative Planet X/Planet 9 should be more robust in the sample of ETNOs with $q > 40$ au, since these objects avoid strong perturbations by Neptune. If we restrict our ETNOs to these 8 objects, $\mathcal{P}_{\text{joint}}$ increases to 94%. Finally, we analyze the subset of objects which are either stable or metastable in the presence of the putative Planet X/Planet 9 (Batygin et al., 2019): 2015 TG_{387}, 2013 SY_{99}, 2015 RX_{245}, 2014 SR_{349}, 2012 VP_{113}, 2013 RA_{109}, and 2013 FT_{28}. For this subset $\mathcal{P}_{\text{joint}} = 82\%$.

For the sake of completeness, we also use a more traditional sampling method to determine the significance of the clustering of ETNOs. We begin by performing a Gaussian kernel density estimate on each survey's posterior distributions. We then perform 10^5 iterations in which we randomly draw N points from each survey's posterior distribution (where N is the number of ETNOs detected by the survey) and multiply each of the N probabilities together to calculate a likelihood. Finally, we calculate the same metric for each survey's actual detections and compare the value to the distribution of our samples. As before, the probability for each survey is the fraction of the 10^5 sample likelihood values that the survey's actual likelihood value exceeds. Rounded to the nearest 1%, the probability for each survey is as follows: $\mathcal{P}_{DES} \sim 0.06$, $\mathcal{P}_{OSSOS} \sim 0.41$, and $\mathcal{P}_{ST} \sim 0.43$. The joint probability is thus 17%.

For a more physically intuitive representation of the survey bias, refer to Figure 3.6. Here the radial quantity represents the barycentric distance, and the azimuthal quantity represents true longitude (the true anomaly + ϖ). The edge of the black circle is at 30 au. The white regions represent the combined surveys' sensitivity (brighter regions correspond to higher sensitivity),

weighted by the number of real ETNO detections. The red dots represent the real ETNOs at the epoch of discovery. The observations are in good agreement with the combined selection function, qualitatively confirming the conclusions of our formal statistical analysis performed on canonical variables.

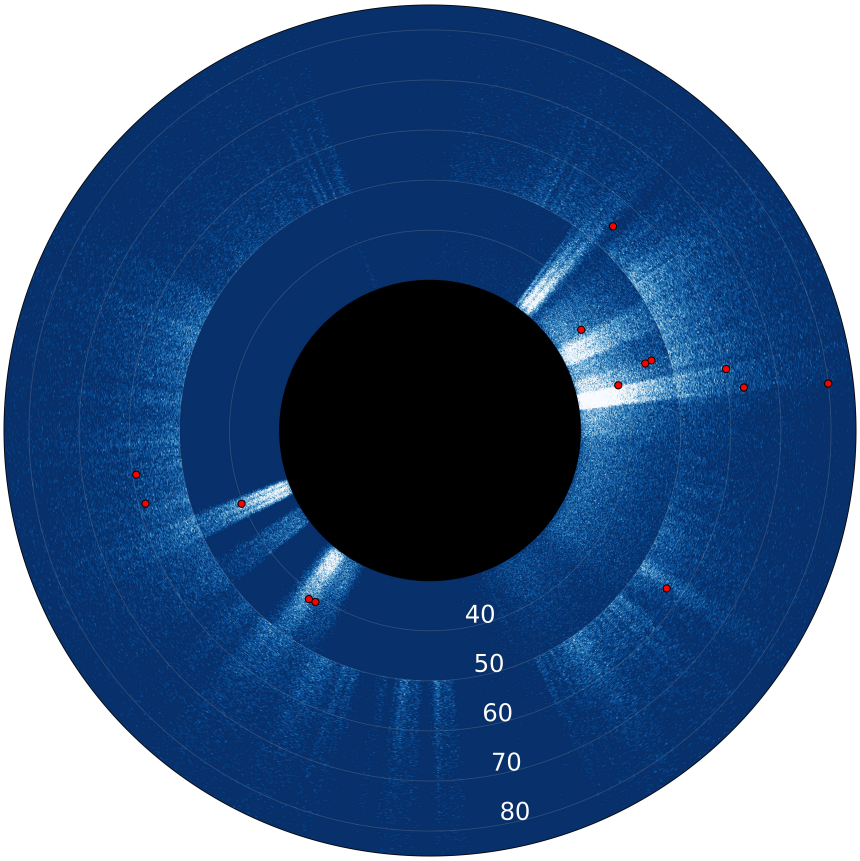


Figure 3.6: Combined ETNO selection function for all three surveys. The radial quantity is the ETNO's barycentric distance, and the azimuthal quantity is true longitude. The edge of the black circle is at 30 au. The white regions represent the combined surveys' sensitivity (brighter regions correspond to higher sensitivity), weighted by the number of real ETNO detections. The red dots represent the real ETNOs at the epoch of discovery. The outer ring is caused by the 50 au tracking criterion imposed by ST.

3.4.1 DES Supernova Fields

The ETNO 2013 RF_{98} was discovered in the deep DES supernova fields (DES SN hereafter). Since the DES SN fields are so small, they suffer from severe selection bias. Additionally, since

their observing cadence and depth ($\sim$24.5 in the r-band) are significantly different than the wide survey, they need to be treated independently. We generated 1,829 simulated detections in the DES SN fields (since the fields are so small, it is computationally prohibitive to generate 10^5 synthetic detections as we do for DES, OSSOS, and ST) from the population model defined in Section 3. We show the posteriors in $\{a, e, i, H, \Omega, \varpi\}$ in Figures 3.A.1, 3.A.2, 3.A.3, 3.A.4, 3.A.5, 3.A.6, and 3.A.7. Figure 3.7 shows a kernel density estimate of the detections in $\{x, y, p, q\}$ space. In all parameters, 2013 RF_{98} appears to be a rather ordinary detection for the DES SN fields.

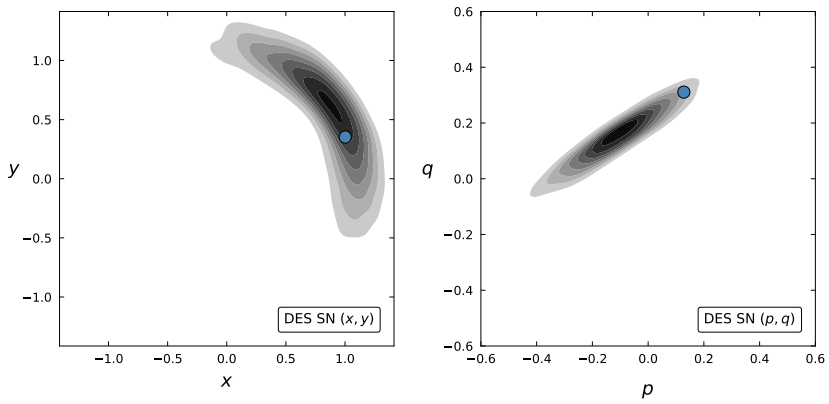


Figure 3.7: Kernel density estimates of the DES SN selection function in the canonical xy-space (left) and pq-space (right). The contours represent simulated detections (the contours scale linearly, and darker contours are more densely populated), while the blue dots represent 2013 RF_{98}.

Since there is only one data point here, we can just numerically integrate to find $\mathcal{P}_{\text{DES SN}} = 0.33$ (i.e. a p-value of 0.33). Treating DES SN as its own survey, we may use Equation 3.4 to calculate the 4-survey joint probability to find $\mathcal{P}_{\text{joint}} = 25\%$.

3.5 Discussion and Conclusions

We use quantified selection bias calculations on all the ETNOs discovered by the three most productive ETNO surveys, each with with a quite different survey strategy and selection function, to test the consistency of the ETNOs with a uniform underlying distribution. Given a joint

probability between 17% and 94% (i.e. a p-value between 0.17 and 0.94), we conclude that the sample of ETNOs from well-characterized surveys is fully consistent with an underlying parent population with uniform distributions in the longitudes ϖ and Ω. Our result differs drastically from the corresponding value in BB19 of 0.2%. Closer inspection sheds some light on the apparent discrepancy. If we examine only the overlapping set of ETNOs used this work and in BB19 (2015 BP_{519}, 2013 RF_{98}, 2013 SY_{99}, 2015 RX_{245}, 2015 GT_{50}, 2015 KG_{163}, 2013 FT_{28}, 2014 SR_{349}, and 2014 FE_{72}), $\mathcal{P}_{\text{joint}}$ drops to < 0.005. This indicates an expected issue: small number statistics are sensitive to fluctuations. For example, when BB19 performed their analysis a small but important set of the ETNOs had not yet been reported to the MPC. As a concrete demonstration of the importance of the omission of a few ETNOs from BB19, consider DES. Of the five ETNOs discovered by the DES wide survey, BB19 included only 2015 BP_{519}. From Figure 3.4 it is clear that this object lands in an extremely low-probability region. This drives down $\mathcal{P}_{\text{joint}}$, and thus gives a satisfactory answer as to why the result of this work differs so significantly from that of BB19.

It is important to note that our work does not explicitly rule out Planet X/Planet 9; its dynamical effects are not yet well enough defined to falsify its existence with current data. This work also does not analyze whether some form of clustering could be consistent with the 14 ETNOs we consider. For example, the ETNOs could happen to be clustered precisely where current surveys have looked. In that case, a survey with coverage orthogonal to the regions shown in Figure 3.6 would find far fewer ETNOs than expected. Various realizations of Planet X/Planet 9 predict clustering of various widths, modalities, and libration amplitudes and frequencies; we do not test for consistency with any of these distributions. Instead, we have shown that given the current set of ETNOs from well-characterized surveys, there is no evidence to rule out the null hypothesis. Increasing the sample of ETNOs with ongoing and future surveys with different selection functions such as the Deep Ecliptic Exploration Project (DEEP) (Trilling et al., 2019) and the Legacy Survey of Space and Time (LSST) at the Vera Rubin Observatory (Schwamb et al., 2018) will allow for more restrictive results. Despite other lines of indirect evidence for

Planet X/Planet 9, in the absence of clear evidence for clustering of the ETNOs the argument becomes much weaker. Future studies should consider other mechanisms capable of giving the outer solar system its observed structure, while preserving a uniform distribution of ETNOs in the longitudes Ω and ϖ.

3.A Supplementary Figures

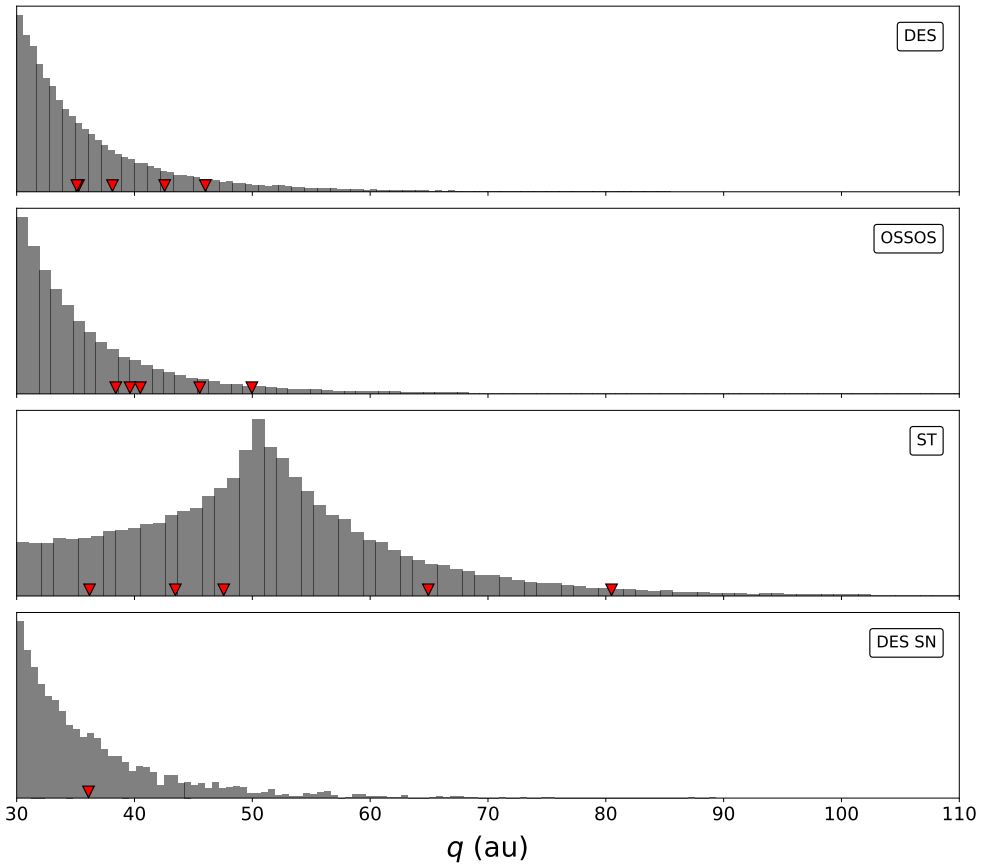


Figure 3.A.1: Posterior pericenter distance distributions of simulated detections. The red triangles are the real ETNO detections by each survey. Note that DES has partially overlapping data points at $q \approx 35$ au.

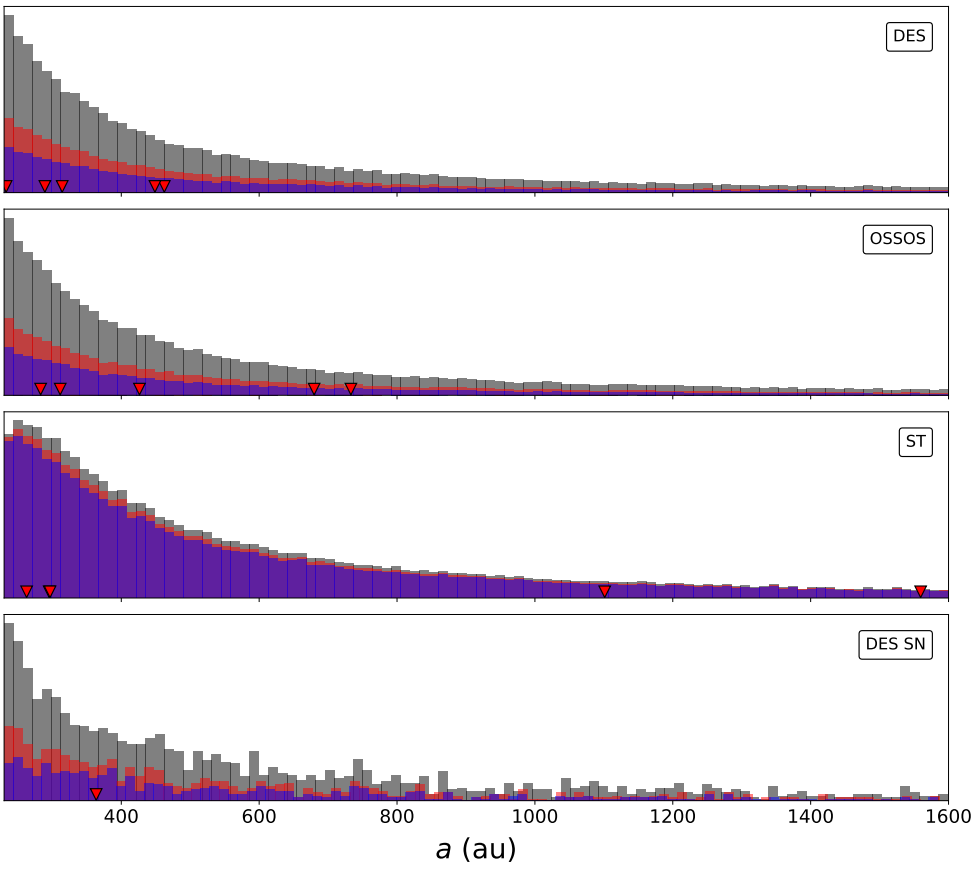


Figure 3.A.2: Posterior semi-major axis distributions of simulated detections. The red triangles are the real ETNO detections by each survey. Note that ST has partially overlapping data points at $a \approx 296$ au. The grey, red, and blue histograms correspond to cuts with $q > 30, 35$ and 38 au, respectively.

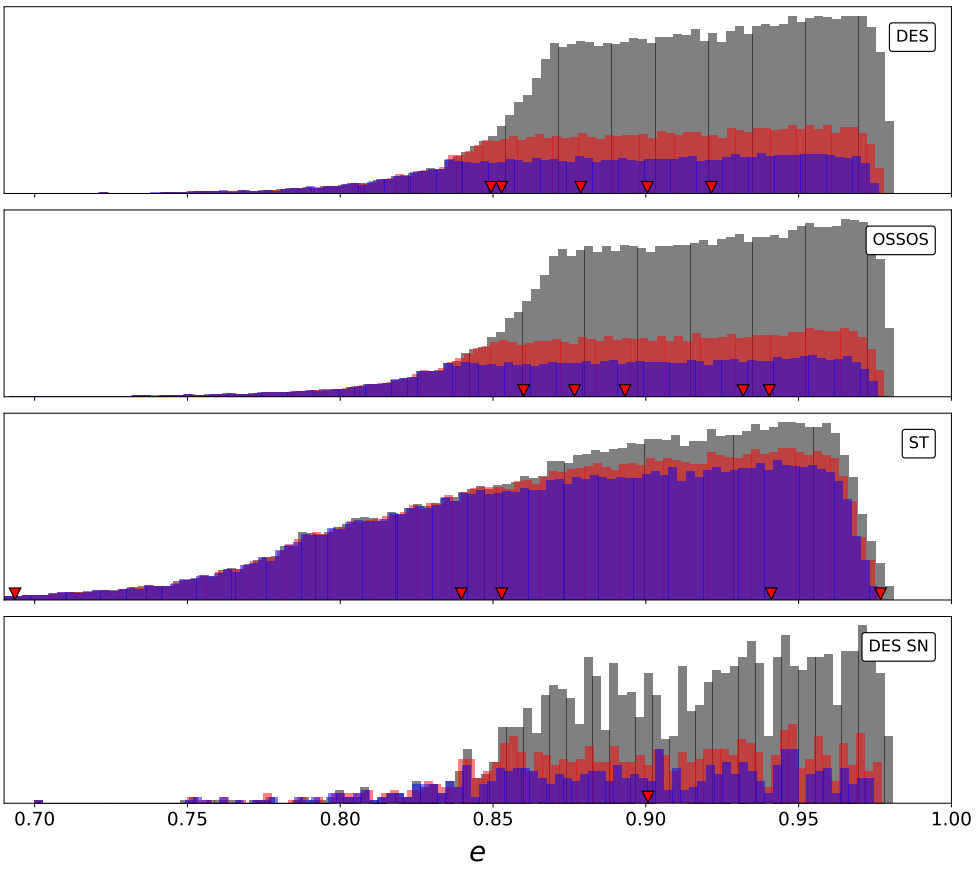


Figure 3.A.3: Posterior eccentricity distributions of simulated detections. The red triangles are the real ETNO detections by each survey. Note that DES has overlapping data points at $e = 0.85$. The grey, red, and blue histograms correspond to cuts with $q > 30, 35$ and 38 au, respectively.

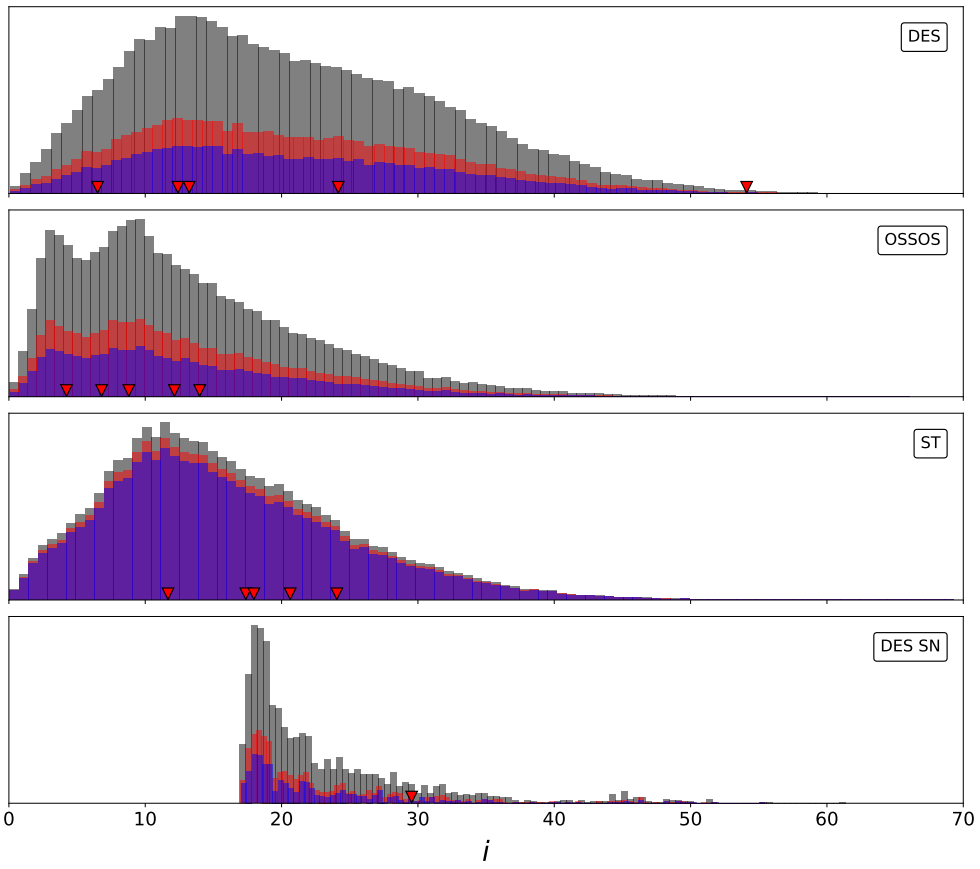


Figure 3.A.4: Posterior inclination distributions of simulated detections. The red triangles represent the real ETNO detections by each survey. The grey, red, and blue histograms correspond to cuts with $q > 30, 35$ and 38 au, respectively.

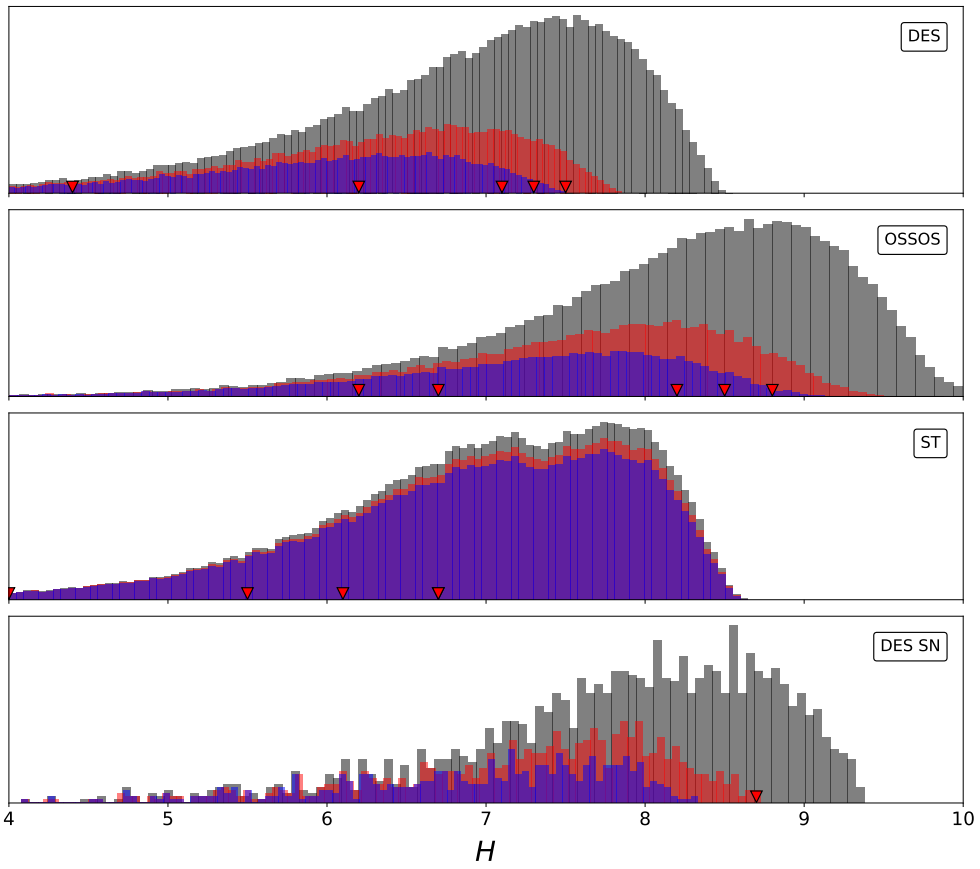


Figure 3.A.5: Posterior absolute magnitude distributions of simulated detections. The red triangles represent the real ETNO detections by each survey. Note that ST has overlapping data points at $H = 6.7$. The grey, red, and blue histograms correspond to cuts with $q > 30, 35$ and 38 au, respectively.

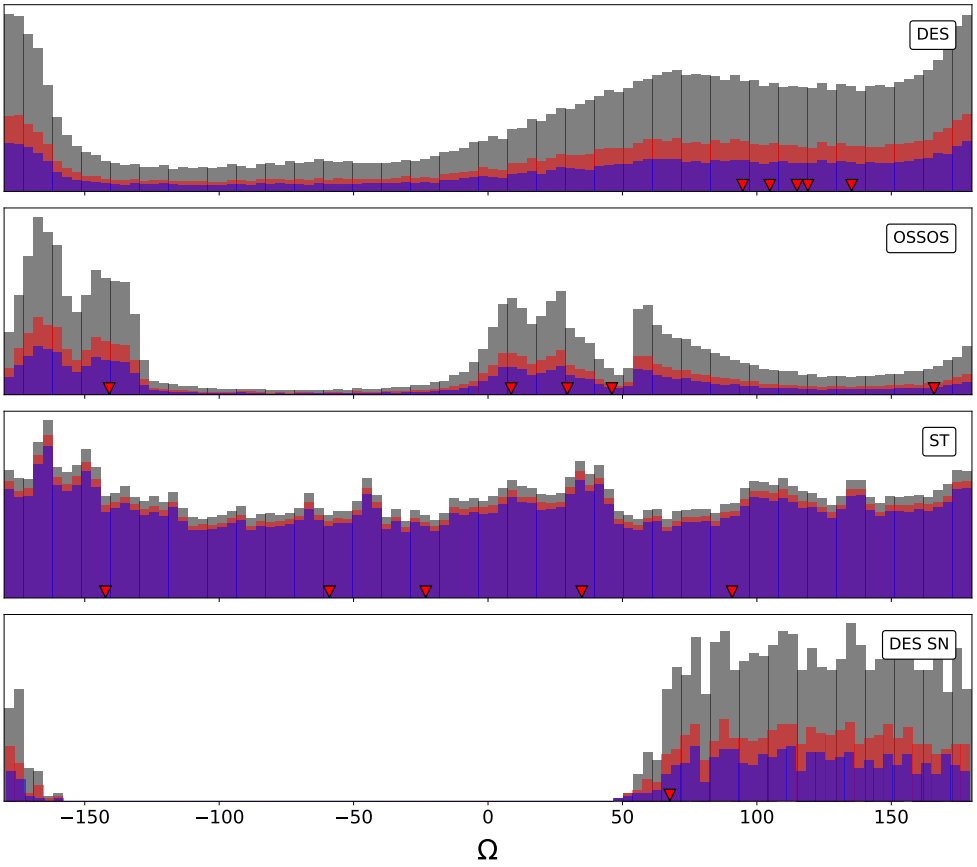


Figure 3.A.6: Posterior longitude of ascending node distributions of simulated detections. The red triangles represent the real ETNO detections by each survey. The grey, red, and blue histograms correspond to cuts with $q > 30, 35$ and 38 au, respectively.

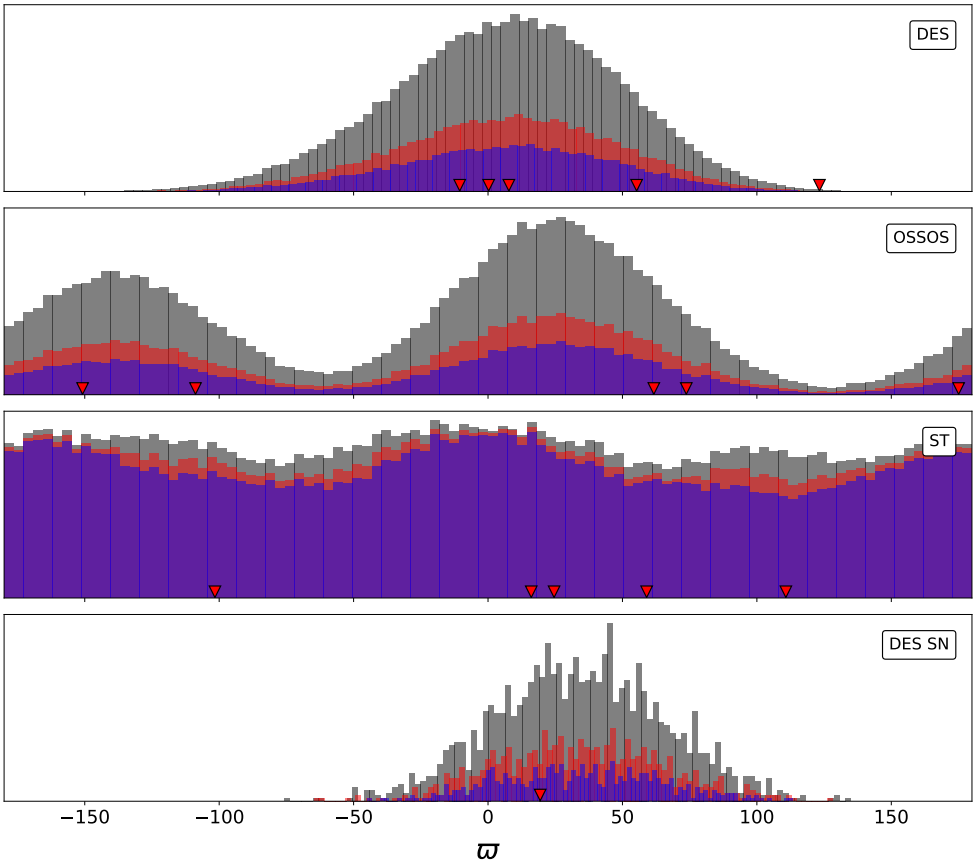


Figure 3.A.7: Posterior longitude of pericenter distributions of simulated detections. The red triangles represent the real ETNO detections by each survey. The grey, red, and blue histograms correspond to cuts with $q > 30, 35$ and 38 au, respectively.

Chapter 4

On the Capture of Interstellar Objects by Our Solar System

> Look, if you had one shot, or one opportunity,
> to seize everything you ever wanted in one moment,
> would you capture it, or just let it slip?
>
> — Eminem, *Lose Yourself*

This chapter is published as Napier et al. (2021a).

4.1 Introduction

The past few years have witnessed the detection of two interstellar bodies passing through the solar system on hyperbolic orbits. The discoveries of the irregular body 'Oumuamua (Meech et al., 2017) and the comet Borisov (Jewitt & Luu, 2019) sparked immediate interest in characterization of these objects and facilitated wide-ranging speculation regarding the possibility that our solar system is more broadly contaminated by minor bodies of extra-solar origin (e.g., Siraj & Loeb 2019; Namouni & Morais 2020). Although no current evidence indicates that any specific objects in the solar system are of extrinsic origin (Morbidelli et al., 2020), the question of whether or not any such objects reside in interplanetary or trans-Neptunian space is of considerable interest. Motivated by these issues, this paper reconsiders the capture of external bodies by our solar system. The calculation of the capture cross sections is the first step in assessing whether or not

the solar system presently contains quasi-permanently trapped interstellar bodies. This treatment also provides constraints on the expected orbits of any such material.

The dynamics of the outer solar system represents one of the oldest problems in theoretical astrophysics. Starting more than two centuries ago, classic studies include the long-term stability of the solar system (Lagrange, 1776; Laplace, 1799-1825), the origin of comets (Laplace, 1846), and orbital anomalies that led to the discovery of Neptune (Le Verrier, 1846; Adams, 1846). Over recent decades, the outer solar system has revealed itself to be increasingly complicated, with the discovery of the Kuiper Belt (Luu & Jewitt, 2002), dozens of dwarf planets (starting with Brown et al. 2004, or perhaps Tombaugh 1946), high-inclination objects (Becker et al., 2018), and aligned extreme trans-Neptunian objects (Sheppard & Trujillo, 2016) that led to the hypothesis of a possible ninth planet (Batygin & Brown, 2016; Batygin et al., 2019). The more recent discovery of interstellar objects (Meech et al., 2017; Jewitt & Luu, 2019) adds to the intrigue. Both the complex orbital architecture of the solar system and the presence of interloping objects motivates this present study. The goal is to determine cross sections for the capture of foreign bodies by the solar system, and to obtain a deeper understanding of the capture process.

The possible capture of interstellar bodies by the solar system also has a long history. The general problem of interacting binaries was considered by Heggie (1975), where the subset of 'resonant' encounters lead to capture. Subsequent studies have carried out numerical explorations of the capture process specifically for our solar system, often considering only the Sun-Jupiter system (see, e.g., Valtonen & Innanen 1982; Valtonen 1983; Siraj & Loeb 2019). Additional studies consider capture for specific scenarios, including capture by compact objects (Pineault & Duquet, 1993), capture of interstellar objects from the field (Lingam & Loeb, 2018; Hands & Dehnen, 2020), the formation of wide binaries (Kouwenhoven et al., 2010), and the possible capture of Planet Nine (Li & Adams, 2016; Mustill et al., 2016). Most of these previous studies calculate the capture rate by sampling a given distribution of encounter speeds between the incoming body and the solar system. These studies generally use the field star velocity distribution, with dispersion $\sim$ 40 km/s (Binney & Tremaine, 2008) or that appropriate for the solar birth

cluster (Portegies Zwart, 2009; Adams, 2010; Pfalzner, 2013; Parker, 2020), where the velocity dispersion is expected to be ~ 1 km/s (Lada & Lada, 2003). Notice, however, that the velocity distribution for rocks (or planets) ejected by solar systems will not generally have a simple Maxwellian form.[1]

The objective of this paper is to extend the aforementioned previous work concerning the capture of interstellar bodies by the solar system. Whereas most studies determine capture rates and cross sections for a given distribution of velocities, this work finds the cross section $\sigma(v_\infty)$ as a function of relative velocity. The results can then be integrated (after the fact) for any distribution of velocities of interest. This approach is much more computationally expensive than previous treatments, but is made possible with current computational capabilities. Specifically, this paper reports the results from $\sim 5 \times 10^8$ fly-by simulations. In addition, we carry out the simulations for solar systems models including all four giant planets. Although earlier work (Heggie, 1975; Pineault & Duquet, 1993) provides analytic estimates for the cross sections, exact forms are not available (primarily due to the lack of an analytic solution to the gravitational three-body problem). We revisit this issue using a different (but equivalent) set of approximations. We then compare the numerical and analytic results for the cross section as a function of velocity, and find good agreement.

4.2 Dynamics of the Rock Capture Process

This section presents an analytic description of the rock capture process. The capture of an incoming body occurs through the time dependence of the gravitational potential of the solar system. In this treatment, we consider the incoming orbit in two regimes. In the outer regime, at large distances, the rock executes a hyperbolic orbit about the center of mass of the solar system. In the inner regime, at closer distances, the rock can enter into the sphere of influence of individual solar system members (e.g., the Sun or Jupiter), and then be described by a hyperbolic orbit around that body. Under favorable conditions, the deflection by the solar system body

[1] As one example, the distribution of speeds for planets ejected from crowded solar systems has the approximate form $dP/dv = 4v/(1+v^2)^3$ (e.g., Moorhead & Adams 2005).

during the close encounter can lead to energy loss and capture in the center of mass frame. This effect is essentially the inverse of the gravitational slingshot mechanism by which satellites are boosted through planetary encounters. Note that by dividing the orbit into two regimes, we are implicitly assuming that 3-body effects are not important.

For the sake of definiteness, we consider only one planet at a time, and work in the limit where the masses of the rock μ, the planet m, and the star M obey the ordering

$$\mu \ll m \ll M \,. \tag{4.1}$$

The incoming orbit of the rock is characterized by its asymptotic speed v_∞ and impact parameter b. For given input variables (v_∞, b), we can define the orbital elements and related physical quantities, including the specific energy and angular momentum,

$$E = \frac{1}{2} v_\infty^2 \qquad \text{and} \qquad J = b v_\infty \,, \tag{4.2}$$

the semi-major axis and eccentricity,

$$|a| = -a = \frac{GM}{v_\infty^2} \qquad \text{and} \qquad e^2 = 1 + b^2/a^2 \,, \tag{4.3}$$

and the perihelion distance

$$r_p = p = a(1-e) = |a|(e-1) \,. \tag{4.4}$$

Note that, by convention, the semi-major axis $a < 0$. To fully characterize the orbit, one must also specify the inclination angle of the incoming trajectory.

It is useful to define the effective cross section for hyperbolic orbits to enter the giant planet region of the solar system. In order for the incoming rock to experience the time-dependence of the gravitational potential, the perihelion r_p must be smaller than the semi-major axis a_p of the planet of interest. This condition implies that the impact parameter b is bounded from above by $b^2 \leq a_p^2 + 2a_p|a|$, where a is the semi-major axis of the incoming orbit. The nominal cross section σ_0 for orbit crossing is thus given by

$$\sigma_0 = \pi \left[a_p^2 + 2a_p|a| \right] \approx 2\pi \frac{GM}{v_\infty^2} a_p \,, \tag{4.5}$$

where the final equality holds for essentially all incoming speeds of interest ($v_\infty^2 < GM/a_p$). The capture cross section will be some fraction of the fiducial cross section (4.5).[2]

4.2.1 Gravitational Slingshot Mechanism for Close Encounters

For the inner regime defined above, we consider close encounters of incoming rocks on initially hyperbolic trajectories with much larger target bodies (either the Sun or one of the giant planets). We can define the coordinate system so that the rock approaches the target body from the $+\hat{x}$ direction and from the $+\hat{y}$ direction, where the angle of the incoming trajectory is θ in the center of mass frame (see the first panel in Figure 4.1). The rock initially has speed v_∞ in the inertial reference frame and the target body has speed U. In the frame of the target, the incoming rock has velocity

$$\mathbf{v}_{1tar} = (-[v_x - U], v_y) \tag{4.6}$$

and the outgoing velocity has the form

$$\mathbf{v}_{2tar} = (+[v_x - U], v_y) \ . \tag{4.7}$$

This second equation assumes that the encounter is symmetric, i.e., the outgoing trajectory of the rock is the mirror image of the incoming trajectory. This approximation thus assumes that the larger body does not change its velocity (consistent with the ordering of equation [4.1]) and that the encounter time is short compared to the orbital period. In the center of mass reference frame, the incoming velocity has the form

$$\mathbf{v}_{1cm} = (-v_x, v_y) = (-v\cos\theta, v\sin\theta) \ , \tag{4.8}$$

whereas the outgoing velocity becomes

$$\mathbf{v}_{2cm} = (+[v_x - 2U], v_y) = (+[v\cos\theta - 2U], v\sin\theta) \ . \tag{4.9}$$

The final speed is then given by the expression

$$v_2^2 = [v\cos\theta - 2U]^2 + v^2\sin^2\theta = v^2 - 4Uv\cos\theta + 4U^2 \ . \tag{4.10}$$

[2]Note that the interpretation of this fiducial cross section would be more complicated if the planetary orbit had significant eccentricity. Nonetheless, one can always scale the results to the expression of equation (4.5).

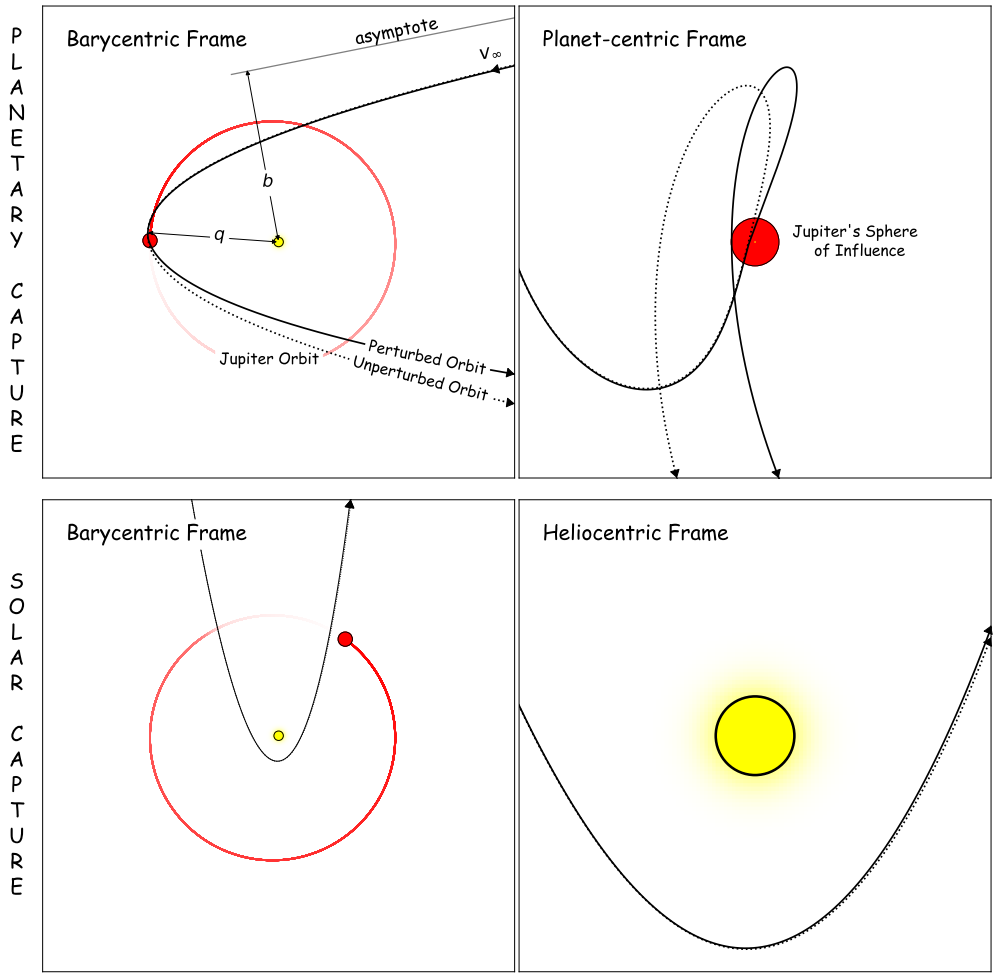


Figure 4.1: Schematic diagrams of capture events. In each frame, the dotted black line denotes the rock's initial (unbound) orbit; the solid black line denotes the rock's trajectory after a perturbation; and the arrows specify the direction of the orbit. The yellow circle represents the Sun, the filled red circle represents Jupiter's sphere of influence drawn at the epoch of the rock's closest approach, and the red line marks Jupiter's orbit. The top row shows a capture by Jupiter: the left panel is in the frame of the solar system's barycenter, and the right panel is in Jupiter's rest frame. Note that the rock gets well inside of Jupiter's sphere of influence, but does not actually collide with the planet. The bottom row shows a capture by the Sun: the left panel is in the frame of the solar system's barycenter, and the right panel is in the Sun's rest frame. If the target (Sun or planet) has a component U of its velocity moving away from the incoming rock as it approaches periapsis, then the encounter causes the rock to lose energy in the inertial reference frame, thereby allowing the rock to potentially enter into a bound orbit.

Note that this discussion assumes that the encounter is symmetric in the reference frame of the capturing body. This approximation is expected to be valid because only close encounters with the body result in capture events, and such close encounters will be symmetric to leading order.

The discussion thus far has implicitly assumed that the target is moving in the $-\hat{x}$ direction. In general, however, the target can also have a $\hat{y}$ component to its velocity. Because of the geometry of the encounter, however, only the $\hat{x}$ component of the rock velocity changes (in this approximation). We can thus incorporate the more general case by interpreting the velocity U as the component of the target velocity in the $-\hat{x}$ direction. With this definition of U, the final speed still obeys equation (4.10).

4.2.2 Solar Close Encounters

When the rock is far away from the Sun, it effectively orbits the center of mass of the system. When the radial distance of the rock becomes sufficiently small, however, its orbit is determined by the location of the Sun. We can delineate the boundary between these two regimes by considering the acceleration (and hence forces) in an accelerating reference frame. As expected, the orbit of Jupiter provides an approximate boundary between the outer problem (hyperbolic orbit about the center of mass of the solar system) and the inner problem (close encounter with the Sun).

When the rock enters the sphere of influence of the Sun, its speed is given by

$$v^2 = v_\infty^2 + \frac{2GM}{r}\,, \tag{4.11}$$

where $r < a_J$ is the location of the rock. It will then execute a (hyperbolic) orbit around the Sun. Due to the motion of the Sun about the center of mass of the solar system, the post-encounter velocity will be given by

$$v_2^2 = v^2 - 4Uv\cos\theta + 4U^2\,, \tag{4.12}$$

where U is the component of the solar velocity in the direction of the perihelion of the orbit and θ defines the shape of the hyperbola. Capture of the rock requires that $v_2^2 < v^2 - v_\infty^2$, so that we obtain the constraint

$$v_\infty^2 < 4Uv\cos\theta - 4U^2\,. \tag{4.13}$$

Here, the angle θ is determined by the parameters of the original hyperbolic orbit about the Sun, so that

$$\cos\theta = \frac{1}{e} = \frac{|a|}{(a^2+b^2)^{1/2}}\,, \tag{4.14}$$

where a is the semi-major axis and b is the impact parameter. The capture constraint thus becomes

$$v_\infty^2 < 4\alpha\frac{m}{M}\left(\frac{GM}{a_J}\right)^{1/2}\left[v_\infty^2+\frac{2GM}{r_p}\right]^{1/2}\frac{|a|}{(a^2+b^2)^{1/2}}+\mathcal{O}\left(\frac{m^2}{M^2}\right)\,, \tag{4.15}$$

where we have assumed that the speed U is some fraction of the speed of the Sun in its orbit about the center of mass. Specifically, we define the parameter α such that

$$U \equiv \alpha\frac{m}{M}\left(\frac{GM}{a_J}\right)^{1/2}\,, \tag{4.16}$$

where m is the mass of Jupiter. Since the speed of the incoming rock $v \gg U$ for the close encounters of interest, we ignore the U^2 term. Finally, we evaluate the rock velocity at the perihelion position r_p (see equation [4.4]), as this location corresponds to where the close encounter takes place. Working to consistent order, the expression for the capture constraint can be written in the form

$$v_\infty^2 < 4\alpha\frac{m}{M}\left(\frac{GM}{a_J}\right)^{1/2}\left[\frac{4GM|a|}{b^2}\right]^{1/2} = 8\alpha\frac{m}{M}\left(\frac{GM}{a_J}\right)^{1/2}\frac{GM}{bv_\infty}\,. \tag{4.17}$$

The constraint can be written as a limit on the impact parameter, i.e.,

$$b < 8\alpha\frac{m}{M}\left(\frac{GM}{a_J}\right)^{3/2} v_\infty^{-3}a_J \approx 93 \text{ au } \alpha\left(\frac{v_\infty}{1 \text{ km/s}}\right)^{-3}\,. \tag{4.18}$$

If one requires that the rocky body is not only captured, but is captured into an orbit with semi-major axis less than some maximum value $a_{\rm max}$, then the left-hand-side of equation (4.17) can be replaced with $v_\infty^2 + v_{\rm x}^2$, where $v_{\rm x}^2 \equiv GM/a_{\rm max}$. Finally, note that this treatment implicitly assumes that $U > 0$. If the Sun is moving in the opposite direction, the encounter would cause the incoming rocky body to gain energy, and capture does not take place.

Given the approximations presented above, the resulting cross section for capture can be written in the from

$$\sigma = 64\pi\alpha^2 a_J^2\left(mM\right)^2\left(GMa_J\right)^3\frac{1}{v_\infty^2(v_\infty^2+v_{\rm x}^2)^2}\,. \tag{4.19}$$

This cross section is specified up to the constant α, which is expected to be of order (but less than) unity. This form is consistent with those derived earlier by other means (Heggie, 1975; Pineault & Duquet, 1993; Valtonen, 1983). Notice that this derivation breaks down for sufficiently high incoming speeds, $v_\infty \gtrsim 8$ km/s, as shown in Appendix 4.A.

4.2.3 Planetary Close Encounters

Another channel for capture occurs through close encounters with the giant planets, most often Jupiter, which will be considered in this discussion. Equation (4.5) represents the cross section for an incoming rock to enter the sphere of radius a_J. Only a fraction of the incoming trajectories $f_1 = R_{Sol}^2/4a_J^2$ will enter the sphere of influence of Jupiter,[3] delineated by $R_{Sol} \approx a_J(m/M)^{2/5}$ (Bate et al., 1971). However, not all of the orbits that enter the sphere of influence will pass close enough to the planet to experience significant deflection. As a result, we must estimate the smaller fraction f_2 of trajectories that allow for capture.

As a rough approximation, significant deflection requires $\cos\theta$ to be of order (but still less than) unity, which in turn implies $b_{hp} \sim |a|_{hp}$ (equation [4.14]), where (a_{hp}, b_{hp}) correspond to the elements of the hyperbolic orbit around the planet. When the rock encounters the planet, its speed in the solar reference frame is given by equation (4.11) evaluated at $r \approx a_J$. The asymptotic speed $(v_\infty)_{hp}$ for the hyperbolic orbit about the planet depends on the planetary motion, but will typically be of the same order. We can thus write

$$(v_\infty^2)_{hp} = v_\infty^2 + \beta\frac{GM}{a_J} \equiv v_\infty^2 + v_z^2\,, \tag{4.20}$$

where β is a dimensionless factor of order unity and where the second equality defines the velocity scale v_z. The semi-major axis of the hyperbolic planetary encounter is given by

$$(|a|)_{hp} = \frac{Gm}{(v_\infty^2)_{hp}} = \frac{Gm}{v_\infty^2 + v_z^2} \sim \frac{m}{M}a_J\,. \tag{4.21}$$

[3]Note that the sphere of influence, as defined here, corresponds to the location where the incoming trajectory switches from a two-body problem with central mass M to a two-body problem with central mass m in the matched conics approximation. The boundary R_{Sol} is comparable to, but not equivalent to, the Hill radius $R_H = a(m/3M)^{1/3}$.

Since we require $b_{hp} \lesssim |a|_{hp}$ and $|a|_{hp} \ll R_{sol}$, the fraction $f_2 = a_{hp}^2/4a_J^2$. The resulting cross section for capture due to planetary encounters has the form

$$\sigma = \frac{\gamma\pi}{2}\frac{GM}{v_\infty^2}a_J\left(\frac{Gm}{a_J\left(v_\infty^2+v_z^2\right)}\right)^2 = \frac{\gamma\pi a_J^2}{2}\left(\frac{m}{M}\right)^2\left(\frac{GM}{a_J}\right)^3\frac{1}{v_\infty^2\left(v_\infty^2+v_z^2\right)^2}\,, \tag{4.22}$$

where we have introduced a dimensionless factor γ that is expected to be of order unity. Note that this expression has a form similar to that of equation (4.19), which corresponds to the capture cross section for solar encounters. Keep in mind, however, that the velocity scales are different and are expected to obey the ordering $v_x < v_z$.

4.2.4 Energy Distribution of Newly Bound Orbits

Using the results from the previous section, we can write the post-encounter speed of the rock in the form

$$v_2^2 \approx v^2 - 4Uv\cos\theta\,. \tag{4.23}$$

The semi-major axis a_b of the bound orbit is defined so that

$$\frac{GM}{a_b} = 2\frac{GM}{r} - v_2^2 = 4Uv\cos\theta - v_\infty^2\,. \tag{4.24}$$

Let us now define a scale length b_0 according to

$$b_0 \equiv 8\frac{m}{M}\left(\frac{GM}{a_J}\right)^{3/2}\frac{a_J}{v_\infty^3}\alpha \sim 100\text{ au}\left(\frac{v_\infty}{1\,\text{km/s}}\right)^{-3}\,. \tag{4.25}$$

With this construction, the semi-major axis of the bound orbit is given by

$$\frac{GM}{a_b v_\infty^2} = \frac{b_0}{b} - 1 \qquad \Rightarrow \qquad a_b = \frac{a_0 b}{b_0 - b} = \frac{a_0\chi}{1-\chi}\,, \tag{4.26}$$

where we let $a_0 = |a|$ denote the (magnitude of) the semi-major axis of the initial hyperbolic orbit, and where the final equality defines $\chi \equiv b/b_0$. The criterion for obtaining a bound orbit (from the previous section) is equivalent to the requirement $b < b_0$ ($\chi < 1$). Since the cross section depends on b^2, the distribution of impact parameters will be weighted towards larger values. This finding, in turn, implies that typical bound orbits will have final semi-major axes

comparable to the starting (negative, hyperbolic) semi-major axis of the incoming orbit. For $v_\infty = 1$ km/s, for example, bound orbits will typically have $a_b \sim 1000$ au. In order to obtain tighter orbits comparable to the size of the solar system (or even the Kuiper belt), we need $a_b \lesssim 100$ au, which in turn implies that $b \lesssim b_0/10 \sim 10$ au.

If we assume that the impact parameters b are uniformly distributed over an area, with a maximum value b_0, then the probability distribution for the dimensionless quantity χ has the simple form $dP = 2\chi d\chi$. Using equation (4.26), we can determine the probability distribution for the semi-major axes of the bound orbits, i.e.,

$$\frac{dP}{da_b} = \frac{2a_b a_0}{(a_b + a_0)^3} \,. \tag{4.27}$$

As written, this distribution is normalized over the interval $0 < a_b < \infty$.

Note that the distribution of equation (4.27) corresponds to the semi-major axes of the bodies when they are captured. The orbital elements of the captured objects will continue to evolve (e.g., through continued close encounters with the planets), so that quasi-stable orbits will display a different distribution (which should be explored further in future work).

4.3 Numerical Results

The cross sections derived in the previous section made use of a number of approximations. In this section we use a suite of more than 500 million simulations to numerically compute the capture cross section.

4.3.1 Simulation Details

We sample rocks of mass $10^{-9}\ M_\odot$ isotropically on the sphere at a barycentric distance of 10^9 au. Each rock's velocity unit vector is uniquely defined by its position on the sphere, pointing directly toward the solar system barycenter. We then assign each rock an impact parameter at some random angle in its plane tangent to the sphere. We randomly sample the impact parameter uniformly given the condition that the maximum pericenter distance $q_{\rm max} \leq 12$ au—comfortably

above the largest pericenter distance for capture not attributable to chance close encounters with a giant planet. Finally we scale the rock's velocity unit vector by a factor

$$v = \sqrt{v_\infty^2 + \frac{2\mu}{r}} \tag{4.28}$$

where $\mu = G \sum_i m_i$ and $i \in \{\text{Sun, Jupiter, Saturn, Uranus, Neptune}\}$. In Equation 4.28, v_∞ is the rock's field (or cluster) velocity at infinity, and the second term accounts for the kinetic energy that the rock gains by falling from infinity to a barycentric distance r.

The above procedure gives us a state vector, from which we compute a body's Keplerian orbital elements. To save computation time, we use these elements to propagate each rock along its unperturbed hyperbolic orbit to a barycentric distance of 1,000 au. This approximation (that the solar system is a point mass with all of its mass at the barycenter) should be accurate to about one part in 10^9, since the solar system's quadrupole term goes like r^{-3}. Once we have performed the analytic propagation of the rock, we use NASA's development ephemerides to initialize the solar system at a random date in a 200–year range around the arbitrarily chosen Julian Date 2459010.5. This ensures that our results are not affected by some exceptional coincidence in the initial phases of the giant planets' orbits.

When we have initialized our rock and the solar system, we use `Rebound's IAS15` integrator (Rein & Spiegel, 2015) to evolve the system numerically. For each simulation, we conserve the system's total energy to better than one part in 10^{14}—much smaller than the fraction of the system's energy attributable to the rock. Therefore we are confident in the accuracy of our integrations.

For each integration, there are three possible outcomes: the rock may be captured; undergo a collision with another body; or be ejected from the system. If at any point during the simulation the rock's energy drops below zero, we consider it to be captured and end the simulation. If the rock undergoes a collision or if the rock is unbound and exiting the solar system with a barycentric distance greater than 40 au, we end the simulation and determine that the rock was not captured. We then follow up on our captured objects, integrating for 51% of an orbital period to ensure that each object is truly bound (as opposed to having a transient bound osculating semi-major axis

due to the phases of the giant planets). If during our followup the object's apocenter distance exceeds 1 parsec, we consider it to be lost to cluster or galactic tides.

Current models of solar system formation predict that the giant planets formed in a more compact arrangement, and then migrated to their current orbits. To account for this we ran a set of simulations with the compact solar system model presented in Tsiganis et al. (2005). The cross section we calculate with this model differs from that calculated using the solar system at the current epoch by less than 1 percent, so our calculations should be equally applicable to the pre-and-post-instability architectures of the solar system.

4.3.2 Capture Cross Section

Since we sampled events uniformly in impact parameter, we can calculate the capture cross section as

$$\sigma = \frac{2\pi b_{\max}}{N} \sum_i b_i \delta_i \tag{4.29}$$

where $b_{\max}$ is the maximum impact parameter sampled, N is the number of events, and δ_i is a Kronecker delta that is 1 if the event resulted in capture, and 0 otherwise. We display our results in Figure 4.2. As we expect, $\sigma(v_\infty)$ goes like v_∞^{-2} in the low-speed limit, and like v_∞^{-6} in the high-speed limit. To facilitate the use of the cross section in analytic calculations we fit $\sigma(v_\infty)$ with the simple function

$$\sigma(v_\infty) = \frac{\sigma_0}{u^2(u^2+1)^2} \tag{4.30}$$

where $u \equiv v_\infty / v_\sigma$ and v_σ is a velocity scale determined by the properties of the planet ejecting the rock. We find the data are best fit by parameter values $\sigma_0 = 232,250$ au^2 and $v_\sigma = 0.4179$ km/s. Keep in mind that these cross sections apply for capture into any bound orbit.

The scale σ_0 for the cross section obtained from fitting our numerical results can be compared to the analytic estimates of the previous section. If we evaluate equation (4.19) in the high speed limit, then agreement between the analytic and numerical estimates implies that $\sigma_0 = 64\pi\alpha^2 a_J^2 (m/M)^2 (v_\sigma / v_J)^6$, where v_J is the orbital speed of Jupiter. The expressions are equal if the dimensionless parameter $\alpha \approx 0.21$. The analytic and numerical results are in agreement

for all incoming speeds if we identify the scales v_x and v_σ, which is equivalent to considering captures with a maximum (post-encounter) semimajor axis $a_{\max} \approx 5090$ au. Notice also that $v_\sigma \sim v_x \sim (Gm/a_J)^{1/2}$ (see also Appendix 4.C). Similarly, equation (4.22) agrees with the numerical result in the high speed limit if the dimensionless parameter $\gamma \approx 5.8$.

Note that the ensemble of numerical simulations is confined to speeds $v_\infty \leq 15$ km/s. This upper limit is invoked for a number of reasons. Due to the steep power-law fall-off of the capture cross section, relatively few capture events take place at higher speeds, so additional computation leads to diminishing returns. In addition to the steep dependence with v_∞, the numerical data indicate that the power-law begins to break at a comparable speed. Some type of break is expected: For v_∞ greater than ~ 10 km/s, capture by close encounters with the Sun becomes ineffective (see Appendix 4.A). It is noteworthy that the capture cross section at $v_\infty \approx 13$ km/s is comparable to the geometrical area of the Sun ($\sim 7 \times 10^{-5}$ au^2). For larger encounter speeds, incoming rocky bodies are thus more likely to collide with the Sun than be captured into a bound orbit.[4]

4.4 Analysis of Captured Objects

In this section we examine the orbital elements of our captured objects to gain insight into the mechanics of the capture process. In Figure 4.3, we show the impact parameter (and pericenter distance) distribution of the unperturbed orbits of our captured objects for asymptotic speeds v_∞ of 1 and 2 km/s. Each histogram displays a clear relative peak at the pericenter distances corresponding to the orbit of Jupiter, along with a much smaller peak for the orbit of Saturn. Comparison of the two histograms indicates that somewhere between 1 and 2 km/s, the dominant capture process switches from that due to the motion of the solar system barycenter to close encounters with a giant planet (especially Jupiter).

In Figure 4.4, we show the post-capture eccentricity as a function of semi-major axis for a subset of the captured objects with v_∞ of 1 km/s (top panel) and 2 km/s (bottom panel). The

[4]For completeness, we note that due to gravitational focusing, the collision cross section with the Sun is larger than the capture cross section for speeds $v_\infty > 2 - 3$ km/s.

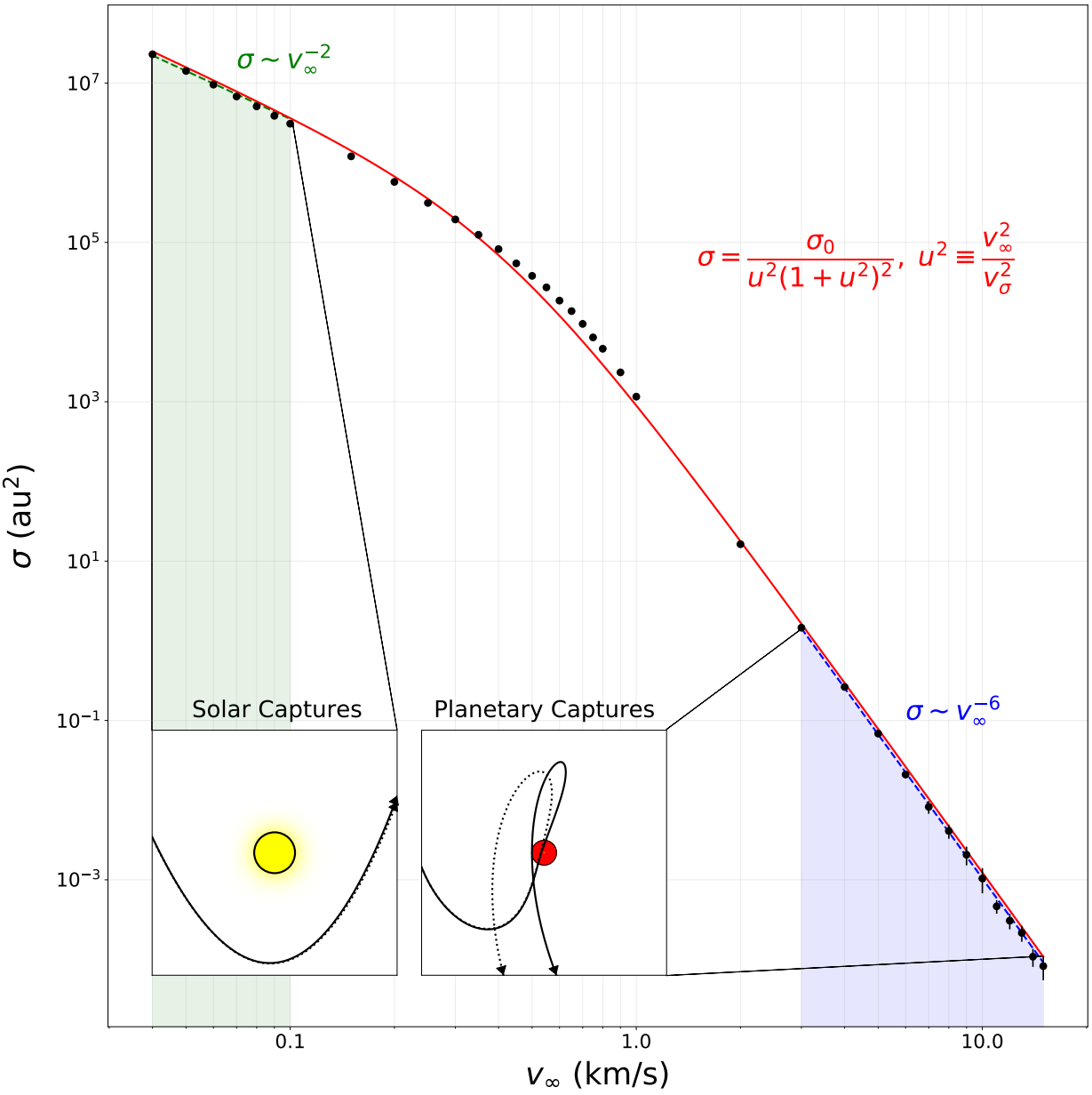


Figure 4.2: Capture cross section (in au^2) as a function of the asymptotic speed v_∞ (in km/s). The black points represent the numerically calculated cross sections, and the corresponding error bars represent one standard deviation. The red curve represents the best fit of equation (4.30) to the data. The blue dashed line shows a power-law of the form $\sigma \sim v_\infty^{-6}$, as expected in the limit of high velocity. The green dashed line shows a power-law of the form $\sigma \sim v_\infty^{-2}$, as expected in the low velocity limit. The shaded regions indicate the parameter space where captures due the Sun (green) and planets (blue) dominate, although the boundaries are not sharp.

figure also includes equi-pericenter curves corresponding to integer multiples of the spheres of influence of Jupiter and Saturn. The numerical results for captures display a relative overdensity

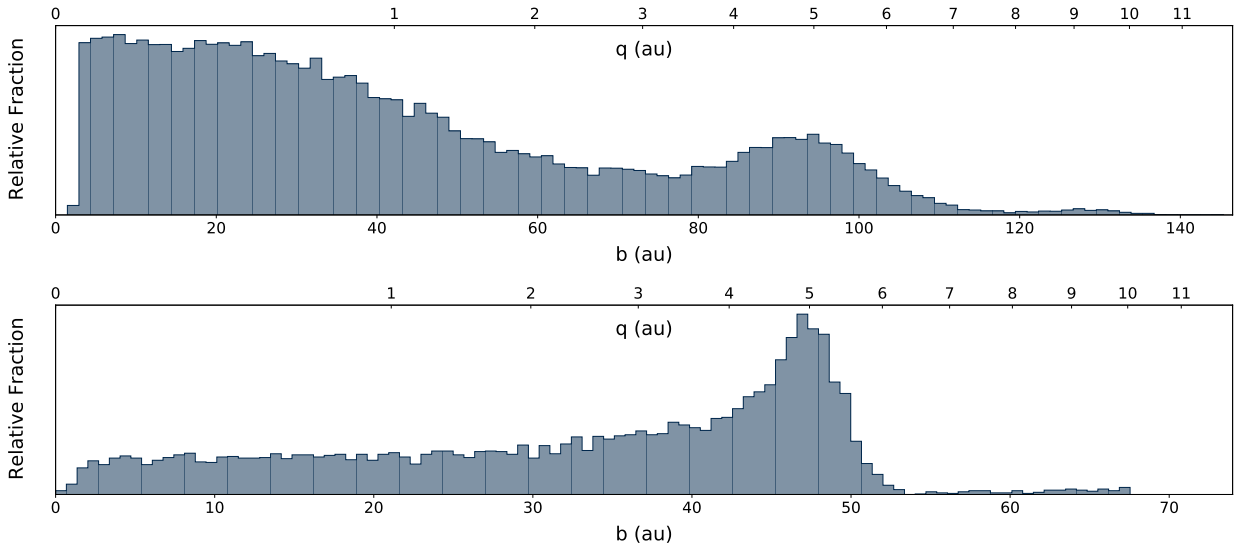


Figure 4.3: Histograms of the (unperturbed) impact parameter distribution of captured objects for asymptotic speeds v_∞ of 1 km/s (top) and 2 km/s (bottom). For convenience, we also indicate the pericenter distance of the unperturbed orbit. Both plots show relative peaks at pericenter distances corresponding to the orbits of Jupiter and Saturn. As v_∞ increases, close encounters with the giant planets become more important for capture.

of points with pericenter distances at Jupiter and Saturn, indicating that these captures are (likely) attributable to close encounters.

In Figure 4.5, we show the kernel density representations for the post-capture inclination semi-major axis, and eccentricity for $v_\infty = 0.5$, 1, and 2 km/s. While captures become increasingly rare for higher-velocity events, the resulting semi-major axes of the captured objects are typically smaller than those for objects captured in low-velocity events. This trend is important for assessing object retention, as captured bodies with semi-major axes $a \gtrsim 1000$ au are more likely to be stripped from the solar system by interacting with passing stars (in the solar birth cluster) or the galactic tides (in the field).

It is noteworthy that capture events readily produce highly-inclined and even retrograde objects. This finding indicates that capture is yet another potential channel for the production of the observed populations of highly-inclined and retrograde centaurs, which are currently best explained by the putative Planet Nine (Batygin et al., 2019). However, it is important to note that the orbits of the captured objects will evolve over time. As a result, the captured objects do

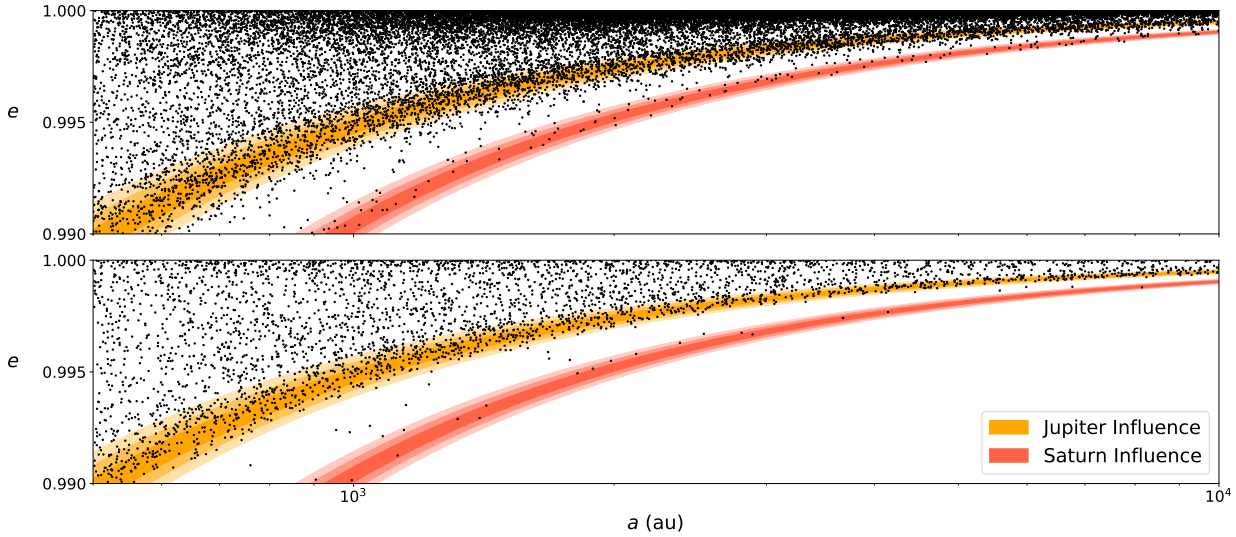


Figure 4.4: Post-capture eccentricity versus semi-major axis of captured objects for incoming speeds of 1 km/s (top) and 2 km/s (bottom). The orange and red regions correspond to integer multiples of the radius of influence centered at the the semimajor axis of the orbits of Jupiter and Saturn, respectively.

not represent a long-term stable population. As the orbits of the captured objects evolve, some will become more eccentric until they collide with the Sun; some will undergo scattering events or interactions with the Galactic tides and be ejected from the solar system; some will be frozen into the inner Oort cloud by passing stars; and some will continue to evolve on stable or quasi-stable orbits (either by becoming caught in resonances with the giant planets, or by achieving orbits that otherwise avoid close encounters).

Because most capture events resulted in high-eccentricity orbits, we have rather low statistics for small values of eccentricity. Despite the data limitations, though, it is clear that as v_∞ increases, the low-eccentricity tail of the distribution becomes fatter.

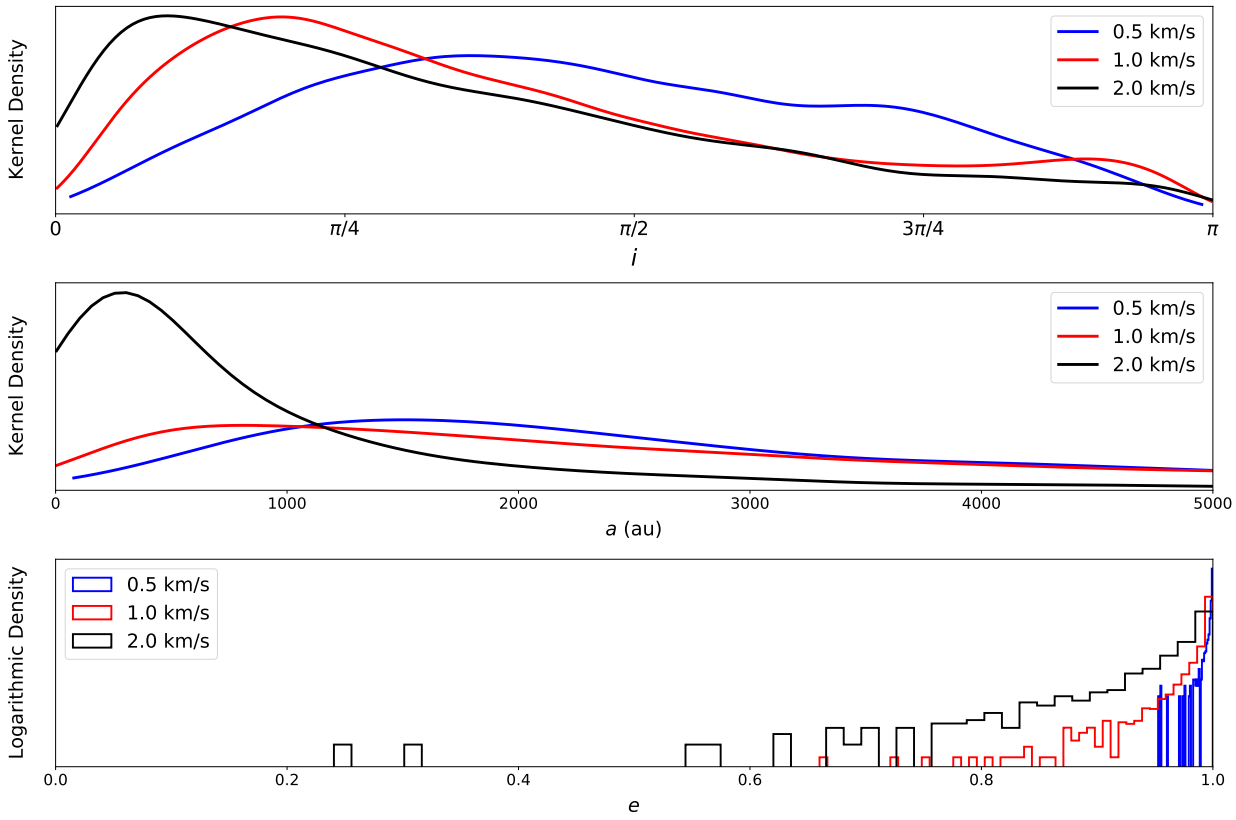


Figure 4.5: Post-capture inclination, semi-major axis, and eccentricity distributions. (Top) Gaussian kernel density estimate of the post-capture inclination distribution (with i measured from the ecliptic plane) of captured objects at speeds $v_\infty = 0.5$, 1, and 2 km/s. (Center) Gaussian kernel density estimate of the post-capture semi-major axis distribution of captured objects for $v_\infty = 0.5$, 1, and 2 km/s. (Bottom) Relative fraction (in logarithmic scale) of the post-capture eccentricity distribution of captured objects for $v_\infty = 0.5$, 1, and 2 km/s. Note that these curves represent probability distribution functions; there will be fewer captures in the high-speed case than in the low-speed case.

4.5 Applications

4.5.1 Velocity Averaged Cross Sections

This paper determines the velocity dependent cross section $\sigma(v_\infty)$, which can be fit with a function of the form

$$\sigma(v_\infty) = \frac{\sigma_0}{u^2(1+u^2)^2} \qquad \text{where} \qquad u \equiv \frac{v_\infty}{v_\sigma}\,. \tag{4.31}$$

The capture rate for rocky bodies by our solar system is given by

$$\Gamma = n_R \langle \sigma v_\infty \rangle \,, \tag{4.32}$$

where n_R is the number density of rocks that the solar system encounters. The capture rate depends on the velocity-averaged cross section, which is given by the integral

$$\langle \sigma v_\infty \rangle = \int_0^\infty v_\infty f(v_\infty) \sigma(v_\infty) dv_\infty \,, \tag{4.33}$$

where $f(v_\infty)$ is the distribution of encounter velocities of the rocky bodies.

The distribution of relative speeds $f(v_\infty)$ depends on the environment. In the solar birth cluster, $f(v_\infty)$ is determined by the processes that eject the rocky bodies from their original planetary systems. In general, the clusters are not sufficiently long-lived for the rocks to attain a thermal distribution of speeds. Instead, they are expected to retain the velocity distribution resulting from the ejection process. If ejection occurs through scattering interactions with giant planets, then $f(v_\infty)$ takes the approximate form

$$f(v_\infty) = \frac{4 v_\infty / v_p}{(1 + v_\infty^2 / v_p^2)^3} \,, \tag{4.34}$$

where the velocity scale $v_p^2 \approx GM_*/a_p$, where a_p is the semi-major axis of the planet that scatters the rocks (e.g., see Moorhead & Adams 2005 for a derivation).

Note that the distribution (4.34) is normalized over the entire interval $0 \leq v_\infty \leq \infty$. In practice, the distribution will have a maximum value determined by the escape speed from the planets that scatter the rocky bodies. Notice also that the full distribution will be a convolution of the distribution of ejection speeds from each planet that scatters rocky material. As an approximation, we consider only a single distribution and interpret the velocity scale v_p as a typical value. As a result, v_p is expected to be comparable to the orbit speed of outer planets, i.e., $v_p \sim 10$ km/s. Finally, we are assuming that equation (4.34) corresponds to the distribution of *relative* speeds between the rocks and the solar system (e.g., see the discussion of Binney & Tremaine 2008).

Putting the above considerations together, we can write the velocity averaged cross section in the form

$$\langle \sigma v_\infty \rangle = \sigma_0 \frac{v_\sigma^3}{v_p^2} \int_0^\infty \frac{4du}{(1+\eta^2 u^2)^3 (1+u^2)^2} = \sigma_0 \frac{v_\sigma^3}{v_p^2} I(\eta) \,, \tag{4.35}$$

where $u = v_\infty / v_\sigma$ (as before), we have defined $\eta \equiv v_\sigma / v_p$, and where the second equality defines the integral function $I(\eta)$. The dimensionless function $I(\eta)$ can be evaluated to obtain

$$I(\eta) = \pi \frac{(1+\eta)(1+3\eta) + 3\eta^3/4}{(1+\eta)^4} \,. \tag{4.36}$$

Note that $I \to \pi$ in the limit $\eta \to 0$, and in practice $\eta \sim 1/10$. As a result, a good approximation for the capture cross section takes the form

$$\langle \sigma v_\infty \rangle \approx \pi \sigma_0 \frac{v_\sigma^3}{v_p^2} \,. \tag{4.37}$$

For comparison, we can determine the velocity averaged cross section for the scenario where the cluster rocks are virialized and have the same (Maxwellian) velocity distribution as the stars.[5] In this limit, $\langle \sigma v_\infty \rangle$ can be written in the form

$$\langle \sigma v_\infty \rangle = \sigma_0 \frac{v_\sigma^2}{s} \sqrt{\frac{2}{\pi}} \int_0^\infty \frac{v_\infty dv_\infty}{s^2} \frac{\exp[-v_\infty^2/2s^2]}{(1+v_\infty^2/v_\sigma^2)^2} \,, \tag{4.38}$$

where s is the velocity dispersion of the distribution. Note that the value of s for the distribution of relative speeds is larger than the value s_0 for the velocity distribution of the stars in the clusters ($s = \sqrt{2} s_0$). Here we define the variable $w \equiv v_\infty / s$ and the parameter $\xi = s / v_\sigma$, so that

$$\langle \sigma v_\infty \rangle = \sigma_0 \frac{v_\sigma^2}{s} \sqrt{\frac{2}{\pi}} \int_0^\infty w dw \frac{\exp[-w^2/2]}{(1+\xi^2 w^2)^2} \equiv \sigma_0 \frac{v_\sigma^2}{s} \sqrt{\frac{2}{\pi}} J(\xi) \,, \tag{4.39}$$

where the second equality defines the integral function $J(\xi)$. The exact form for $J(\xi)$ can be found. If we define $\mu = 1/(2\xi^2) = v_\sigma^2/(2s^2)$, then

$$J(\xi) = J(\mu) = \mu \left[1 - \mu \, e^\mu \, E_1(\mu) \right] \,, \tag{4.40}$$

[5] In general, we expect the stars to reach virial equilibrium much faster than the rocky ejecta. The stars start their cluster trajectories with sub-virial speeds, but then fall toward the cluster core where interactions take place, and equilibrium is rapidly realized (in a few Myr, e.g., Adams et al. 2006). In contrast, the rocks are ejected with speeds much larger than the virial speed and have little chance for interactions to slow them down. Moreover, the stellar virialization starts as soon as stars form, whereas the planet formation and the subsequent ejection of rocks occurs many Myr later.

where $E_1(x)$ is the exponential integral (Abramowitz & Stegun, 1972). In the limit $\xi \to 0$ ($\mu \to \infty$), the function $J = 1$; in the opposite limit $\xi \gg 1$ ($\mu \to 0$), $J = \mu = 1/(2\xi^2)$. A good working approximation for the cross section of equation (4.39) thus has the form

$$\langle \sigma v_\infty \rangle \approx \sigma_0 \frac{v_\sigma^4}{s} \frac{\sqrt{2/\pi}}{v_\sigma^2 + 2s^2}\,, \tag{4.41}$$

which is exact in the limits and has a relative error less than $\sim 20\%$ over the entire range $0 \leq \xi \leq \infty$. One can use equations (4.39) and (4.40) if higher accuracy is required.

4.5.2 Rock Capture in the Birth Cluster

Using the results derived above, we can estimate the total mass in rocky bodies that were captured while the Sun remained in its birth cluster. The capture rate is given by equation (4.32), with velocity-averaged cross section specified through equations (4.39) and (4.40). The rocky bodies will have a distribution of sizes $g(R)$, which is defined here such that

$$n_R = \int_0^\infty g(R) dR \qquad \text{and} \qquad \rho_R = \int_0^\infty g(R) m(R) dR\,, \tag{4.42}$$

where $m(R)$ is the mass of the rock as a function of its size. With these definitions, the capture rate Γ can be converted into a mass accretion rate given by

$$\dot{M} = \rho_R \langle \sigma v_\infty \rangle\,, \tag{4.43}$$

where ρ_R is the mass density of the cluster in the form of rocks. Given that each planetary system in the cluster is expected to eject a few Earth masses of rocky material (e.g., Rice & Laughlin 2019), the density ρ_R is given by

$$\rho_R = \frac{\alpha M_\oplus N_*}{V} = \alpha M_\oplus n_*\,, \tag{4.44}$$

where α is a dimensionless factor of order unity and n_* is the number density of stars. For completeness, note that the inclusion of icy planetesimals will increase this density estimate. In any case, the total mass in rocky bodies captured by the solar system during its cluster phase can be written in the form

$$(\Delta M)_R = \alpha M_\oplus \left[\int_0^\infty n_* dt \right] \langle \sigma v_\infty \rangle \equiv \alpha M_\oplus \langle n_* \rangle \tau \langle \sigma v_\infty \rangle\,. \tag{4.45}$$

The final equality defines the mean density of the cluster, where τ is its effective lifetime. A number of studies have found upper bounds on the product $\langle n_* \rangle \tau$ by requiring that the solar system is not overly disrupted, including considerations of the planetary orbits (Adams & Laughlin, 2001; Adams, 2010; Li & Adams, 2015), the Kuiper Belt (Moore et al., 2020), and the orientation of the plane of the cold classicals (Batygin et al., 2020). This work indicates that the product is bounded by $\langle n_* \rangle \tau \lesssim 2 \times 10^4$ pc^{-3} Myr.

If we take $v_\sigma = 0.5$ km/s, $v_p = 10$ km/s, and $\sigma_0 = 2 \times 10^5$ au^2, then the velocity averaged capture cross section becomes $\langle \sigma v_\infty \rangle \approx 800$ au^2 km/s. With this cross section, the total mass in captured rocky material from equation (4.45) is about $(\Delta M)_R \sim 10^{-3} M_\oplus$. Of course, most of this material will be ejected back into the cluster or the field. The retention rate of material in the inner Oort cloud is $\sim 1\%$ (Brasser et al., 2006), so we would expect $\sim 10^{-5} M_\oplus$ to be captured in the inner Oort cloud.

Note that if rocks ejected from planetary systems in clusters follow the velocity distribution of equation (4.34), then some fraction of the material will leave the cluster during its first crossing. The high speed tail of the velocity distribution will thus be de-populated. In practice, however, most of the capture events arise from the low-speed portion of the distribution, so that the correction for the loss of high speed material is modest.

We can also estimate the mass of rocks captured while the solar system is in the field. In this case, we expect the rocky material to encounter the solar system with a velocity distribution comparable to that of the field stars, i.e., a Gaussian distribution with $s \sim 40$ km/s. In this case, the velocity averaged cross section $\langle \sigma v_\infty \rangle \approx 0.08$ au^2 km/s. If we also assume that each planetary system ejected the same mass in rocks during its formative phases, then the density of rocky material will be proportional to the stellar density (we are thus assuming negligible losses). As a result, the product $\langle n_* \rangle \tau \sim 460$ pc^{-3} Myr, and the expected mass in captured rocks is about $(\Delta M)_R \approx 2 \times 10^{-9} M_\oplus$. Using the approximate retention rate of 1%, we would expect only $\sim 2 \times 10^{-11} M_\oplus$ of these rocks to remain in the inner Oort cloud. This inventory of captured alien material from the field is exceedingly small, roughly the equivalent of one 5

km body. Rock capture during the birth cluster phase is thus expected to produce the dominant contribution (by roughly a factor of one million). These latter objects are expected to have radiogenic ages comparable to ordinary solar system bodies, but might be identified by different (unusual) chemical composition.

Note that the values presented here are highly approximate. Not all of the rocks will be captured in the inner Oort cloud, so that the retention fraction could be smaller than assumed here (most of the captured interstellar bodies initially have Jupiter-crossing orbits, whereas the planetesimals in the Oort cloud could have different origins). In any case, most of the captured objects will be ejected, and some will eventually collide with the Sun. Although these calculations provide working order-of-magnitude estimates, in forthcoming work we will refine these projections by numerically investigating the long-term behavior of the captured bodies from this work.

4.6 Conclusions

This paper has revisited the problem of capturing interstellar objects on initially hyperbolic trajectories into bound states. Using an ensemble of 500 million numerical fly-by simulations, the main result of this study is the determination of our solar system's capture cross section as a function of encounter speed (see Figure 4.2). The resulting capture cross section shows the power-law velocity dependence $\sigma \sim v_\infty^{-2}$ in the limit of low speeds and the dependence $\sigma \sim v_\infty^{-6}$ in the limit of high speeds. The capture cross section $\sigma(v_\infty)$ over the entire range of asymptotic speeds can be fit with the function given in equation (4.31).

This paper also presents an analytic treatment of the capture problem using the approximation of matched conics and the (inverse) gravitational slingshot effect (Section 4.2). These arguments show that capture by both close encounters with the Sun (Section 4.2.2), and by close encounters with a giant planet (Section 4.2.3), have the same nearly velocity dependence as that seen in the numerical simulations (namely $\sigma \sim v_\infty^{-2}$ at low speeds and $\sigma \sim v_\infty^{-6}$ at high speeds).

The capture events can be classified as either close encounters with the Sun or close encounters with giant planets. At low speeds, encounters with the Sun dominate the capture cross section.

At higher speeds, close encounters with Jupiter dominate. Close encounters with the other giant planets contribute to the cross section, but do not dominate the dynamics. More specifically, for the particular case of $v_\infty = 1$ km/s, capture events due to close encounters with Jupiter are ~ 100 times more likely than captures due to Saturn. The frequency of close encounters with Uranus and Neptune are smaller (than for Saturn) by an additional factor of ~ 100.

With the capture cross section as a function of velocity specified, the effective mean cross section $\langle v\sigma \rangle / \langle v \rangle$ can be determined for any distribution of encounter speeds. For the case of a Maxwellian distribution and a power-law distribution motivated by rock ejection, the mean cross section can be evaluated analytically (see Section 4.5.1).

Finally, as an application of the capture cross section, we estimate the total mass $(\Delta M)_R$ in the Oort cloud that originates from other planetary solar systems (Section 4.5.2). The mass accreted while the Sun lived within its birth cluster is of order $(\Delta M)_R \sim 10^{-5} M_\oplus$, about a million time larger than the mass subsequently accreted from the field.

Although the capture cross section for the solar system is now well-characterized, many avenues for future research remain. The simulations of this paper consider the capture of interstellar objects and the resulting cross sections include all capture events, independent of their residence time in the solar system as bound objects. Future work should determine how long captured bodies can remain bound to the Sun, since many such objects are expected to be ejected from the system or to collide with other solar system members. The residence time (ejection time) should thus be determined for each type of orbit displayed by the captured objects. With these results in place, one can make a refined estimate of the current population of alien objects in the solar system, along with their expected orbital properties.

4.A Upper Bound on Incoming Speed for Capture

In this Appendix, we find upper limits on the asymptotic speed, i.e., the largest speed for which a rock can be captured by the solar system. We consider both close encounters with the Sun and with one of the giant planets.

We start with the capture criterion of equation (4.13), which we reproduce here,

$$v_\infty^2 < 4Uv\cos\theta - 4U^2 < 4Uv\,, \tag{4.46}$$

where the second inequality follows because $\cos\theta < 1$ and $4U^2 > 0$. The largest speed U possible for the Sun, relative to the center of mass, is the orbit speed due to the giant planets, where

$$U = \frac{m}{M}\left(\frac{G(M+m)}{a_p}\right)^{1/2} \approx \frac{m}{M}\left(\frac{GM}{a_p}\right)^{1/2}\,, \tag{4.47}$$

where the second expression is consistent with the ordering approximation and where we have included the reflex speed due to only one planet. The maximum possible speed is given by the sum of the contributions of all of the planets. We can account for this complication by writing the limit in the form

$$U < \frac{2m_J}{M}\left(\frac{GM}{a_J}\right)^{1/2}\,. \tag{4.48}$$

Thus far, our upper bound has the form

$$v_\infty^2 < 8\frac{m_J}{M}\left(\frac{GM}{a_J}\right)^{1/2} v = 8\frac{m_J}{M}\left(\frac{GM}{a_J}\right)^{1/2}\left[v_\infty^2 + \frac{2GM}{R_\odot}\right]^{1/2}\,, \tag{4.49}$$

where we have taken the minimum distance (and maximum speed) of the orbit to be given by the radius of the Sun. If we drop the first term in square brackets, the expression simplifies to the form

$$v_\infty^2 < 8\frac{m_J}{M}\left(\frac{GM}{a_J}\right)^{1/2}\left(\frac{2GM}{R_\odot}\right)^{1/2} = 8\sqrt{2}\frac{Gm_J}{(a_J R_\odot)^{1/2}} \approx (8\,\mathrm{km/s})^2\,. \tag{4.50}$$

To derive a rigorous version of the upper bound, we define two velocity components

$$u \equiv 8\frac{m_J}{M}\left(\frac{GM}{a_J}\right)^{1/2} \qquad \text{and} \qquad w \equiv \left(\frac{2GM}{R_\odot}\right)^{1/2}\,, \tag{4.51}$$

so that the limit takes the form

$$v_\infty^2 < uw\left[1 + \frac{u^2}{4w^2}\right]^{1/2} + \frac{u^2}{2}\,. \tag{4.52}$$

Since u is *of order* the orbital speed of the Sun (specifically, a small fraction of 1 km/s) and w is the escape speed of the Sun (~ 620 km/s), we find $u \ll w$. In this limit, the full expression

of equation (4.52) reduces to the form $v_\infty^2 < uw$, which corresponds to the approximation of equation (4.50).

We can also consider the case where the incoming rock enters into the sphere of influence of a planet and loses energy through an inverse gravitational assist from the encounter. In order for the rock to be captured by the solar system, the final speed must be sufficiently small, i.e.,

$$v_2^2 < 2v_{\rm orb}^2 \,, \tag{4.53}$$

where the orbital speed $v_{\rm orb}$ of the planet is a measure of the depth of the gravitational potential well at the location of the planet (the location of the encounter). We also assume that the incoming rock obeys conservation of energy so that

$$\frac{1}{2}v_\infty^2 = E = \frac{1}{2}v^2 - \frac{GM}{r} \approx \frac{1}{2}v^2 - v_{\rm orb}^2 \qquad \Rightarrow \qquad v^2 = v_\infty^2 + 2v_{\rm orb}^2 \,. \tag{4.54}$$

Using this expression in the result for the post-encounter speed from before, we obtain

$$v_\infty^2 + 4U^2 < 4U\cos\theta[v_\infty^2 + 2v_{\rm orb}^2]^{1/2} \,. \tag{4.55}$$

If we take $U = v_{\rm orb}\cos\phi$, this expression can be rewritten in the form

$$v_\infty^4 < 8v_{\rm orb}^2 v_\infty^2 \cos^2\phi(2\cos^2\theta - 1) + 16v_{\rm orb}^4\cos^2\phi(2\cos^2\theta - \cos^2\phi) \,. \tag{4.56}$$

We can thus obtain an upper limit on the asymptotic speed for an object to be captured by taking $\cos\theta = 1 = \cos\phi$, i.e.,

$$v_\infty^4 < 8v_{\rm orb}^2 v_\infty^2 + 16v_{\rm orb}^4 \,, \tag{4.57}$$

which leads to the bound

$$v_\infty < 2v_{\rm orb}(1+\sqrt{2})^{1/2} \,. \tag{4.58}$$

The fastest orbit is that of Jupiter, where $v_{\rm orb} \approx 13$ km/s, so we have the limit $v_\infty \lesssim 40$ km/s.

We thus find that the maximum possible speed (for capture) is larger for the channel involving close approaches to planets (in particular Jupiter) than for close approaches to the Sun. On the other hand, planet encounters are expected to occur much less frequently. Note that these limits do not account for effects such as radiation pressure or atmospheric drag. It is therefore possible in principle to capture rocks with larger v_∞, but such events would be exceptionally rare.

4.B Rock Capture by Circumstellar Disks

Rock capture could also take place due to gas drag if interstellar bodies enter the solar nebula while it retains its gaseous component. This Appendix explores the efficacy of this process.

If a rock has speed v_0 when it enters the gaseous region of the disk, it will have final speed v_f given by

$$v_f = v_0 \exp[-\rho A \ell / m] \,, \tag{4.59}$$

where ρ is the gas density, A is the rock area, m is the rock mass, and ℓ is the length traveled. If a rock comes from infinity with speed v_∞, then it will have a larger velocity when it hits the disk, so that v_0 is given by

$$v_0^2 = v_\infty^2 + \frac{2GM}{r} \,. \tag{4.60}$$

In order for the rock to lose enough energy to enter into a bound orbit, the final speed must be less than the limit

$$v_f^2 = \left[v_\infty^2 + \frac{2GM}{r} \right] \exp\left[-\frac{2\rho A \ell}{m} \right] < \frac{2GM}{r} \,. \tag{4.61}$$

The capture criterion thus becomes

$$\frac{2\rho A \ell}{m} > \ln\left[1 + \frac{v_\infty^2 r}{2GM} \right] \,. \tag{4.62}$$

If the rock passes through the disk vertically, then $\rho \ell = \Sigma$, where the surface density for the solar nebula can be written in the form

$$\Sigma(r) = \Sigma_1 \left(\frac{1\mathrm{au}}{r} \right)^{3/2} \,, \tag{4.63}$$

where $\Sigma_1 \approx 3000$ g/cm^2. If the rock passes through the nebula at an arbitrary angle θ, then $\rho \ell = \Sigma / \cos\theta$. We can also write $A = \pi R^2$ and $m = \rho_R (4\pi/3) R^3$ where R is the size of the rock. If we work in the low speed limit with $v_\infty = 1$ km/s, and write R in units km and r in units of au, then the criterion for capture reduces to the simple form

$$\frac{36}{\cos\theta} > R r^{5/2} \,. \tag{4.64}$$

For the fiducial case where $\cos\theta = 1$, the solar nebula can thus capture incoming rocks with radii $R \lesssim 1$ km at radial location $r = 1$ au. Somewhat larger rocks can be captured for typical inclined trajectories (e.g., the limit becomes $R \lesssim 2$ km for $\cos\theta = 1/2$). Smaller rocks are readily captured, and rocks as large as dwarf planets are highly unlikely to be captured (see also Brasser et al. 2007). For incoming trajectories that are confined to the plane of the disk, the above treatment must be modified to include the disk structure (i.e., the above treatment does not apply in the limit $\cos\theta \to 0$).

4.C Dimensional Analysis

The capture problem has more than one dimensionless field, so that the cross section of interest cannot be directly determined from dimensional analysis. Nonetheless, such an analysis is useful for understanding the result.

Consider the solar system to consist of only the Sun-Jupiter binary. The physical variables required to characterize the systems are thus the masses (M, m) and the orbital radius of Jupiter a_J. Note that the eccentricity of the Jovian orbit is too small to matter, and the radii of the bodies also do not contribute. The other variables in the problem are the gravitational constant G and the speed of the incoming rock v_∞. Here we want to determine the cross section as a function of these variables. We thus want to determine the impact parameter b of the incoming orbits that allow for capture.

The first estimate for the cross section is thus

$$\sigma_1 = \pi a_J^2 \,. \tag{4.65}$$

The remaining (relevant) variables can be combined to construct two dimensionless fields, i.e.,

$$\Lambda_1 \equiv \frac{GM}{v_\infty^2 a_J} \qquad \text{and} \qquad \Lambda_2 \equiv \frac{Gm}{v_\infty^2 a_J} \,. \tag{4.66}$$

Alternatively, one could define the second field to be the mass ratio m/M, and the quantity $\Lambda_2 = \Lambda_1(m/M)$. At low speeds, the capture cross section has the form

$$\sigma_2 = \pi a_J^2 \Lambda_1 \,, \tag{4.67}$$

whereas at high speeds the cross section becomes

$$\sigma_3 = \pi a_J^2 \Lambda_1 \Lambda_2^2 \,. \tag{4.68}$$

We can combine all of the above results to write the capture cross section in terms of the dimensionless fields of the problem:

$$\sigma_4 = \pi a_J^2 \frac{\Lambda_1 \Lambda_2^2}{(1+\Lambda_2)^2} \,. \tag{4.69}$$

Chapter 5

On the Fate of Interstellar Objects Captured by Our Solar System

> Put that thing back where it came from, or so help me!
>
> — Mike Wazowski, *Monsters, Inc.*

This chapter is published as Napier et al. (2021b).

5.1 Introduction

The recent discoveries of the irregular body 'Oumuamua (Meech et al., 2017) and the comet Borisov (Jewitt & Luu, 2019) have precipitated a boom in studies of interstellar objects. A pertinent open question of both scientific and general interest is whether our solar system currently harbors any material of extrinsic origin. Currently no objects present overwhelming evidence of having originated outside of our solar system (Morbidelli et al., 2020). Given the complex dynamical architecture of the outer solar system it is not straightforward to determine whether an object is of interstellar origin. A more severe difficulty is that we do not know the types of orbits on which such objects might reside. To address this issue, one must first determine the kinds of orbits onto which interstellar objects can be captured, and then determine which objects can survive for long enough to be considered quasi-permanent members of the solar system, rather than members of a transient population.

The general process of gravitational capture, along with its application to our solar system, has been studied through both analytic and numerical treatments (Heggie, 1975; Valtonen & Innanen,

1982; Valtonen, 1983; Hands & Dehnen, 2020; Lehmann et al., 2021; Napier et al., 2021a). Our recent contribution (Napier et al. 2021a; hereafter Paper I) used an ensemble of $\sim$ 500 million flyby numerical simulations to compute a capture cross section as a function of hyperbolic excess velocity. The resulting cross section, in conjunction with environmental parameters, allows one to compute a capture rate for any velocity distribution of interstellar objects. Although the specification of the capture rate is robust, we also need to know the dynamical lifetime of the captured objects in order to estimate the standing population of alien bodies.

The final piece of this puzzle is thus to determine the dynamical stability and lifetimes for the population of captured interstellar objects. Previous related studies have investigated the dynamical stability of objects such as short-period centaurs (Bailey & Malhotra, 2009), short-period comets (Levison & Duncan, 1994), and long-period comets (Malyshkin & Tremaine, 1999). However, these treatments cannot be directly applied to interstellar objects because the latter are captured on particular types of orbits – with extremely high eccentricities, large semi-major axes, and perihelia lying in the giant planet regime (Napier et al., 2021a). While several previous studies have now considered the capture of interstellar objects, none have offered a complete treatment of the long-term dynamics of the captured objects.

The objective of this paper is to examine the dynamical stability of captured interstellar objects and characterize their dynamical lifetime. We extend our previous work by performing long-term (up to 1 Gyr) simulations of 276,691 synthetic captured interstellar objects (Section 2). These results are then used to compute a dynamical lifetime function, i.e., the fraction of captured objects that remain in bound orbits as a function of time since capture (Section 3). We then consider (in Section 4) the secular evolution of the orbital elements of the captured objects in order to uncover the dynamics governing their survival at long times, and comment on the implications for our solar system. Finally, we combine the dynamical lifetime function of this paper with the capture cross sections to obtain a refined estimate (in Section 5) of the total mass of alien rocks on bound orbits in the solar system (this estimate is limited to objects with pericenter within the orbit of Neptune). The paper concludes (in Section 6) with a summary of

our results and a discussion of their implications.

5.2 Numerical Simulations

In order to understand and quantify the survival rate for captured interstellar objects, we use an ensemble of numerical simulations to study their long-term dynamical evolution and stability. More specifically, we perform numerical integrations for each of the 276,691 synthetic captured interstellar objects resulting from the simulations of Paper I. Note that these captured objects arise from encounters with given values of the initial (pre-capture) asymptotic speed v_∞. We can thus determine how/if survival depends on v_∞.

The numerical simulation details are as follows. We use `Rebound's IAS15` integrator (Rein & Spiegel, 2015) to track the orbit of each captured object for up to 1 Gyr (note that most of the objects are lost on shorter times scales). We use the same simulation details as Paper I, but instead of halting the simulation when an object becomes bound we continue the integration until one of the following events occurs: the object becomes unbound from the system, its apocenter distance exceeds the somewhat arbitrarily-chosen value of 50,000 au, it collides with a massive body, or the integration reaches 1 Gyr. Our simulations did not account for galactic tides and passing stars, as we assume that orbital diffusion is dominated by interactions with giant planets. We record the state vector of each captured object every 100 years for later use in studying the secular evolution of its orbit.

Here we analyze the simulation data at the conceptual level to get a feel for the dynamics. We perform a more in-depth analysis of the long-lived objects in Section 5.4. In Figure 5.1 we show the time-series of each object's pericenter distance and inclination. Each line represents a single object, and truncates when the object was ejected from the system. It is difficult to make out any structure for the first $\sim$ 10 Myr, but soon thereafter the density of points falls off and some trends begin to emerge. There are a few features of note:

- To survive for more than a few million years, captured objects must somehow lift their pericenters beyond Jupiter (survival in this context means that the object remains bound

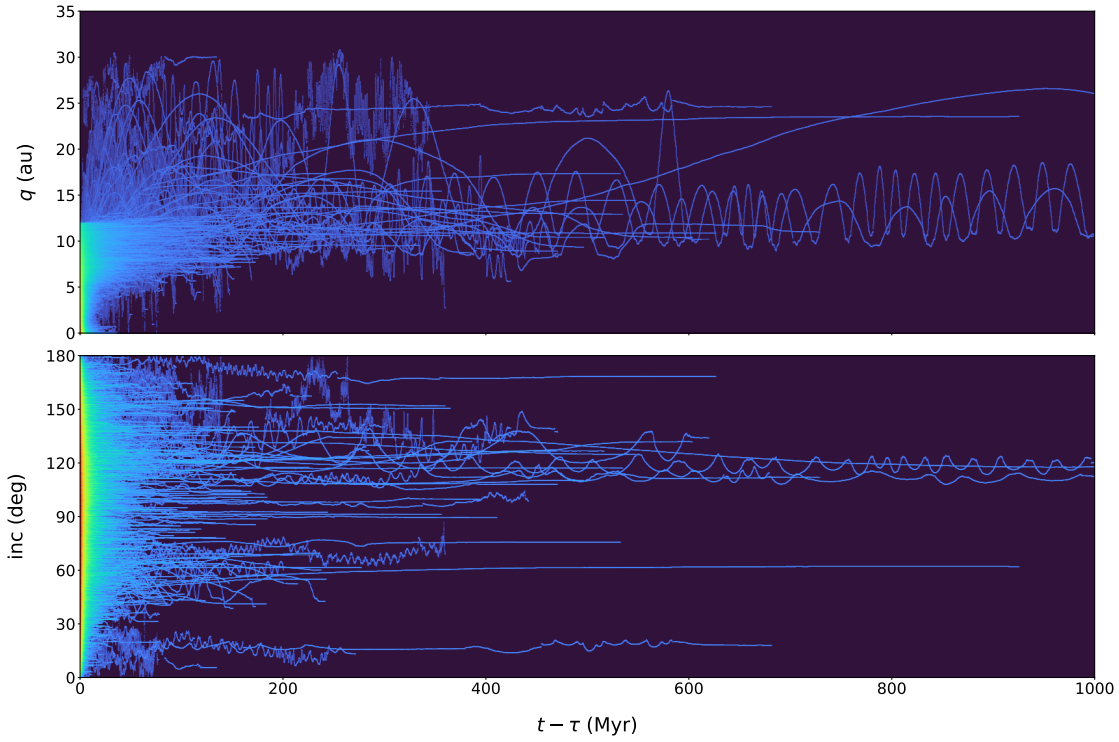


Figure 5.1: Time-series histogram of osculating pericenter distance (top) and inclination (bottom) for each captured object. The color gradient corresponds to the point density. After approximately 100 Myr, individual objects become distinguishable.

to the system).

- Objects on highly inclined orbits tend to survive for longer than those on planar orbits.

- No object achieved permanent trans-Neptunian status (i.e., $q > 30$ au).

These trends all make sense intuitively, and we briefly address each here. Objects that fail to lift their pericenters beyond Jupiter are likely to eventually have a close encounter with the planet and thus be scattered. Changes in pericenter can be achieved via mechanisms such as scattering or the von Ziepel-Lidov-Kozai (vZLK) effect (von Zeipel, 1910; Lidov, 1962; Kozai, 1962; Ito & Ohtsuka, 2019). Objects that are on highly inclined orbits are less likely to encounter a planet because they spend most of their time out of the plane of the solar system. However, objects on orbits nearly perpendicular to the plane of the solar system are subject to the polar instability, which can cause them to evolve onto unstable orbits and eventually be ejected from the system

(Namouni & Morais, 2020). Finally, in the absence of perturbations beyond Neptune, attaining a pericenter distance greater than 30 au requires some combination of serendipitous scattering events and vZLK resonances.

For completeness, we note that our simulations make a number of minor approximations. First, the simulations include only the four giant planets and the Sun. We do not account for external forces such as out-gassing, radiation pressure during close approaches to the Sun, or drag due to orbits that cross through planetary atmospheres (which is an exceedingly rare event). In addition, we neglect the higher-order gravitational harmonics that arise from modeling the shape and rotation of massive bodies. Each of these approximations is rather modest, so that including them would make relatively little difference to our conclusions. We reiterate that our simulations did not model galactic or cluster tides, and did not include perturbations by passing stars. Incorporating these effects could allow some of the captured objects to lift their perihelia beyond Neptune, which would make them more likely to survive for the age of the solar system. These effects should be explored more thoroughly in future work.

5.3 Dynamical Lifetime

In this section we use the results of our simulations to characterize a dynamical lifetime for captured interstellar objects. Specifically we calculate the probability that any object still resides in the solar system after a given time t, and we designate this function as $f(t)$.

5.3.1 Numerical Results

The combined set of numerical data can be used to construct the fraction $f(t)$ of rocky bodies that survive as a function of time, where the result is shown in Figure 5.2. In addition, we can write down a simple function that fits the data, and find that the expression

$$f(t) = \frac{1}{u^{\beta} + 1} \qquad \text{where} \qquad u \equiv \frac{t}{\tau} \tag{5.1}$$

agrees with our data over the entire range of lifetimes under consideration. Here the parameter τ acts as a characteristic timescale for the system, and β is the power law index for depletion at long times. The best-fit values, with uncertainties calculated using a bootstrapping technique (see, e.g., Efron & Tibshirani 1986), are $\tau \approx 0.84 \pm 0.04$ Myr, and $\beta \approx 1.6 \pm 0.04$ (note that we can take $\beta = 8/5$). In addition to our numerically determined survival function, Figure 5.2 displays the best fit to Equation (5.1) as the solid blue curve. The grey curve (shaded region) represents the uncertainty in the curve of best fit. For comparison, the figure also shows a previously obtained prediction for $f(t)$ from a diffusion model as the dotted red curve (see Yabushita 1980 and Section 5.3.2).

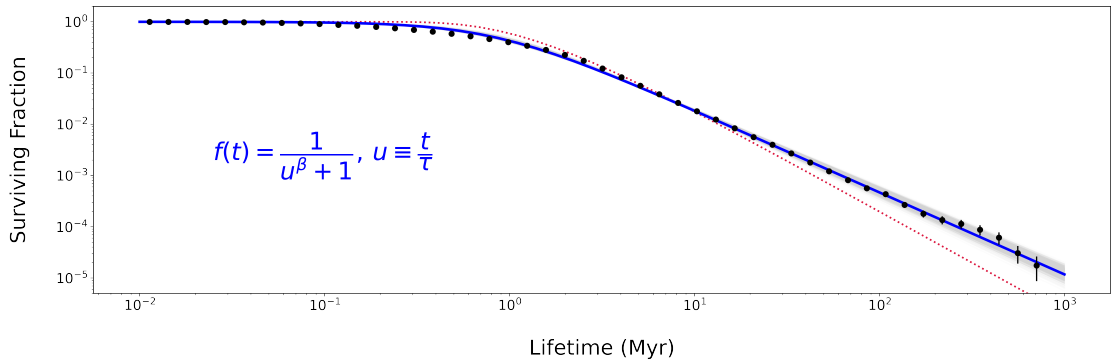


Figure 5.2: Surviving fraction of captured objects as a function of integration time. The black points represent the data, and the blue curve is the best fit using the expression of Equation (5.1). The best-fit values are $\beta = 1.6 \pm 0.04$ and $\tau = 0.84 \pm 0.04$ Myr. The grey curve represents the uncertainty in the best fit. The dashed red line is an incomplete gamma function with index 2, as predicted by the diffusion model from Yabushita (1980).

These data show that the population of rocky bodies remains relatively constant for a few million years before the system steadily ejects the captured objects. After this time, the lifetime function $f(t)$ takes the form of a power law. The fact that $f(t)$ transitions to a power law after a few million years has interesting implications. First we note that the orbits of captured object typically have semi-major axes $a \sim 1000$ au with periods $P \sim 30,000$ yr. It thus takes (at least) ~ 1 Myr for enough orbits (here, more than 30) to accumulate to lead to ejection. In addition, the Lyapunov time of the outer solar system is roughly comparable to the transition time scale (e.g., Laskar 1989). This finding suggests that the system can be influenced by chaos over the

time scales on which bodies are ejected. This result, in turn, affects theoretical descriptions of the ejection process (see Section 5.3.2). The full theoretical interpretation is also complicated by effects such as secular resonances and vZLK resonances, which we discuss in Section 5.4.

Although Equation (5.1) fits the numerically determined lifetime function quite well, we note a few shortcomings. First, the sparse statistics in the number of long-lived objects (longer than, say, 100 Myr) makes the tail of the distribution somewhat uncertain. Although it is possible that the power law decay continues over even longer times than those explored here, it is also possible that the survival fraction flattens out in the limit $t \to \infty$. In this case, we are underestimating the number of permanently captured objects. On the other hand, the large uncertainty for long lifetimes also allows the survival function to be somewhat steeper than the best fit power-law (with slope $\beta = 8/5$).

Next we note that the ensemble of objects considered here represents all of the objects captured in the numerical exploration of Paper I. As a result, this sample does not necessarily correspond to the starting conditions that would result from a particular capture scenario. On one hand, the starting ensemble of captured objects determines their initial distribution of energies, which affects their lifetimes. On the other hand, the velocity scale v_σ in the cross section determines the typical speed of captured objects for bodies sourced from an initial velocity distribution with relatively high speeds (e.g., from the field).[1] In addition, we find that the survival time is largely independent of the speed v_∞ that characterizes the initial (pre-capture) orbit of the incoming bodies. Figure 5.3 shows the numerically determined dynamical lifetimes as a function of the initial hyperbolic velocity v_∞. Although the velocity bins at extremely low speeds ($v_\infty \lesssim 0.1$ km/s) show a wider distribution of lifetimes, the distributions (and their means) are largely constant over the range of incoming speeds (this result is consistent with Lehmann et al. 2021).

[1] The capture rate is determined by the integral $\langle v\sigma \rangle$. If we consider capture from the field, for example, the integrand is proportional to $v_\sigma v^2 \exp[-v^2/2s^2]$, where s is the velocity dispersion. In the limit $s \gg v_\sigma$, the integrand peaks near $v = v_\sigma/\sqrt{3}$ and thus the integral has most of its support near $v \sim v_\sigma \sim 0.42$ km/s.

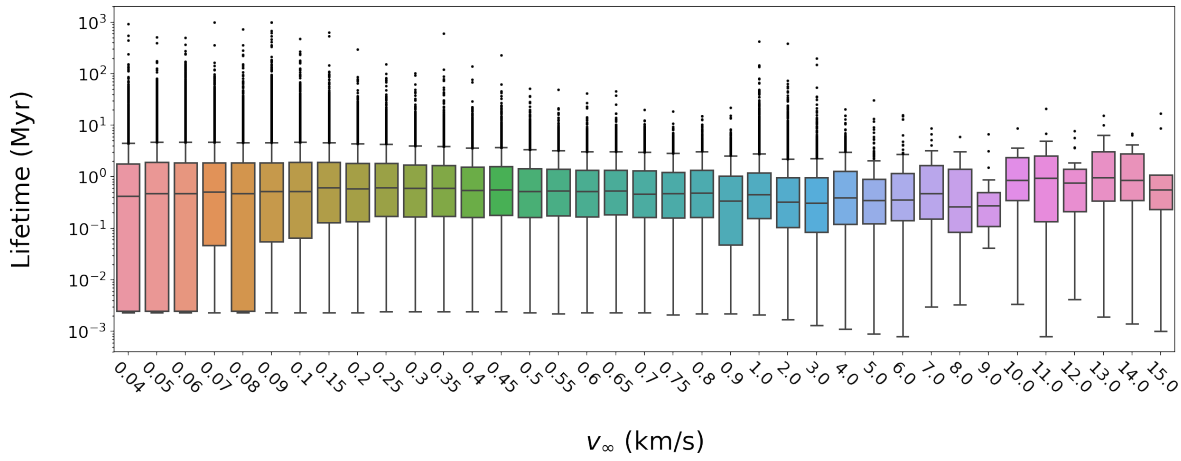


Figure 5.3: Box-and-whisker plots of dynamical lifetime as a function of hyperbolic excess velocity, v_∞. There is not a noticeable correlation between v_∞ and dynamical lifetime (and this vindicates our decision to simply integrate the objects from Paper I). Note that the data is not binned linearly in v_∞.

5.3.2 Diffusion Approach

The simulations show that most captured objects survive for a time scale of order 1 Myr, and that the population steadily decreases at later times. It is useful to have a theoretical description of how long it takes interactions between the rocky bodies and the planets to increase the orbital energy to an unbound state. The surviving fraction depends on time according to $f \sim 1/t^\beta$, with $\beta = 8/5$, for times spanning $t = 1$ Myr to 1 Gyr. Most of the captured objects have large semimajor axes, $a \sim 10^3 - 10^4$ au, with periheila falling in the realm of the giant planets. The captured rocks spend most of their time at large distances from the Sun, but periodically enter the giant planet region where their orbits are perturbed. In the regime of interest — long survival times — the orbital periods are much shorter than the lifetimes, so that many perturbations (perihelion passages) are generally required for ejection. Moreover, given the four giant planets and the wide range of orbital elements (a, e, i) for the captured bodies, the perturbations are largely uncorrelated. To leading order, one expects a random walk in the energy of the orbits, with the perihelion and inclination angle held nearly constant (Duncan et al., 1987; Dones et al., 1996). The rocks become unbound when the energy becomes positive.

While the literature does not contain a complete exploration of the dynamical lifetime of captured interstellar objects, previous works have attempted to characterize the dynamical lifetime of comets. In particular, Yabushita (1980) studied the survival of long-period comets under the approximation that comets undergo one-dimensional diffusion in energy, and found an analytic solution for $f(t)$ in the form of an incomplete gamma function. However, later papers (e.g., Duncan et al. 1987; Dones et al. 1996; Malyshkin & Tremaine 1999) point out that the diffusion approximation (which assumes that the perturbations are truly random) cannot capture resonance sticking and related effects. In this spirit, Malyshkin & Tremaine (1999) study the the dynamical lifetime using a Keplerian map, including a rather complicated interpolation function that describes the energy kicks experienced by a comet at perihelion as a function of the perturbing planet's phase (a *kick function*). The $f(t)$ functions computed in this previous paper, while correct in context, are limited (for the present application) because they were computed using the three-body approximation with a narrow range of test particle inclinations. Due to the complexity of our (current) numerical simulations, which use all four giant planets and consider the full range of orbital inclination, the corresponding kick function would be considerably more complicated. This complexity, in turn, makes a random (diffusion) description more applicable. Although the challenge of constructing an analytic theory for the survival fraction $f(t)$ remains for future work, here we briefly describe results from the diffusion approximation.

The dimensionless form of the diffusion equation can be written in the form

$$\frac{\partial n}{\partial t} = D \frac{\partial^2}{\partial x^2} \left(x^{3/2} n \right) \ , \tag{5.2}$$

where x is the absolute value of the dimensionless energy (so that objects become unbound as $x \to 0$). The factor of $x^{3/2}$ arises from Kepler's law, i.e., orbits with lower energy have larger semimajor axis $a = 1/x$, longer periods $P \sim x^{-3/2}$, and thus diffuse more slowly. Note that with the orbital time accounted for, the dimensionless diffusion parameter D is a constant (see Yabushita 1980 for further discussion).[2] Yabushita (1980) derives the Green's function for the

[2]Note that the diffusion field n appearing in Equation (5.2) represents the distribution of energies. Equivalently, one can define the quantity $\nu(x, t) = x^{3/2} n(x, t)$, which represents the distribution of bodies passing through

diffusion equation (5.2) and finds the solution for $n(x, t)$ corresponding to an initial distribution of energies $\delta(x - x_0)$, where x_0 is the starting dimensionless energy for all of the orbits. The resulting solution for the total number of surviving bodies $N(t)$ becomes

$$N(t) = \int_0^\infty n(x, t)dx = \gamma[2, 4\sqrt{x_0}/(Dt)] \rightarrow \frac{8x_0}{(Dt)^2}\,, \tag{5.3}$$

where $\gamma(2, \xi)$ is an incomplete gamma function (Abramowitz & Stegun, 1972). The final expression shows the survival probability in the asymptotic limit $t \rightarrow \infty$. The full form of the function is shown as the red dotted curve in Figure 5.2. Although the asymptotic slope is somewhat steeper than the best fit to the numerical data, this straightforward result provides a good working description.[3]

5.4 Analysis of long-lived objects

In this section we more closely analyze the secular dynamics of the captured objects in order to uncover the effects governing the long-term evolution of their orbits. Malyshkin & Tremaine (1999) found that the tail of their lifetime function was influenced by comets being locked into mean motion resonance with a giant planet. However, as previously discussed, this work considered the three-body problem with low inclinations. In the present problem, the bodies under consideration are on planet-crossing orbits in the presence of all four giant planets and explore a wide range of orbital inclination. These complications make it more difficult for a body to be captured into or remain in a mean motion resonance. Though we did not explicitly test our simulations for mean motion resonance, we did not note any obvious signs when examining the orbital elements of long-lived objects.

A notable feature of our data is that none of the prograde captured objects survived the full Gyr-integration (although one object came close). Three retrograde objects survived the

perihelion per unit time. The scaled function $\nu(x, t)$ obeys a diffusion equation with the $x^{3/2}$ factor appearing outside the spatial derivative operator (compare Malyshkin & Tremaine 1999 and Yabushita 1980).

[3]For completeness, we note that the inclusion of resonance sticking (Malyshkin & Tremaine, 1999) predicts a survival function $f(t) \sim t^{-p}$, where the power-law index is smaller than the numerical results. Specifically, $p \approx 2/3$ for the standard case and $p \approx 4/3$ if orbits are considered to become unbound with a small but negative energy.

full simulation, and all three had similar inclinations ($i \sim 120^\circ$). Though these statistics are not robust enough to make a definitive statement, we note an apparent trend in Figure 5.1. The objects that can achieve and maintain inclinations $i \sim 60^\circ$ with respect to the ecliptic (or invariable) plane are more likely to survive for the age of the solar system. This trend suggests that a common mechanism is driving the dynamical evolution of the longest-lived objects. Motivated by this finding, we investigate our data for signatures of vZLK resonances by plotting pericenter distance (q) vs argument of pericenter (ω) in Figure 5.4. We note that as dynamical lifetime increases, features resembling vZLK dynamics become more prominent.

Given the numerical results, we now investigate the possible presence of vZLK dynamics analytically. The Hamiltonian for this system can be written as

$$\mathcal{H} = \mathcal{H}_0 + \mathcal{P} \tag{5.4}$$

where $\mathcal{H}_0$ is the unperturbed Keplerian part of the Hamiltonian, and $\mathcal{P}$ accounts for planetary perturbations. By limiting $\mathcal{P}$ to first order in planetary mass, one can average over the mean anomaly of both the test particle and planetary orbits. This averaging has the effect of fixing the semi-major axes of all bodies in the problem, meaning that we can simply drop the $\mathcal{H}_0$ term from the Hamiltonian, as it becomes a constant offset. Then, up to this constant offset, the Hamiltonian becomes the sum of doubly-averaged planetary perturbations,

$$\mathcal{K} = \frac{1}{4\pi^2} \sum_k m_k \int_0^{2\pi} \int_0^{2\pi} \mathcal{P}_k(M_k, a, e, i, \omega, \Omega, M)\, dM_k dM. \tag{5.5}$$

In this expression the index k runs through the giant planets, masses (m_k) are in units of solar masses, and distances are in units of au. By making the approximation that the planets are on circular orbits with zero inclination, $\mathcal{K}$ becomes rotationally invariant about Ω (for definiteness, we set $\Omega = 0$). Thomas & Morbidelli (1996) showed that after evaluating the integral with respect to M_k, the function takes the form

$$\mathcal{K} = \frac{1}{4\pi^2} \sum_k m_k \int_0^{2\pi} \frac{4}{\sqrt{r^2 + a_k^2}} \sqrt{1 - \frac{\xi}{2}} K(\xi)\, dM\,, \tag{5.6}$$

where

$$\xi = \frac{4a_k\sqrt{x^2+y^2}}{r^2+a_k^2+2a_k\sqrt{x^2+y^2}}\,. \tag{5.7}$$

Here, the coordinates x and y denote the test particle position in the plane of the planets, r is its heliocentric distance, and $K(\xi)$ is the complete elliptic integral of the first kind[4]. While there is not a general closed-form analytic solution for $\mathcal{K}$, the expression can be integrated numerically for any appropriate values of the variables a, e, i, and ω.

In Figure 5.5, we plot the level curves of $\mathcal{K}$ in the parameter space of q versus ω by holding fixed a and $H = \sqrt{1-e^2}\cos i$. In each panel, we overlay in red the simulation data of the object used to generate the level curves. The theoretical curves match the numerical data rather well, indicating that for these objects the vZLK effect provides a good working description of their long-term dynamics. We find that this theoretical framework matches the numerical data for 11 of the 13 objects that survived for more than 500 Myr. We do not expect this theory to hold for objects that are close to mean-motion resonances (in which case one would have to consider extra terms in the Hamiltonian to account for resonance), or those that experience close encounters with the giant planets (in which case their a and H values are not well-approximated as constant). The interpretation that emerges here is that the vZLK mechanism acts as an effective phase-protection mechanism. Objects that can attain $q \sim 10$ au execute vZLK cycles and return to pericenter only at $\omega \sim 90^\circ$ or 270°. This geometry implies that they come to pericenter directly above and below the plane, so they are shielded from encounters with the giant planets.

5.5 Population Estimates for Captured Alien Rocks

This section utilizes the newly derived lifetime function to estimate the current population of alien rocky bodies in our solar system (i.e., rocks that have been captured from other planetary systems). We thus refine our earlier estimates from Paper I. Note that rocks can be captured by

[4]To avoid ambiguity caused by the discrepancy between the notation used for function tables and numerical implementations, we give the explicit form $K(\xi) = \int_0^{\pi/2}(1-\xi^2\sin\theta)^{-1/2}\,d\theta$. Note that the modulus here is different than the one used in Abramowitz & Stegun (1972).

the solar system during two regimes of interest: early times when the Sun still resides in its birth cluster and later times when the Sun resides in the field.

In Paper I, we estimated that the total mass in rocky bodies captured during the cluster phase is approximately $M_{R0} \approx 10^{-3} M_\oplus$. Note that the residence time of the solar system in its birth cluster is expected to be much shorter than its current age (Adams, 2010). As a result, all of the rocks captured from the cluster have approximately the same chances of expulsion, so that the mass in captured rocks remaining at the present epoch is given by

$$M_{RC} = M_{R0} f(t_{\rm age}) \approx M_{R0} \left(\frac{\tau}{t_{\rm age}} \right)^{8/5} \approx 10^{-9} M_\oplus \approx 6 \times 10^{18} \, {\rm g} \, , \tag{5.8}$$

where $f(t)$ is the fraction of captured objects remaining as a function of time and $t_{\rm age}$ is the age of the solar system.[5]

In addition to the contribution from the birth cluster, the solar system will steadily accumulate additional rocky bodies from the field. These newly added rocks are continually ejected, so that this population will reach a steady state. In the field, the solar system captures rocky objects at a rate given by

$$\Gamma = n_R \langle \sigma v \rangle \, , \tag{5.9}$$

where n_R is the number density of rocky bodies in the solar neighborhood and the velocity-averaged cross section $\langle \sigma v \rangle$ is evaluated for local conditions. Over most of the age of the solar system, the parameters that define the capture rate (5.9) are expected to be constant in time. Here, the cross section as a function of incoming speed has the form derived in Paper I,

$$\sigma(v) = \frac{\sigma_0}{(v/v_\sigma)^2 [1 + (v/v_\sigma)^2]^2} \, , \tag{5.10}$$

where $\sigma_0 \approx 232,000$ AU^2 and $v_\sigma \approx 0.42$ km/s. If we combine this cross section with a Maxwellian distribution of velocities with dispersion s, we find the approximate result

$$\langle \sigma v \rangle \approx \frac{\sigma_0 v_\sigma}{\sqrt{2\pi}} \left(\frac{v_\sigma}{s} \right)^3 \, , \tag{5.11}$$

[5] Note that this scale is ~ 1000 times less massive than the Earth's atmosphere.

where this expression has been evaluated in the limit $s \gg v_\sigma$ (see Paper I). The velocity dispersion of stars in the solar neighborhood is ~ 40 km/s (Binney & Tremaine, 2008). Since the value of s corresponds to the relative speed between bodies, we use $s \approx \sqrt{2}(40) \approx 57$ km/s for the estimates of this paper.[6]

Although the size distribution and number density of rocky material in the interstellar medium are not known, we can make the following estimate. The rocky bodies must originate from planetary systems. Moreover, each planetary system is expected to eject a given mass m_1, where we expect $m_1 \approx$ few $\times M_\oplus$ (Rice & Laughlin, 2019). The mass density in rocky bodies is thus given by $\rho_R = m_1 n_*$, where n_* is the number density of stars. The mass accretion rate of the solar system can be found by combining the above results:

$$\dot{M} = \rho_R \langle \sigma v \rangle = m_1 n_* \frac{\sigma_0 v_\sigma}{\sqrt{2\pi}} \left(\frac{v_\sigma}{s}\right)^3 . \tag{5.12}$$

The analysis of the previous section shows that after rocks are captured, the fraction of the objects remaining in bound orbits is a decreasing function of time $f(t)$. The steady-state mass contained in (initially alien) bound objects is thus given by

$$M_{SS} = \int_0^{t_{\rm age}} \dot{M} dt f(t_{\rm age} - t) = \int_0^{t_{\rm age}} \frac{\dot{M}}{1 + (t/\tau)^\beta} dt \,, \tag{5.13}$$

where $t_{\rm age}$ is the age of the solar system. Using $\beta = 8/5$ and changing variables ($u = t/\tau = z^5$), the expression becomes

$$M_{SS} = 5\dot{M}\tau \int_0^Z \frac{z^4 dz}{1 + z^8} \to \frac{5\pi}{2\sqrt{2\pi}} \frac{\cos(\pi/8)}{2 + \sqrt{2}} m_1 n_* \sigma_0 v_\sigma \tau \left(\frac{v_\sigma}{s}\right)^3 . \tag{5.14}$$

The final expression can be obtained by contour integration and by taking the limit $Z \to \infty$, which is valid in the regime where $t_{\rm age} \gg \tau$ (note that $t_{\rm age}/\tau \approx 5500$, so that the correction is of order $\sim 0.5\%$). If we evaluate this expression using typical parameter values for the solar

[6]For completeness, we note that the solar system's current velocity with respect to the local standard of rest, $v_{\rm LSR}$, is somewhat slower than 40 km/s, and could have been even smaller in the distant past. If we use a conservative value $v_{\rm LSR} = 20$ km/s, the velocity dispersion becomes $s \sim 45$ km/s. Using this value of s increases our estimate of the steady state mass of captured rocks in the field by a factor of $(57/45)^3 \approx 2$. This correction factor is less than other uncertainties inherent in the calculation.

neighborhood ($n_* = 0.1$ pc^{-3} and $s = 57$ km/s; Binney & Tremaine 2008), we find

$$M_{SS} \approx 7 \times 10^{-14} m_1 = 4 \times 10^{14}\,\mathrm{g} \left(\frac{m_1}{M_\oplus}\right) . \tag{5.15}$$

Comparing equations (5.8) and (5.15) we find that the mass in rocks captured from the solar birth cluster greatly exceeds the steady state mass captured from the field.[7]

Note that we can use the same procedure to estimate the steady-state mass of captured dark matter particles. The velocity of dark matter particles in the Milky Way's dark matter halo is assumed to follow a Maxwellian distribution with a dispersion of ~ 200 km/s, so that $s_{DM} \sim 200\sqrt{2}$ km/s.[8] Although the vast majority of the particles will be moving too fast to be captured, the low-velocity tail of the distribution provides an ample regime of parameter space for capture. In practice, the dark matter particles only interact gravitationally, and should therefore be captured onto orbits similar to those of the interstellar rocks in our simulations (both the rocks and the dark matter particles act as test particles). As a result, the formalism leading to Equation (5.14) remains applicable, so that the total mass of captured dark matter particles in steady-state is given by

$$M_{DM} \approx M_{SS} \times \frac{\rho_{DM}}{n_* m_1} \left(\frac{s_{rocks}}{s_{DM}}\right)^3 \approx 10^{17}\,\mathrm{g}\,, \tag{5.16}$$

where we have used $\rho_{DM} = 0.43$ GeV cm^{-3} for the dark matter density (e.g., Salucci et al. 2010; Read 2014). It is interesting, though not consequential, that this value for dark matter capture is close to the geometric mean of the (baryonic) mass captured in the field and in the birth cluster.

Finally we consider how the steady state mass of captured objects compares to the instantaneous mass of interstellar rocks passing through our solar system without being captured. Rocks in the field have a typical hyperbolic excess velocity that is larger than the escape speed of the solar system (unless $r \lesssim 1$ au) so this estimate can neglect gravitational focusing by the Sun. In this approximation, the mass of interstellar material passing through the solar system is given by $M_{\rm int} = \rho_{\rm int} V_{\rm SS}$. Here, the mass density of interstellar objects $\rho_{\rm int} \approx 0.1$ M$_{\odot}$ pc^{-3}. The effective

[7] Note that this scale is ~ 1000 times less massive than Mount Everest.

[8] Note that this estimate may change if the Milky Way's dark matter halo is rotating with respect to the inertial frame.

volume of the solar system depends on the chosen outer boundary R, such that $V_{SS} = (4\pi/3)R^3$. Since the orbit of Neptune defines the edge of the solar system for many applications, we start by taking $R = 30$ au, and find the total interstellar mass $M_{int} \approx 10^{16}$ g. This value is about 600 times smaller than the mass contained in the solar system's inventory of captured objects. However, the interstellar objects that remain in the solar system for an appreciable time tend to be captured with semi-major axis $a \sim 1000$ au and perihelion $q \sim 10$ au. Such objects spend only about 0.5% of their time inside 30 au. As a result, the mass of captured objects within 30 au and the mass of unbound objects passing through that same volume of the solar system are roughly comparable.

5.6 Discussion and Conclusions

In this paper we have examined the long-term stability of interstellar objects captured by our solar system. Using an ensemble of numerical simulations for 276,691 synthetic captured interstellar objects, the main result of this study is the determination of a dynamical lifetime function, i.e., the fraction of objects remaining bound as a function of time (Figure 5.2). The resulting lifetime function shows that — after a few million years — the system reaches a state where the survival probability approaches a power-law form such that $f(t) \sim t^{-1.6}$. The lifetime function $f(t)$ over the entire range of time can be fit with the function given in Equation (5.1). These results are broadly consistent with a one-dimensional diffusion model (Yabushita, 1980), although the latter produces a somewhat steeper power-law $f(t) \sim t^{-2}$ (see Section 5.3.2 and Figure 5.2).

This survival fraction $f(t)$ can be used in combination with the capture cross section (see Paper I) to estimate the number of objects of interstellar origin trapped in our solar system (Section 5.5). The mass in alien rocks remaining from capture events in the solar birth cluster ($M_{RC} \sim 10^{-9} M_{\oplus}$) is much larger than the steady-state mass due to capture from the field ($M_{SS} \sim 10^{-13} M_{\oplus}$). The estimated (steady-state) mass in dark matter particles has an intermediate value ($M_{DM} \sim 10^{-11} M_{\oplus}$; see also Lehmann et al. 2021). These results should also be useful in understanding phenomena such as panspermia (see, e.g., Melosh 2003, Adams & Spergel 2005).

We also analyze the dynamical evolution of the captured objects (Section 5.4). We find that not all kinds of orbits are equally likely to survive for long times. Objects that can attain pericenter distances beyond Jupiter (and preferably beyond Saturn as well) are preferentially longer-lived. The numerical data suggest that the objects that can attain and maintain high inclinations and/or become locked into vZLK oscillations are more likely to avoid close encounters with the giant planets, and thus remain in the solar system for longer times. Nonetheless, it would be useful to have additional long-term integrations to provide better statistics.

This work has some limitations, many of which could be improved with a considerable amount of computing. The most apparent shortcoming is that the tail of our lifetime function is rather sparsely populated. As only 13 objects survived for more than 500 Myr, and only 3 objects survived for the entire 1 Gyr integration, it is not clear whether the power-law lifetime continues indefinitely. We also reiterate that this work used the synthetic captured interstellar objects from Paper I as a starting point, so it is possible that the v_∞ distribution of the captured objects has a different form than assumed herein. However, the distribution of ejection times is largely independent of v_∞ (see Figure 5.3), so that this issue is higher order (see the discussion of Section 5.3.1). These complications, combined with uncertainties in the amount of time our solar system spent in the birth cluster versus in the field, and the velocity dispersion of the ambient rocks in each setting, make our estimate of the steady-state population of captured interstellar objects uncertain as well. Another direction for future work is to incorporate galactic tides into the numerical simulations. The objects tend to be captured onto extremely elliptical and distant orbits, and spend much of their time beyond a few thousand au where galactic (or cluster) tides are no longer negligible.

As a final comment, we note that this study did not account for the hypothesized Planet Nine (see, e.g., Batygin et al. 2019). Since the capture process explored in Paper I relies on an initially hyperbolic orbit entering a planet's sphere of influence (and the Planet Nine sphere of influence is minuscule compared to the volume of its orbit), the inclusion of Planet Nine would have little effect on the capture cross section. However, since objects are typically captured with

semi-major axes of order $\sim$ 1000 au, they spend most of their time in the trans-Neptunian and Inner Oort Cloud regions where the secular effects of Planet Nine become important. To get a sense for the role that Planet Nine plays in the orbital evolution of our captured objects, we performed a set of simulations on $\sim$ 40,000 of our captures in which we included a 10 Earth-mass body with $a = 500$ au, $e = 0.6$, and $i = 18^\circ$ at a random phase. These simulations yielded rich dynamics that did not appear in the simulations including only the four known giant planets. While no trans-Neptunian objects (TNOs) were produced in our simulations with only the four known giant planets (though we emphasize again that they may have been generated if we had included galactic or cluster tides), our simulations including Planet Nine produced several such cases. Some objects displayed the well-understood behavior in which Planet Nine causes their orbital planes to tilt, sometimes even switching between prograde and retrograde. Finally, we note that some objects in these simulations were able to achieve osculating pericenter distances of up to 600 au (i.e., Planet Nine injected them into the Inner Oort Cloud). It is worth noting that this result is essentially the inverse of the process discovered by Batygin & Brown (2021) in which Planet Nine injects Inner Oort Cloud objects into the trans-Neptunian region of the solar system, and may warrant further investigation.

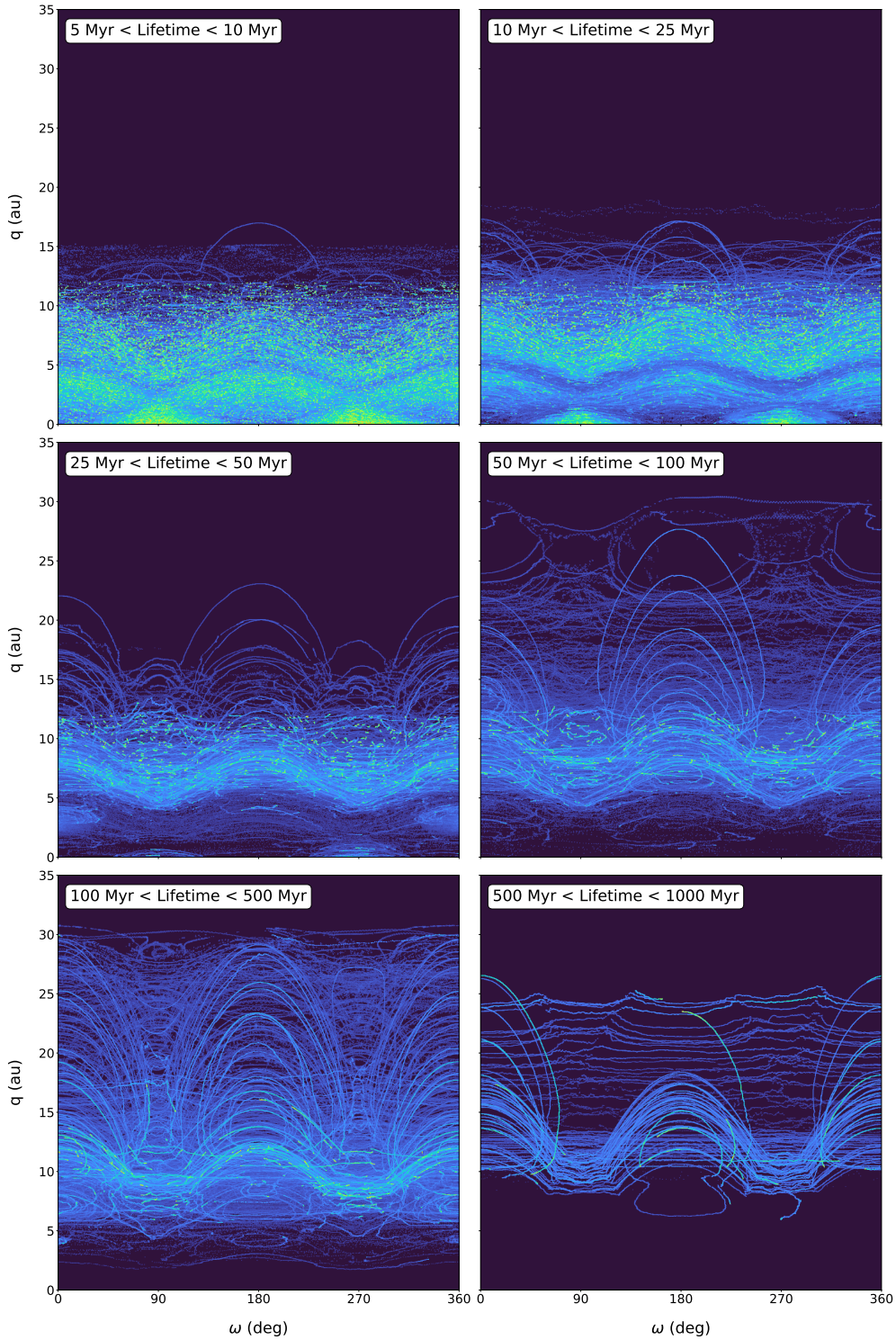


Figure 5.4: Two-dimensional histograms of q vs ω for the captured objects, grouped into lifetime ranges. The color gradient corresponds to the point density. At long times, it becomes apparent that the captured objects are locked into vZLK oscillations (both librating and circulating).

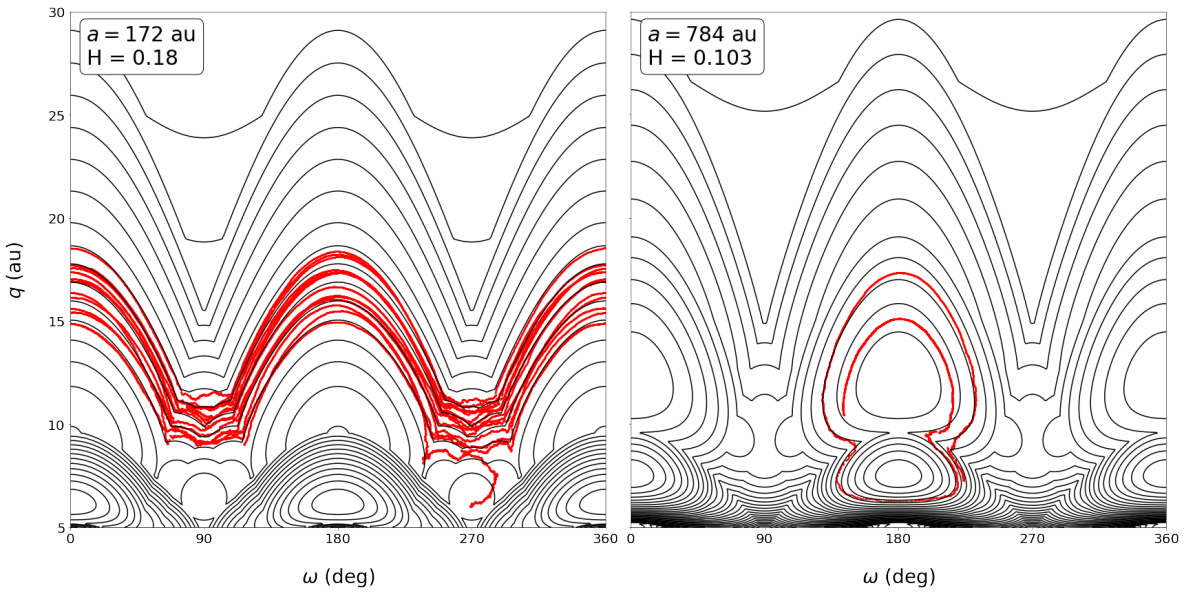


Figure 5.5: (Black) Hamiltonian level curves in the parameter space of q vs. ω for two objects that survived for more than 500 Myr. The level curves were computed by holding constant the object's median semi-major axis a and reduced action $H = \sqrt{1-e^2}\cos i$, as indicated in the upper-left of each panel. (Red) Numerical data from our simulations. The theory and numerics match remarkably well for these objects, suggesting that their dynamics are well-described by a simple vZLK model.

Chapter 6

A Collision Mechanism for the Removal of Earth's Trojan Asteroids

When the great crisis comes,
our daring and our resolution vanish;
our soul is agitated, paralyzed;
and we run around the walls
seeking to save ourselves in flight.
Nevertheless, our fall is certain.

— Constantine P. Cavafy, *Trojans*

This chapter is published as Napier et al. (2022).

6.1 Introduction

The restricted circular three-body problem predicts five points of equilibrium called the Lagrange points. Three of these points (L1, L2, and L3) are colinear with the primary and secondary masses, and the remaining two (L4 and L5) lead and trail the secondary mass by 60 degrees respectively. The three colinear points are all saddle points of the potential, and are therefore unstable. However, L4 and L5 are points of linearly stable equilibrium (as long as the secondary/primary mass ratio $\mu < 0.039$, see Murray & Dermott 2000), and as such can host long-term stable populations colloquially known as Trojan asteroids.

Because Trojan asteroids are in the 1:1 mean-motion resonance with their host planet, they can remain stable for billions of years. As a result, they have the potential to be of primordial origin and can provide important probes of the solar system's dynamical and chemical history. The Jupiter Trojans are of considerable interest, as their quantity and chemical composition may encode sensitive information about the composition of the protoplanetary disk and the migration history of the giant planets. Understanding this population is of such high importance that NASA's Lucy spacecraft will fly by seven Jupiter Trojans over the next decade, four of which are in binary systems (Levison et al., 2021). Neptune Trojans encode similar information and are now being studied in detail, with characterization efforts including a Cycle 1 James Webb Space Telescope program to obtain near-infrared reflectance spectra (Markwardt et al., 2021).

Only two Earth Trojans (ETs) have been found to date, and both are transient; simulations show that they are stable over time scales much shorter than the age of the solar system (Connors et al., 2011; Hui et al., 2021). The prospect of discovering long-term stable ETs is enticing, as their proximity and low delta-v with respect to Earth would make a spacecraft mission imminently possible. Additionally, it would likely be possible to return samples of ETs. This feature would be extremely valuable for studying the composition of the protoplanetary disk if the objects are indeed of primordial origin. If such Trojans are not of primordial origin, the samples would still be useful for gleaning insight about the reservoirs that populate near-Earth orbits. In light of the potential scientific payoff of discovering ETs, several surveys have been dedicated to searching for such objects (Whiteley & Tholen 1998; Cambioni et al. 2018; Yoshikawa et al. 2018; Markwardt et al. 2020; Lifset et al. 2021). While these searches have failed to discover any long-term stable ETs, they have placed significant observational upper limits on the possible population. Specifically, the number of ETs with sizes larger than approximately 400 meters is limited to at most 100 objects (Markwardt et al., 2020; Lifset et al., 2021).

These observations are seemingly in tension with numerical simulations showing that ETs can be stable over Gyr timescales (see, e.g., Tabachnik & Evans 2000; Ćuk et al. 2012; Marzari & Scholl 2013; Zhou et al. 2019; Malhotra 2019; Christou & Georgakarakos 2021). Furthermore,

while it has been shown that the Yarkovsky effect (Bottke et al., 2006; Marzari & Scholl, 2013; Zhou et al., 2019) can remove the small ETs (smaller than tens of kilometers) from resonance, the effect is not strong enough to remove the larger bodies on time scales comparable to the age of the solar system. Current theories therefore suggest that long-term stable ETs can exist. To date, any such object has eluded discovery.

Although the existence of any large ETs has not been ruled out by observational surveys, the continued absence of even a single long-term stable ET merits the consideration of a gap in our understanding of their dynamics. In this work we consider how collisions between asteroids and the Earth, such as those hypothesized to have occurred during the Moon-forming impact (see, e.g., Canup 2012; Ćuk & Stewart 2012) or while Earth was acquiring its Late Veneer (see, e.g., Kleine 2011; Raymond et al. 2013; Brasser et al. 2016), would affect an initially-stable ET population. In Section 6.2 we derive the circumstances of the collisions between the Earth and the asteroids that may have been responsible for delivering its Late Veneer. In Section 6.3, we describe our numerical experiments in which we simulated the bombardment of Earth by large asteroids. In Section 6.4 we present the results of our simulations and analyze their outcomes. In Section 6.5, we summarize our results and interpret them in the broad context of the possible Earth Trojan population. For completeness, Appendix 6.A provides a quantitative assessment of some of our approximations and Appendix 6.B presents an order of magnitude estimate for the orbital changes required to remove Trojans from bound states. Finally, we consider perturbations of Earth's orbit due to close encounters that do not result in collisions (Appendix 6.C) and direct collisions between incoming asteroids and Trojans (Appendix 6.D).

6.2 Analytic Theory

In this section we compute the distributions of relative velocity and orientation for potential Earth impactors (i.e., asteroids on Earth-crossing orbits). We use the resulting distributions to compute the change in Earth's momentum for our simulations of asteroid-Earth collisions in Section 6.3.

6.2.1 Scales and Ordering

Consider an asteroid on an Earth-crossing orbit. Here the asteroid has mass m, the Earth has mass $M_\oplus$, and the Sun has mass $M_\odot$. The masses obey the ordering

$$m \ll M_\oplus \ll M_\odot \,, \tag{6.1}$$

so that the orbits can be considered in the hierarchical limit. Using the mass of Earth, the mass ratio $\mu = M_\oplus / M_\odot \sim 3 \times 10^{-6}$.

As a first approximation, we take Earth to have a circular orbit with semi-major axis $a_\oplus = 1$ au. The potential impactor has initial orbital elements (a, e) and is considered to have the same orbital plane as Earth ($i = 0$). The Hill sphere of Earth is given by

$$R_H = a_\oplus \left(\frac{M_\oplus}{3 M_\odot} \right)^{1/3} \approx 0.01 \text{ au}, \tag{6.2}$$

whereas the Sphere of Influence has the form

$$R_S = a_\oplus \left(\frac{M_\oplus}{M_\odot} \right)^{2/5} \approx 0.006 \text{ au}. \tag{6.3}$$

Since the orbital length scales obey the ordering

$$R_S \sim R_H \ll a_\oplus < a$$

we can use intermediate asymptotics to describe the dynamics of the bodies that collide with Earth. In other words, the orbit of the asteroid has a Keplerian orbit around the Sun until it reaches the vicinity of Earth. Once the asteroid enters the Earth's sphere of influence, its subsequent dynamics are governed by the gravitational potential of Earth. We can thus find the velocity vector of the asteroid at the point where it crosses the orbit of Earth. This velocity vector then provides the outer boundary condition for the inner problem where the asteroid collides or interacts with Earth.

6.2.2 Relative Velocity for Earth Encounters

We now calculate the speed of the asteroid with respect to Earth at the point of orbit crossing. The asteroid orbit has specific energy and angular momentum given by

$$E = -\frac{GM_\odot}{2a} \qquad \text{and} \qquad J^2 = GM_\odot a(1-e^2) \tag{6.4}$$

The azimuthal velocity component of the asteroid orbit is determined by conservation of angular momentum so that

$$v_{\phi\odot}^2 = \frac{J^2}{r^2} = \frac{GM_\odot a}{r^2}(1-e^2) \tag{6.5}$$

Note that the azimuthal velocity is defined here in the solar reference frame, as indicated by the subscript. The radial velocity of the asteroid with respect to the Sun is then given by

$$v_r^2 = 2E - \frac{J^2}{r^2} + \frac{2GM_\odot}{r} = \frac{2GM_\odot}{r} - \frac{GM_\odot a}{r^2}(1-e^2) - \frac{GM_\odot}{a}. \tag{6.6}$$

Because Earth is assumed to have a circular orbit, it has no radial velocity component, so that the radial velocity of the asteroid is the same in both reference frames. In this approximation, the speed of the asteroid with respect to Earth is thus given by

$$v_\infty^2 = v_r^2 + (v_{\phi\odot} - v_\oplus)^2 = v_r^2 + v_{\phi\odot}^2 + v_\oplus^2 - 2v_{\phi\odot}v_\oplus \tag{6.7}$$

where $v_\oplus = (GM_\odot/a_\oplus)^{1/2} \approx 30$ km/s is the orbital speed of Earth, and where all quantities are evaluated at $r = a_\oplus$.[1] Using the results obtained above we find

$$v_\infty^2 = \frac{GM_\odot}{a_\oplus}\left[3 - \frac{a_\oplus}{a} - 2\left(\frac{a}{a_\oplus}\right)^{1/2}(1-e^2)^{1/2}\right]. \tag{6.8}$$

We see that the orbital speed of Earth $v_\oplus$ sets the velocity scale for this collision problem. Let us define the dimensionless variables

$$u_\infty \equiv \frac{v_\infty}{v_\oplus} \qquad \text{and} \qquad \xi \equiv \frac{a}{a_\oplus} \tag{6.9}$$

[1] We have assumed in this calculation that the impactors are on prograde, bound orbits. If the orbits are bound and retrograde, then the minus sign in Equation (6.7) becomes a plus sign.

so that the expression for the hyperbolic speed at infinity reduces to the form

$$u_\infty^2 = 3 - \frac{1}{\xi} - 2\xi^{1/2}(1-e^2)^{1/2} \tag{6.10}$$

The largest possible speed is thus $u_\infty = \sqrt{3}$ or $v_\infty = v_\oplus\sqrt{3} \approx 52$ km/s.[2] This case corresponds to a distant asteroid orbit (large a) with small perihelion (eccentricity e close to unity) so that the asteroid velocity is nearly radial when it encounters Earth. For more typical cases, we expect perihelion $q = \xi(1-e) \sim 1$ in dimensionless units, and $\xi \sim 2$, so that $u_\infty \sim 0.2$ and $v_\infty \sim 6$ km/s. Note that the expressions derived above also apply for orbits with $a < a_\oplus$ (although such objects do not originate from the asteroid belt).

6.2.3 Distributions of Impact Variables

Using the results derived above, we can find the distribution of relative speeds by assuming a simple model for the initial population of asteroid impactors. We can also determine the distribution of angles at which the asteroids intercept the orbit of Earth. A great deal of previous work has quantified the distributions of speeds (and other properties) for near Earth asteroids for purposes of planetary defense (e.g., see Bottke et al. 2000; Jeffers et al. 2001; Steel 1998; and many others). These distributions are consistent with the model developed below (although the present-day and primordial distributions could be different).

Let us consider the case where the asteroids have a surface density distribution of the form $\sigma \propto 1/a$, so that the distribution of semi-major axes of the asteroids is uniformly distributed within some range, i.e.,

$$\frac{dN}{da} = constant \qquad \text{for} \qquad a_1 \leq a \leq a_2. \tag{6.11}$$

For the sake of definiteness, here we assume $a_1 = 1$ au and $a_2 = 5$ au, so we are sampling the full annulus between the current orbits of Earth and Jupiter. For a given value of a, the eccentricity

[2]Since we have assumed bound orbits, the incoming asteroids travel on elliptical orbits. In principle, incoming objects from outside the solar system could encounter Earth on hyperbolic orbits. This would allow for larger encounter speeds. Notice also that these types of interactions could either take place in the solar birth cluster or in the field. Such encounters are expected to be rare in the field, but could be more common in the birth cluster (see also Napier et al. 2021a).

must be larger than a minimum value given by

$$e_{min} = 1 - \frac{a_\oplus}{a}. \tag{6.12}$$

If this condition is not met, then the asteroid cannot cross the orbit of Earth. The eccentricity is unlikely to be much larger than that corresponding to an Earth crossing orbit—it takes interactions to pump up the eccentricity, and continued orbit crossings will eventually lead to large orbital perturbations. For the sake of definiteness, let us assume that the maximum eccentricity corresponds to perihelion of $a_\oplus/10$ so that

$$e_{max} = 1 - \frac{a_\oplus}{10a} \tag{6.13}$$

We then assume that the distribution of eccentricities is uniform over the interval $[e_{min}, e_{max}]$. With these choices for the distributions of the asteroid orbital elements (a, e), the resulting distribution of relative speeds v_∞ is shown in Figure 6.1.

We can also find the distribution of directions. Let θ be the angle of the incoming velocity (in the planetary reference frame) with respect to the velocity of Earth. In other words,

$$\cos\theta = \frac{\mathbf{v}\cdot\mathbf{v}_\oplus}{|\mathbf{v}||\mathbf{v}_\oplus|} = \frac{(v_{\phi\odot} - v_\oplus)}{v_\infty}. \tag{6.14}$$

The resulting distributions of both relative speed and angular direction are shown in Figure 6.1. These distributions are used as the starting point for the numerical simulations discussed in the following section. Note that the distribution of angles has two peaks, corresponding to bodies impacting Earth on incoming and outgoing trajectories. Moreover, the two parts of distribution, centered on $\theta = \pi/2$ and $3\pi/2$, each include both positive and negative values of $\cos\theta$. In Appendix 6.A we investigate the validity of the approximations made in this section.

6.3 Numerical Simulations

The primary contribution of this work is an ensemble of numerical experiments where we tested the stability of a synthetic population of long-term stable ETs while Earth was stochastically

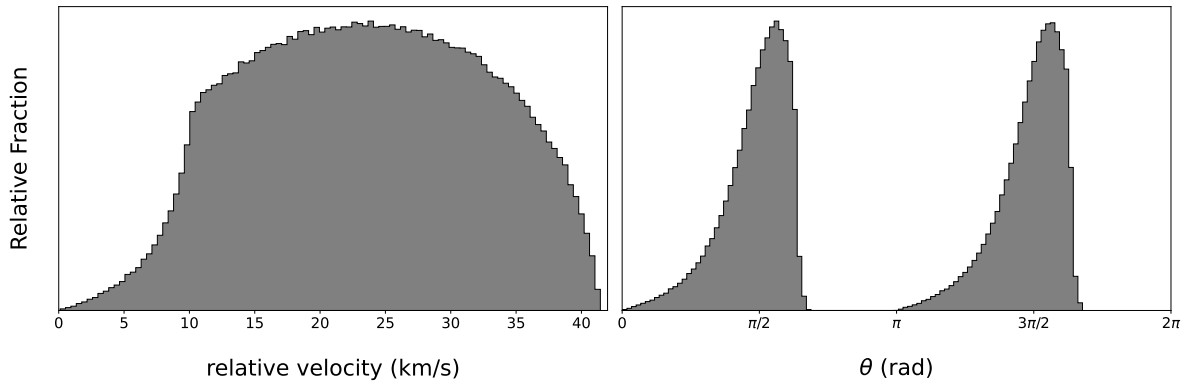


Figure 6.1: Distributions of incoming speeds (left) and incoming direction (right) for asteroids with orbits leading to close encounters with Earth. In both plots, the asteroid orbits have semi-major axes in the range a = 1–5 au, and a uniform distribution of eccentricity. Only those asteroid orbits with sufficiently large values of eccentricity e cross the orbit of Earth and are included in the distribution. The two peaks in the distribution of incoming direction are a result of the impacting bodies being equally likely to hit Earth on the incoming or outgoing trajectories.

bombarded by asteroids. For all of our numerical experiments, we used `Rebound`'s `WHFAST` integrator (Rein & Tamayo, 2015) and a model of the solar system that included the Sun, Venus, Earth, Mars, Jupiter, Saturn, Uranus, and Neptune.[3]

We began by generating a synthetic population of 3,021 ETs, each of which was dynamically stable for at least 10 Myr. This population represents the stable subset of an initial set of candidate ETs (which were integrated numerically to verify stability). The surviving objects were split roughly evenly between L4 and L5. We show the resulting population projected into the ecliptic plane in Figure 6.2. The dynamical properties of our synthetic population match those reported by Zhou et al. (2019). Note that due to the presence of other planets, these stable clouds are wider than one would expect by considering only the circular restricted three body problem (e.g., Murray & Dermott 2000).

Next we ran simulations to test the stability of our synthetic ETs as Earth underwent collisions with asteroids. We began each simulation by setting Earth's mass to 0.99 $M_\oplus$, and integrating for a warm-up period of 100 years with no collisions. After the warm-up period, we perturbed Earth

[3]We found that our results were not changed by resolving the Earth-Moon system. In the interest of saving computation time, we chose to use the Earth-Moon barycenter.

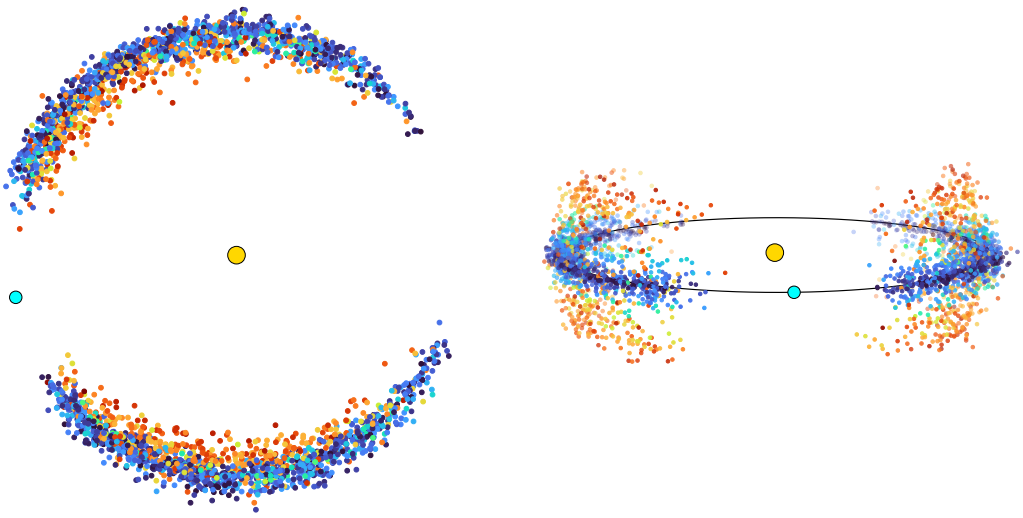


Figure 6.2: Our synthetic population projected into the ecliptic plane (left) and viewed from a higher inclination (right). The large cyan circle represents Earth, the large yellow circle represents the Sun, and the test particles are represented by smaller circles with colors corresponding to their inclinations (blue represents low inclination, and red represents high inclination, maxing out at approximately 40°).

stochastically by modifying its velocity as it would change due to a collision with an asteroid.[4] We randomly drew the impactor velocity and impact angle from the distributions shown in Figure 6.1. We chose the impactor mass by drawing a radius R from a distribution of the form $N(R) \propto R^{-2}$ with values ranging from 500 km to 2000 km, and assuming a mass density $\rho = 3$ g cm^{-3}. While there are certainly many asteroids smaller than 500 km, they will deliver a negligible amount of mass to Earth compared to the mass delivered by a few large objects. Once we had perturbed Earth, we integrated for 10,000 more years (we found that this was enough time for the system to roughly settle back into equilibrium), tracking whether each test particle remained a Trojan, migrated onto a horseshoe orbit, or was ejected from a co-orbital state altogether. If an object was no longer a co-orbital with Earth, we removed it from the simulation. We continued this procedure until Earth's mass reached 1 $M_\oplus$.[5] We repeated this process a total of 5,000 times

[4] For simplicity, we assumed a totally inelastic collision so that the entire mass of the impactor was added to Earth's mass. The velocity perturbations were calculated such that linear momentum was conserved, and were applied instantaneously.

[5] If a collision would have resulted in Earth's mass exceeding 1 $M_\oplus$, we simply reduced the mass of the impactor such that Earth's mass became exactly 1 $M_\oplus$.

to build up our statistics. The numerical values used in these simulations were all chosen to be roughly consistent with the circumstances of Earth's Late Veneer impacts reported in the literature (see, e.g. Raymond et al. 2013, Brasser et al. 2016, and the references therein).

6.4 Analysis

In this section we analyze the data from our simulations. We begin by examining the probabilities with which initially-Trojan objects remained on true Trojan orbits, or were perturbed onto horseshoe orbits. In Figure 6.3, we show box-and-whisker plots of the fraction of the initial population of 3,021 ETs that ended the simulation in a Trojan or a horseshoe orbit, as a function of the number of impactors. Note that there are data points that exceed the vertical range we have chosen for Figure 6.3, but such events are relatively rare. After running many simulations an apparent trend emerges: for any number of impacts, the median surviving fraction of Trojans was less than a few percent. In addition, for a given number of impactors, the median number of objects surviving on horseshoe orbits was always on the order of 1-2 percent. These results indicate that large asteroid impacts with Earth are efficient at clearing the co-orbital population.

This series of collisions is a stochastic process in which different combinations of the number, velocities, and impact angles of the impactors can lead to vastly different outcomes. Indeed, there were some simulations in which an appreciable fraction of the objects remained on Trojan orbits, and as such require a more nuanced analysis. We find that all such simulations were instances where the number of impactors was small, and where all of them had low relative velocities. In these cases, the impactors simply did not impart enough momentum to Earth to remove the majority of the Trojan population. We demonstrate this trend in Figure 6.4 by plotting the surviving fractions of Trojan and horseshoe orbits as a function of the net linear momentum of the impactors. As the net linear momentum imparted by impactors increases, the probability of many objects remaining on Trojan orbits decreases. As the number of impactors increases, the probability of drawing all impactors with low velocities decreases. This decreases the probability of a small net linear momentum, and thus decreases the frequency and severity of outliers. Of

course, there is a limit to this logic. In the case that Earth gradually acquired its Late Veneer from millions of small rocks, the net linear momentum would converge to zero, and the Trojans would remain largely unaffected. In a similar vain, the most common outcome (by far) with 1–5 impactors was that no objects remained in Trojan orbits at the end of the simulation. This outcome occurs because with so few impactors, a large net linear momentum is far more probable than a small net linear momentum (see the distributions in Figure 6.1).

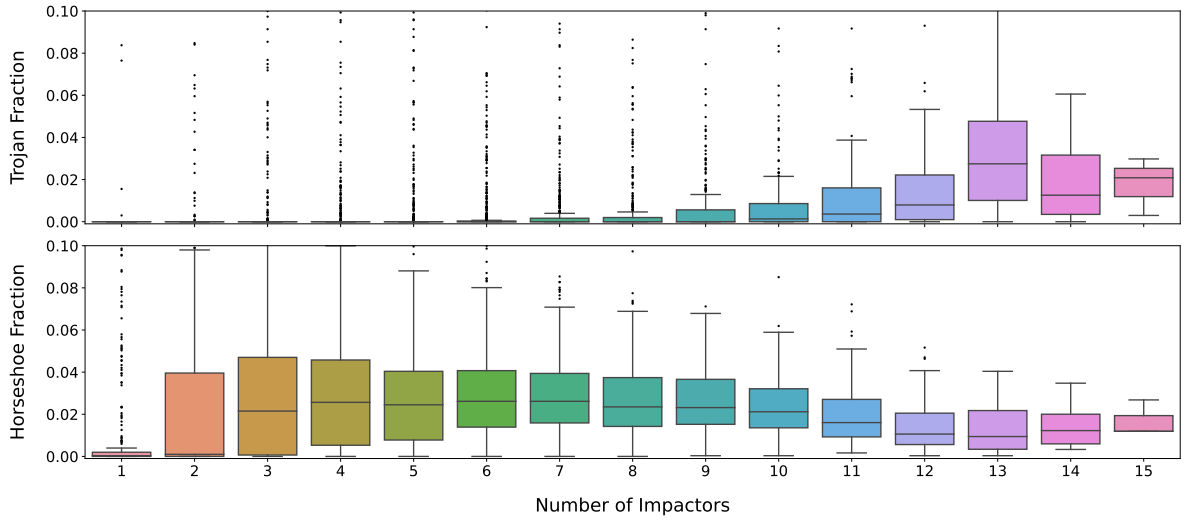


Figure 6.3: Box-and-whisker plots of the fraction of the initial population of 3,021 ETs that ended the simulation in a Trojan (top) or a horseshoe (bottom) orbit, plotted as a function of the number of impactors. There are data points that exceed the vertical range we have chosen for these plots, but such events are relatively rare.

Finally in Figure 6.5, we show the distribution of changes in Earth's semi-major axis and eccentricity during our simulations. We find that Earth's semi-major axis never changes by more than a few percent, and its eccentricity remains reasonably small — always within the range of predicted secular oscillations for the present-day solar system. The distribution of changes in semi-major axis is nearly symmetric about $\delta a = 0$. In contrast, the distribution of changes in eccentricity is slightly skewed toward negative values (decreases in eccentricity, i.e. increases in angular momentum). Note that we began our simulations with Earth's $a = 1$ au and $e = 0.0167$. The exact starting values are not important for this study, as long as they are in the vicinity of today's observed values. If the starting eccentricity is close to zero, however, then the changes

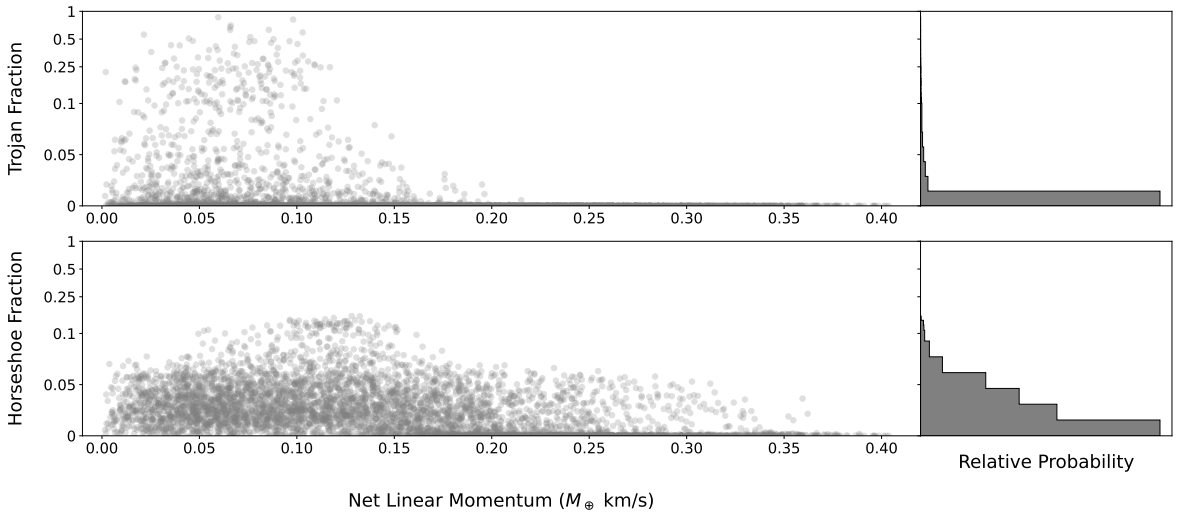


Figure 6.4: Scatter plot of the fraction of the initial population of 3,021 ETs that ended the simulation in a Trojan (top) or a horseshoe (bottom) orbit, plotted as a function of the net linear momentum of the impactors. The panels on the right-hand side are histograms representing the probability with which a given fraction of co-orbitals was retained. The fraction of objects remaining in co-orbitals decreases as the net linear momentum of the impactors increases. Note that the vertical axis is linear in the range $[0, 0.1]$ and logarithmic in the range $[0.1, 1]$.

in eccentricity (while still small) must be skewed toward the positive. Finally, we note that our choice of orbital elements for the incoming rocks includes specific angular momenta that are both larger and smaller than that of Earth. Different choices of these distributions could lead to somewhat different changes in the orbit of Earth (see, e.g., Fernandez & Ip 1984).

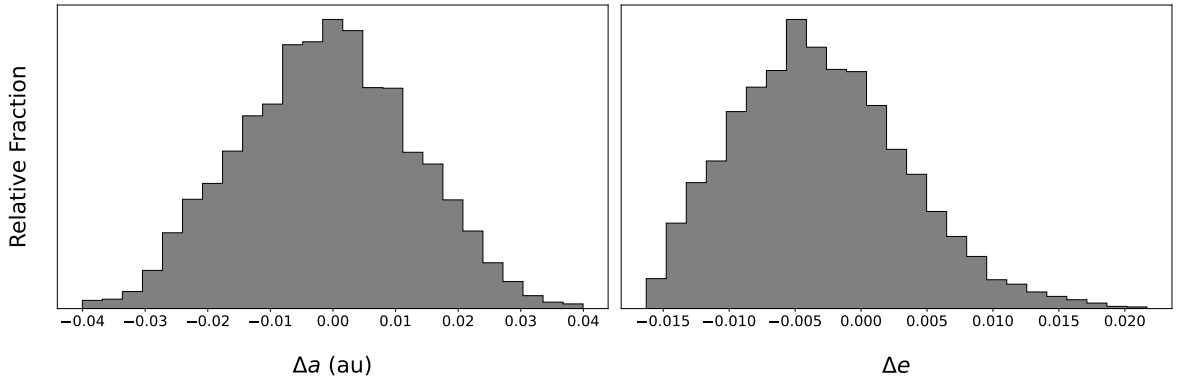


Figure 6.5: Histograms of changes in Earth's semi-major axis (left) and eccentricity (right) during each simulation. The initial values were $a = 1$ au and $e = 0.0167$.

6.5 Discussion and Conclusions

In this paper we have examined the stability of the ETs in the face of collisions between Earth and massive asteroids. The main result of this work is the demonstration that large impacts to Earth can readily remove Trojans from resonance. Using an ensemble of 5,000 numerical simulations, we have demonstrated that if Earth acquired the final 1% of its present-day mass from a handful of asteroidal impactors, the population of Earth Trojans would be reduced by more than 99% (less than 1% remaining; see Figure 6.3). Due to the wide distribution of impactor parameters and the chaotic nature of the system, there are some instances in which a significant fraction of the original Trojans can remain on Trojan orbits. However, such instances are rather rare. In addition, we find that the population of objects on horseshoe orbits exceeds that of the Trojans, provided that the total number of collisions remains small. For the scenarios explored in this work, on average approximately 2% of the primordial Trojan population ends up on horseshoe orbits.[6] As long as the number of impactors remains small, the number of Trojans expected to survive is at least an order of magnitude smaller than the horseshoe objects.[7]

We also analyzed our simulation data for correlations between the impactor parameters and the surviving fraction of Trojans. First we find that as the number of impactors increases to more than of order ~ 10, the surviving fraction of Trojans increases. This finding makes sense intuitively: If we consider the limiting case in which Earth acquired the final 1% of its mass by having 10^{28} grains of sand slowly and isotropically scattered onto its surface, the resulting impulse would be a negligible and the Trojans would remain undisturbed. In a similar vein, we find a trend between the maximum retention fraction of the Trojan population and the net linear momentum imparted to Earth by the collisions (see Figure 6.4). As the net linear momentum increases, the upper bound on the surviving fraction of Trojans decreases. We found that no Trojans survived in any simulation with net linear momentum $\Delta p \gtrsim 0.25\ M_\oplus$ km/s (see also Appendix 6.B). We also analyze the changes in Earth's semi-major axis and eccentricity at the end of the simulations

[6] As an important caveat, note that we did not integrate the objects for the age of the solar system, and therefore can not say whether they would survive until the present day.

[7] As the number of impactors increases, we expect to see more ETs and fewer horseshoe objects.

to make sure that its orbit did not become unreasonably excited. The changes to the orbital elements are relatively small (see Figure 6.5), with $\Delta a \sim 0.01$ au and $\Delta e \sim 0.005$.

This set of results has important implications for our understanding of the fate of primordial Earth Trojans. While numerical simulations show that stable Earth Trojans can exist, dedicated searches still have not discovered any such objects. It has been noted that the Yarkovsky effect can remove small Trojans from resonance on the timescale of the age of the solar system, but it can not explain the apparent absence of large objects. In contrast, the collision mechanism explored in this paper is independent of the size of the Trojans. Moreover, this collision mechanism is of particular interest because previous work suggests that Earth has indeed experienced large impact events after its formation: one example being the impact that formed the Moon (e.g., Canup 2012), and another a series of impacts in which Earth acquired its Late Veneer (e.g., Raymond et al. 2013). The Moon is thought to have formed as the result of a Mars-sized object colliding with the primordial Earth. Our simulations show that such a collision likely to have cleared out the primordial ETs (though perhaps not if the collision was relatively slow, thus imparting a rather small change in net linear momentum). However, a collision this large would probably have re-injected some material into Trojan resonances (note that similar ideas have been put forth in the context of Martian Trojans, e.g., Polishook et al. 2017). On the other hand, the collisions that gave Earth its Late Veneer are hypothesized to have occurred well after the Moon-forming collision. If the Late Veneer was comprised of a small number of large impactors, as favored by recent models (Raymond et al., 2013; Genda et al., 2017), any re-injected material is likely to have been cleared. Furthermore, the impacts comprising the Late Veneer were probably not large enough to re-populate L4 and L5 with ejecta (especially bodies large enough to survive the Yarkovsky effect), but this complication should be explored in future work. Additionally, because the Late Veneer is thought to have occurred late in the formation of the inner solar system, there was probably not a large population of ambient bodies to repopulate the Trojan resonances (Brasser et al., 2016). Despite the large parameter space in which dynamically stable ETs can exist, it is possible that all or most such objects have been removed due to asteroids impacting

Earth.

We note that this work has limitations and can be improved in the future. Given a considerable amount of additional computing, future studies should by more thoroughly explore the distributions of impactor parameters. In particular, changing the size, mass, and eccentricity distributions of the impactors could influence the results. For example, a shallower size distribution would mean that fewer impacts are necessary to deliver the required mass to Earth, while a steeper size distribution would mean that more impacts are necessary. In this vein, the total mass of the impactors is also important, and should be considered in future studies. We also made the approximation that the collisions were co-planar with Earth's orbit, while in reality they would have some distribution of inclinations. One should also address the possibility of impactors coming from outside of the asteroid belt. In this work we considered only impactors on orbits with $a \in [1, 5]$ au. Orbits with larger semi-major axes corresponding to the giant planet region or the Kuiper Belt would have marginally larger relative velocities when entering Earth's Hill sphere. Such collisions would deliver larger impulses to Earth than their low-a counterparts, and would thus be more detrimental to the survival of the Trojan populations. However, this correction is rather small — a body with $a \to \infty$ can achieve a maximum velocity of only ~ 52 km/s. We also did not model the effect of close encounters between Earth and potential impactors, or impacts directly onto the Trojan asteroids. We make analytic estimates of the relative sizes of these effects in Appendices 6.C and 6.D respectively. We find that the expected net velocity perturbation to Earth from close encounters is roughly equivalent to that expected from direct collisions. The perturbations from close encounters will be orthogonal to those from direct collisions, so including close encounters in the model should enhance the net perturbation to Earth by a factor of approximately $\sqrt{2}$. Direct collisions with Trojans, on the other hand, are a less likely source of orbital disruption than perturbations to Earth's orbit by an order of magnitude. These considerations imply that the present work likely represents a lower limit on the disturbance to the Trojan population. Of course, the most important additional work is observational: Future observational surveys will either detect a (presumably small) population of Earth Trojans or place

increasingly tight limits on their existence. With sufficiently tight constraints on the population of primordial ETs, future simulations can in turn constrain the properties of the primordial asteroid belt.

6.A Velocity Corrections

In this section, we assess the severity of the approximations used in Section 6.2. In particular, we have assumed that the orbit of the asteroid only depends on the solar gravitational potential, and we have assumed that the orbit of Earth is circular.

In this approximation, we assume that the asteroid follows an elliptical orbit around the Sun, and then encounters Earth at the point where the orbits intersect. In this phase, the distance $d_\odot$ between the asteroid and Earth in the frame of the solar system thus vanishes ($d_\odot \to 0$) at the intersection. The approximation then switches over to a reference frame where the planet (Earth) is at rest, and the asteroid is assumed to have speed v_∞ and distance $d_\odot \to \infty$. We are thus using the method of matched intermediate asymptotics, where the inner boundary condition for the outer problem ($d_\odot \to 0$) represents the outer boundary condition for the inner problem ($d_\odot \to \infty$). In order to test the consistency of this approach, we need to determine how much the gravitational potential of Earth affects the incoming speed v_∞ calculated previously.

The boundary between the inner problem and the outer problem is delineated by the Hill sphere, as given by Equation (6.2). At this Hill sphere boundary, the additional velocity that the asteroid would gain from the gravitational potential well of Earth is given approximately by

$$v_x^2 = \frac{2GM_\oplus}{R_H} = \frac{2GM_\oplus}{a_\oplus}\left(\frac{3M_\odot}{M_\oplus}\right)^{1/3} = 3^{1/3}\frac{2GM_\odot}{a_\oplus}\left(\frac{M_\oplus}{M_\odot}\right)^{2/3} \tag{6.15}$$

As shown previously, the incoming speed v_∞ is comparable to the orbital speed of Earth. If we let v_T denote the true speed (corrected for Earth's gravity), then

$$v_T^2 = v_\infty^2 + v_x^2, \tag{6.16}$$

which in turn implies

$$\frac{v_T}{v_\infty} = \left[1 + \frac{v_x^2}{v_\infty^2}\right]^{1/2} \approx 1 + \frac{v_x^2}{2v_\infty^2}. \tag{6.17}$$

The size of the correction is thus given by

$$\frac{\Delta v}{v} = \frac{v_T - v_\infty}{v_\infty} \approx 3^{1/3} \left(\frac{M_\oplus}{M_\odot} \right)^{2/3} \sim 0.0003. \tag{6.18}$$

If we use the sphere of influence from Equation (6.3), instead of the Hill radius, the above equation has the form

$$\frac{\Delta v}{v} = \frac{v_T - v_\infty}{v_\infty} \approx \left(\frac{M_\oplus}{M_\odot} \right)^{3/5} \sim 0.0005. \tag{6.19}$$

The correction due to the gravity of Earth is thus relatively small.

Another correction arises from our assumption that Earth has a circular orbit. In this case, the planet has a constant speed $v_\oplus^2 = GM_\odot/a_\oplus$. For small eccentricity, the variation in speed over the course of the orbit is given approximately by the expression

$$\left. \frac{\Delta v}{v} \right|_{max} = \left[\frac{1+e}{1-e} \right]^{1/2} - 1 \approx e \approx 0.0167, \tag{6.20}$$

where we have used the observed eccentricity of Earth. In summary, both corrections are small, less than about 1 percent. The correction for the eccentricity of Earth's orbit is larger than the correction due to the Earth's gravity.

Next we consider the crossing time: The relative asteroid speeds are of order $v_\infty = 10$ km/s and the size of the Hill sphere is of order $R_H = 10^{11}$ cm. The crossing time of the asteroid is thus of order

$$t_{cross} \approx \frac{R_H}{v_\infty} \sim 10^5 \text{ sec} \sim 1 \text{ day}. \tag{6.21}$$

Over the crossing time ($\sim$ 1 year), the planet rotates through about 1/300 of its orbit, so the reference frame of the planet does not rotate appreciably over the course of the encounter. We thus expect the effects of the rotating reference frame to be relatively small, but they can be quantified, as shown below. The encounter takes place in a rotating frame of reference, with rotation rate Ω given by the mean motion of Earth. In a rotating reference frame, the acceleration is given by

$$\mathbf{a}_R = \mathbf{a}_g - 2\Omega\hat{z} \times \dot{\mathbf{r}} - \Omega^2 \hat{z} \times (\hat{z} \times \mathbf{r}), \tag{6.22}$$

where we have taken the angular velocity vector of the rotating frame to point in the $\hat{z}$ direction. On the right hand side of the equation, the first term is the acceleration due to the gravity

of Earth, the second term is the Coriolis acceleration, and the third term is the centrifugal acceleration. In order of magnitude, we can find the relative size of the Coriolis term and the centrifugal term,

$$\frac{a_{cor}}{a_g} = 2\left(\frac{M_\odot r^3}{M_\oplus a_\oplus^3}\right)^{1/2} = \frac{2}{\sqrt{3}}\left(\frac{r}{R_H}\right)^{3/2} \qquad \text{and} \qquad \frac{a_{cen}}{a_g} = \left(\frac{M_\odot r^3}{M_\oplus a_\oplus^3}\right) = \frac{1}{3}\left(\frac{r}{R_H}\right)^3 . \quad (6.23)$$

At the Hill radius, both ratios are thus of order unity, but they both decrease rapidly as the asteroid penetrates the sphere and orbits closer to the planet. The centrifugal term is always smaller than the Coriolis term. We can thus estimate the size of this correction by using only the Coriolis term, and by noting that the change in velocity produced by this additional acceleration is given approximately by

$$\Delta v \sim \int a_{cor} dt \sim \int 2\Omega v dt \sim 2\Omega dr < 2\Omega R_H = 2\left(\frac{GM_\odot}{a_\oplus}\right)^{1/2}\left(\frac{M_\oplus}{3M_\odot}\right)^{1/3} \quad (6.24)$$

We thus obtain the approximate bound

$$\frac{\Delta v}{v} \lesssim 2\left(\frac{M_\oplus}{3M_\odot}\right)^{1/3} \approx 0.03 \quad (6.25)$$

If we use the sphere of influence instead of the Hill radius, the above expression becomes

$$\frac{\Delta v}{v} \lesssim 2\left(\frac{M_\oplus}{M_\odot}\right)^{2/5} \approx 0.01 \quad (6.26)$$

The correction for the rotating frame of reference is thus small, but it is somewhat larger than the corrections outlined earlier due to the influence of Earth in the solar reference frame and the non-circular nature of its orbit.

6.B Trojan Removal

Here we make some rough estimates of the energetic requirements for the removal of Earth Trojans from bound orbits. In this context, a bound orbit is one such that the body is bound to either the L4 or L5 point, and hence executes a Trojan orbit. Unbound orbits are no longer Trojans, but still remain bound to the Sun.

We work in the approximation of the circular restricted three body problem, following the notation of Murray & Dermott (2000) and work in their dimensionless units where the mean motion of Earth is unity and the separation of the two majors bodies is also unity. In these units, the Jacobi constant has the form

$$C_J = (x^2 + y^2) + 2\left(\frac{1-\mu}{r_1} - \frac{\mu}{r_2}\right) - v^2. \tag{6.27}$$

The values of the Jacobi constants for the five Lagrangian equilibrium points are as follows:

$$C_1 = 3 + 3^{4/3}\mu^{2/3} - 10\mu/3 \tag{6.28}$$

$$C_2 = 3 + 3^{4/3}\mu^{2/3} - 14\mu/3 \tag{6.29}$$

$$C_3 = 3 + \mu \tag{6.30}$$

$$C_4 = C_5 = 3 - \mu \tag{6.31}$$

We can thus define the difference in Jacobi constants between Lagrangian points:

$$\Delta_{3-45} = C_3 - C_4 = 2\mu \tag{6.32}$$

and

$$\Delta_{12-45} = \frac{1}{2}(C_1 + C_2) - C_4 = (9\mu)^{2/3} - 3\mu \tag{6.33}$$

As a rough approximation, the velocity increment required to move an object from a Trojan orbit to a horseshoe orbit is then given by

$$\Delta v \sim (\Delta_{3-45})^{1/2}\, v_\oplus = (2\mu)^{1/2} v_\oplus \approx 0.0024 v_\oplus \qquad \Longrightarrow \qquad \Delta v \sim 0.074 \text{ km/s} \tag{6.34}$$

Similarly, the velocity increment required to make a Trojan become unbound from any libration is given approximately by

$$\Delta v \sim (\Delta_{12-45})^{1/2}\, v_\oplus = (9\mu)^{1/3} v_\oplus \approx 0.03 v_\oplus \qquad \Longrightarrow \qquad \Delta v \sim 0.9 \text{ km/s} \tag{6.35}$$

The above values represent upper bounds on the velocity increments required to dislodge the Trojans. Note that when an impact takes place on Earth, the immediate collision changes the

velocity of Earth, so that the velocity of the Trojan in the Earth-centric frame changes accordingly. However, there is another effect: the collision also induces an eccentricity in Earth's orbit, the magnitude of which will depend on the direction from which Earth is struck. As a result of the nonzero eccentricity, the speed of Earth relative to the Trojan continues to change over the course of an orbit, causing additional instability.

6.C Close Encounters

The calculation of the main text shows that direct collisions between incoming rocky bodies and Earth can lead to the disruption of Trojan orbits. In addition to direct impacts, close passages of rocky bodies will also lead to orbital perturbations of Earth. Since these close passages do not add mass to Earth, these encounters will enhance the disruptive effects already considered (i.e., the disruption found in the main text should be considered as a lower limit).

We can estimate the magnitude of this effect. Consider a field of rocky bodies that are streaming toward Earth. As a first approximation, we consider the rocks to have a given mass m, speed v with respect to the Earth, and assume that all of them are coming from the same direction. One could generalize to distributions of (m, v), but we start by comparing the velocity perturbations produced by direct collisions with those induced by close encounters.

The rate at which rocky bodies collide with Earth is given by

$$\Gamma = nv\pi f_{\rm g} R_\oplus^2 \,, \tag{6.36}$$

where n is the number density of rocks and where $f_{\rm g} > 1$ is gravitational focusing factor. It will be useful to work in terms of the number of impacts N, so we write

$$(\Delta t) = \frac{N}{nv\pi f_{\rm g} R_\oplus^2} \,. \tag{6.37}$$

If all of the collisions have the same sign, the velocity perturbation (to Earth) produced over time (Δt) has the form

$$(\Delta v)_{\rm dir} \approx nv^2 \left(\frac{m}{M_\oplus}\right) \pi f_{\rm g} R_\oplus^2 (\Delta t) = Nv \left(\frac{m}{M_\oplus}\right) \,. \tag{6.38}$$

However, are interested in the square of the velocity perturbation. If the incoming rocks produce both positive and negative angular momentum perturbations, as expected, then the velocity perturbation grows as a random walk (Binney & Tremaine, 2008) so that

$$(\Delta v)^2_{\rm dir} \approx N v^2 \left(\frac{m}{M_\oplus}\right)^2 . \tag{6.39}$$

For completeness, we note that if the rocks all arrived from a single direction, then the expression would pick up an additional factor of N.

Now consider the case of near-misses or close encounters. First we note that while the direct collisions produce velocity perturbations in the direction of motion of the rocky impactor, close encounters produce velocity perturbations in the perpendicular direction. Working in the impulse approximation, the velocity perturbation produced by a fly-by encounter (Binney & Tremaine, 2008) has the form

$$\delta v = \frac{2Gm}{bv} , \tag{6.40}$$

where b is the impact parameter of the interaction (we suppress subscripts, but keep in mind that the direction is perpendicular). For close encounters, the interactions are equally likely to occur on either side of the planet, which means that the velocity perturbations have both positive and negative signs. Moreover, the size of the perturbation depends on the impact parameter. The rate of encounters with impact parameter between b and $b + db$ is given by

$$\Gamma_b = nv2\pi b db . \tag{6.41}$$

As a result, the square of the velocity perturbation accumulates according to the relation

$$d\left[\frac{d}{dt}(\Delta v)^2_{\rm cl}\right] = \left(\frac{2Gm}{bv}\right)^2 nv2\pi b db . \tag{6.42}$$

After integrating over all impact parameters, and over time, we obtain

$$(\Delta v)^2_{\rm cl} = \frac{8\pi G^2 m^2 n(\Delta t)}{v} \ln\Lambda , \tag{6.43}$$

where

$$\Lambda \equiv \frac{b_{\rm max}}{b_{\rm min}} . \tag{6.44}$$

The minimum impact parameter must be larger than Earth so that $b_{\rm min} \geq R_\oplus$. In general, we expect $b_{\rm min} \approx \sqrt{f_{\rm g}} R_\oplus$. As a working approximation, we can take the maximum impact parameter to be the Hill radius, i.e.,

$$b_{\rm max} \approx R_H = a_\oplus \left(\frac{M_\oplus}{3 M_\odot} \right)^{1/3} \approx 0.01 \, {\rm AU} \, . \tag{6.45}$$

As result, the so-called Coulomb logarithm $\ln \Lambda \approx 5$. Removing the factor of (Δt) in favor of N we find

$$(\Delta v)^2_{\rm cl} = \frac{8 G^2 m^2 N}{v^2 R_\oplus^2 f_{\rm g}} \ln \Lambda \, . \tag{6.46}$$

Even though the perturbations are in the parallel and perpendicular directions, for direct and close encounter interactions respectively, we are nonetheless interested in their relative magnitude. Since the direct collisions are also subject to a random walk, the ratio of the two contributions is given by

$$\mathcal{R} = \frac{(\Delta v)^2_{\rm cl}}{(\Delta v)^2_{\rm dir}} = \frac{8 \ln \Lambda}{f_{\rm g}} \left(\frac{G M_\oplus}{R_\oplus v^2} \right)^2 \approx \frac{40}{f_{\rm g}} \left(\frac{v_{\rm esc}}{v} \right)^4 \, . \tag{6.47}$$

The distribution of incoming speeds (Figure 1) shows a broad peak near $v \sim 25$ km/s, whereas the escape speed of Earth $v_{\rm esc} \sim 11$ km/s. The gravitational focusing factor $f_{\rm g} = 1 + (v_{\rm esc}/v)^2$, so that the ratio $\mathcal{R} \approx 1$. In other words, in the absence of asymmetries in the collision stream, the velocity perturbations due to close encounters are roughly comparable to those induced by direct collisions. The effects considered in this paper thus represent a lower limit to the full degree of Trojan disruption. Moreover, since the close encounter perturbations are perpendicular to the perturbations produced by direct collisions, the contributions add in quadrature, so we expect the enhancement factor to be $\sim \sqrt{2}$.

6.D Collisions with Trojans

This Appendix considers another channel for the disruption of Trojan orbits. Given the reservoir of rocky bodies on Earth-crossing orbits that interact with Earth and provide the Late Vaneer, direct collisions with the Trojans can occur. For a given Trojan, the expected number of collisions

is given by

$$N_T = n_T \sigma_T v(\Delta t)\,, \tag{6.48}$$

where $\sigma_T = \pi R_T^2$ is the cross section for the Trojan and n_T is the number density of rocky bodies large enough to disrupt the orbit. Eliminating (Δt) using equation (6.37), the above expression can be written in the form

$$N_T = N_\oplus \frac{n_T \sigma_T}{n_\oplus \sigma_\oplus}\,, \tag{6.49}$$

where $n_\oplus$ is the number density of object considered in this paper to provide perturbations of Earth's orbit. Note that as the size R_T of the Trojan decreases, the cross section $\sigma_T \sim R_T^2$ decreases, but the number density of objects n_T large enough to disrupt the body increases. Although the size distribution of Earth-crossing bodies in not well-known, the two effects roughly compensate. We can thus obtain an approximate result using $n_T = n_\oplus$ and taking the cross section σ_T to be that of an $R = 1000$ km body. Using these values along with the size of Earth and the gravitational focusing factor, we find

$$N_T \approx \frac{N_\oplus}{50}\,. \tag{6.50}$$

Here, $N_\oplus$ is the number of (large) impactors from the simulations of the paper, where we find $N_\oplus = 1 - 15$, with typical values $\sim 5 - 10$. This estimate suggests that direct collisions with Trojans are a less likely source of orbital disruption than perturbations to Earth's orbit, by a factor of 5 to 10.

Part II

Discovery

Chapter 7

The Absolute Magnitude Distribution of the Cold Classical Kuiper Belt

He could pick the fly specks out of pepper.

— My Mom

This chapter has been submitted for publication in the *Planetary Science Journal*.

7.1 Introduction

Beyond the orbits of the major planets, our solar system hosts a large population of minor bodies known as Kuiper Belt Objects (KBOs). In the 30 years since the observational establishment of the Kuiper Belt (Jewitt & Luu, 1993), several surveys (e.g., Millis et al. 2002; Bernstein et al. 2004; Petit et al. 2011; Bannister et al. 2018; Bernardinelli et al. 2022) have pushed the inventory of known objects to nearly 4000. These bodies, which are left over from the birth of our planetary system, provide constraints on its formation and dynamical evolution. When taken in aggregate, their dynamics, compositions, and sizes enable us to infer details about the dynamical evolution of the planets, the composition of our solar system's protoplanetary disk, and even the physical processes by which planetesimal formation occurred (see, e.g., Nesvorný 2018, Gladman & Volk 2021).

In particular, the size distribution of the so-called Cold Classicals (CCs)—which are thought to be relics of the birth of the solar system, relatively untouched and uncontaminated in the ~ 4.5 Gyr since their formation—is a sensitive probe of the process of planetesimal formation. If the CCs are truly a quiescent population, a measurement of their size distribution can provide us with a unique opportunity to directly compare a primordial size distribution with predictions made by planetesimal formation models. Such a comparison will enable us to hone our formation models, and better understand the details of the physical processes at play in planetesimal formation, as well as the specific conditions of our own protoplanetary disk.

Recently, the streaming instability (SI; Abod et al. 2019) has begun to emerge as a leading theory of planetesimal formation. Numerical SI simulations predict an exponentially tapered power law absolute magnitude (H) distribution, enabling a direct comparison between theory and observation. Kavelaars et al. (2021) found that the absolute magnitude (H) distribution of the CCs detected by the Outer Solar System Origins Survey (OSSOS; Bannister et al. 2018) is consistent with an exponentially tapered power law. However, the CCs used by Kavelaars et al. (2021) only went as faint as $H_r \sim 8.3$, leaving faint-end consistency with the SI as an open question. Existing literature on measurements of the faint-end of the CC absolute magnitude (H) distribution is in weak tension with SI models of planetesimal formation. In particular, Fraser et al. (2014) find a marginally steeper faint-end size distribution than is predicted by SI simulations. However, a dearth of observed objects at the faint end of the distribution limits the usefulness of such comparisons. We require a larger, deeper set of CC detections by a survey with well-understood biases to thoroughly test any planetesimal formation theory.

The DECam Ecliptic Exploration Project (DEEP) is the first survey with sufficient depth and areal coverage to settle the tension in the shape of the faint end of the H distribution of the CCs. In this paper, we use our detections from 20 DECam fields (comprising an area of approximately 60 sq. deg.) with a mean limiting magnitude of $m_r \sim 26.2$ to reconstruct the luminosity function of the Kuiper Belt as a whole, as well as the luminosity function of the CC population. As the main scientific result of this paper, we use our results to reconstruct the

underlying absolute magnitude (H) distribution of the CCs, and find consistency with models of planetesimal formation via the streaming instability (Abod et al., 2019; Kavelaars et al., 2021), as well as evidence that the H distribution of the CCs continues to flatten out beyond our detection limit.

This paper is organized as follows. In Section 7.2 we outline our observational strategy and discuss the data used in the subsequent analysis. Next we describe our image pre-processing pipeline (Section 7.3) and the pipeline used to carry out the object search (Section 7.4). We present an overview of our detections in Section 7.5. In Section 7.6 we calculate our detection efficiency using implanted synthetic objects. We compute the luminosity function for our KBOs as a whole in Section 7.7. In Section 7.8 we isolate a sample of CCs from our detections. We use this isolated sample to compute the luminosity function for the CCs separately in Section 7.9. Our main scientific result, presented in Section 7.10, is a calculation of the absolute magnitude distribution for the CC population. In Section 7.11 we test the consistency of our absolute magnitude distributions with the results from Bernstein et al. (2004). The paper concludes in Section 7.12 with a summary of our results and a discussion of their implications.

7.2 DEEP Survey Strategy and Data

DEEP is a survey being carried out with the Dark Energy Camera (DECam) on the 4-meter Blanco telescope located at Cerro Tololo Inter-American Observatory in Chile. In 2019 our team targeted four sets of three 3 sq. deg. discovery fields located along the invariable plane in search of KBOs. In 2020-21 we expanded our search to fields that account for the diffusion of the 2019 objects (see Papers I and II for more details). This paper focuses on the data taken in our so-called B1 fields from 2019–2021. The pointings for these fields are shown with the DECam footprint in Figure 7.1, and are given in tabular form in Appendix 7.A.

A DEEP exposure sequence, designed with a cadence ideal for a technique called shift-and-stack (Tyson et al., 1992; Gladman & Kavelaars, 1997; Allen et al., 2001; Bernstein et al., 2004; Holman et al., 2004; Parker & Kavelaars, 2010; Heinze et al., 2015; Whidden et al., 2019),

typically consists of ∼100 consecutive 120-second VR-band exposures of the target field. In the shift-and-stack approach, single-epoch images are shifted at the rate of a moving object (rather than at the sidereal rate) so that a moving object appears as a point source in the co-added stack.

There are two primary reasons why the shift-and-stack technique is preferable to long exposures for the discovery of moving objects. First, since the rate and direction of an object's motion are not known *a priori*, we are able to stack our images in a grid of velocity and direction that spans the space of possible KBO motions. The second benefit is the preservation of the S/N of moving objects. For stationary sources in astronomical images, S/N goes like $t^{1/2}$. A source is effectively stationary if its apparent position changes by less than the pixel resolution of the observing instrument over the course of the exposure. This sets an upper limit on the useful exposure time when searching for moving objects, which we will call t_{max}. Given DECam's typical VR-band seeing of ∼0.9" and the typical KBO rate of motion of a few arcseconds per hour, our value of t_{max} is on the order of a few minutes. When $t \geq t_{\mathrm{max}}$, the S/N of *stationary sources* continues to go like $t^{1/2}$, while smearing causes the S/N of *moving sources* to go like $t^{-1/2}$. Thus moving objects fade into the background while the S/N of background sources continues to grow. Because DECam's CCDs have negligible read noise, we lose no sensitivity by adding together many short images, and thus the S/N of the moving objects continues to increase like $t^{1/2}$.

7.3 Image Pre-processing

In this section we describe how our images are processed in preparation for our shift-and-stack pipeline (described in Section 7.4). The following steps take place after the images have gone through preliminary reductions with the DECam community pipeline (Valdes et al., 2014).

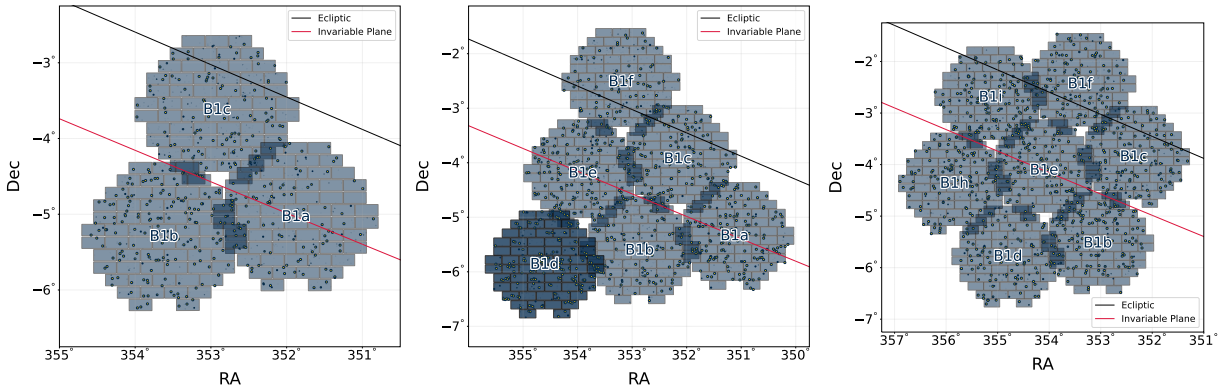


Figure 7.1: The DEEP "B1" TNO search fields used in this analysis. Each hexagonal shaded area represents the DECam focal plane with its 61 active CCDs. The B1a-c fields were observed with integrated exposure times of $\sim$ 3.5 hours in August 2019 (left), and re-observed at suitably shifted positions (not plotted) in September 2019. In October 2020 we observed the B1a-f fields (center). In October 2021 we observed the B1b-f, h, and i fields. The larger areas in 2020 and 2021 account for diffusion of the 2019 detections. The yellow dots represent our real KBO detections. Note that Neptune is near the B1 fields, meaning that there should be relatively few resonant objects among our detections.

7.3.1 Synthetic TNOs

To enable studies of our efficiencies, we generated a population of several thousand synthetic sources to plant in our images. These synthetic sources were not meant to emulate a realistic population, but rather to test efficiency across the space of all possible bound orbits in the Kuiper Belt. They span distances from $\sim$ 20 au to a few hundred au, and include fully retrograde orbits. To enable studies of efficiency as a function of brightness, we have given our fakes apparent magnitudes as bright as 20, and as faint as 27.2, as well as sinusoidal rotation curves with amplitudes as large as 0.5 mag, and rotation periods between a few hours and a few days.

7.3.2 Flux Calibration and Synthetic Source Injection

To calibrate the flux of our synthetic sources we calculate the photometric zero-point for each individual CCD image by cross-matching the non-streaked sources (ellipticity < 0.8[1]) against Pan-

[1]Most of the images can successfully match with enough sources with smaller ellipticity. However, a small number of them may fail due to some issues with the images, e.g. bad telescope guiding. This value ensures that every image can be processed using the same pipeline parameter settings. Since galaxies may also be included

STARRS sources (Magnier et al., 2013) with r_{SDSS} magnitude[2] between 15 and 21. We then use the Python package `SpaceRocks` (Napier, 2020) to calculate the sky position, sky motion, and brightness of each synthetic TNO, including a synthetic rotation lightcurve. With the sky motion of each object and the PSF of the image, we generate a streak model for each synthetic TNO. With the brightness, streak model, and photometric zero-point specified, we inject the synthetic TNOs into the image. Along with the synthetic TNOs, we also inject 12 stationary synthetic point sources with an r-band magnitude of 21 into each CCD image; since there are no static sources remaining after image differencing, we use these sources to re-calibrate each image.

7.3.3 Difference Imaging

After we implant synthetic sources, we prepare the images for the shift-and-stack pipeline. To do this, we must remove every stationary source—even the faintest sources that are not visible in single exposures. We apply the High Order Transform of Psf ANd Template Subtraction code HOTPANTS (Becker, 2015), which implements and improves upon the method of Alard & Lupton (1998) to create difference images. This code formed the basis for the Dark Energy Survey's supernova search pipeline (Kessler et al., 2015), and has consequently been thoroughly exercised on DECam data.

We first assemble the collection of exposures in the same observing run, including images from both the long and short stares (see Papers I and II for details). We generate three different templates by median-combining the single epoch images with seeing (by measuring the FWHM of the in-frame stars) in the top, middle, and bottom 1/3 of the ensemble. We require the minimum time separation between observations to be longer than 0.01 days (14.4 minutes) to ensure that the templates contain minimal flux from the slow movers. The HOTPANTS algorithm then performs seeing matches between science images and the template with the closest match to the image's seeing to generate difference images. The better-seeing images (either single epoch

during this process, we clip photometric outliers to purify the sources for calibration.

[2]The r_{SDSS} magnitude were converted from the Pan-STARRS r_{P1} and g_{P1} magnitude using Tonry et al. (2012).

or template) are convolved to match the images with the worst seeing, and the bright sources (pixel counts > 3000) are masked before performing image subtraction. The final step is masking the bright Gaia sources ($G > 18$) and the pixels above/under the $\pm 2\sigma$ level. This step usually masks less than 1% of the total pixels and not only cleans out most of the artifacts generated by the difference imaging process, but also removes streaks from artificial satellites, thus greatly reducing the false detection rate in shift-and-stack images. However, this masking comes with the caveat of masking bright sources. In practice, we find that sources brighter than VR$\sim$ 24–24.5 are masked. To ensure that we recover the bright objects, we also write out difference images in which we do not mask the pixels above/under the $\pm 2\sigma$ level. While the un-masked images produce significantly more spurious sources after shifting and stacking, we simply ignore the faint sources produced by these images, opting to only consider sources with S/N$\geq$ 30, thus finding all of the bright sources with minimal additional effort.

Finally we use the stationary magnitude 21 fakes to re-scale each difference image such that it has a zero-point of 30. After re-scaling, we compute weight images as the inverse of the variance in the difference images. Using weight images enables us to optimize the S/N of our stacks without manually rejecting images.

7.4 The Detection Pipeline

7.4.1 The Grid

To date, all shift-and-stack searches described in the literature have used regular grids (either rectangular, polar, or hexagonal) to sample a parameter space encompassing the sky motions of interest for moving objects. While this method of rate sampling works in principle, it is not ideal for two reasons. First, a regular grid spacing is susceptible to providing biased coverage for irregularly-shaped parameter spaces. Second, most previously used grids have covered a region of parameter space that includes unphysical orbits. Including such regions of parameter space increases both the computational cost and the rate of false-positive detections. To avoid these

issues, we developed a novel method of generating the shift-and-stack grid, as described below.

We begin by computing the grid bounds. A unique grid is required for each field, as its exact shape depends on the epoch and sky position of a pointing. We must also select the range of topocentric distances (Δ) of interest. For the search described in this paper, we consider $\Delta \in [35, 1000]$ au.[3] With RA, Dec, and Δ fixed, an object's position vector in the topocentric frame, $\vec{x}_T$ is uniquely determined. We then change the origin from the topocentric to the barycentric frame so that we have $\vec{x}_B$.[4] After changing the origin to the barycentric frame, we assign a velocity v_{bound} such that an object at the position $\vec{x}_B$ would be barely bound to the solar system (specifically, we use the speed appropriate for a semi-major axis $a = 200{,}000$ au). We then uniformly sample a collection of velocity vectors $\vec{v}_B$ on the surface of the sphere of radius v_{bound}. With the velocity vectors specified, the state vectors are fully determined, allowing for the computation of the corresponding RA rates and Dec rates as observed in the topocentric frame.[5] We repeat this process at several discrete values of Δ, and then use the Python package `alphashape` to draw a concave hull bounding the computed rates. This hull encompasses the full region of physically possible sky motions for the distances of interest.

Once we have computed the boundary of the concave hull for a given pointing, we choose a finite set of rates at which to stack our data. Toward this end, we employ a new method in which we fill the hull with a large sample of random points, and then use a K-means clustering algorithm to divide the region into N clusters. One can use dimensional analysis to determine that $N \propto A_{\text{hull}}/\epsilon^2$, where A_{hull} is the area of the hull, and ϵ is the desired grid spacing to minimize the maximum source trailing (roughly determined by the PSF width divided by the duration of the exposure sequence). We take the centroid of each cluster as a grid point. When compared to a rectangular grid, this method allows us to use $\sim$20% fewer grid points, and simultaneously reduce both the mean and maximum distance of any given point in the hull to the nearest grid

[3] Note that since we are searching through retrograde orbits, we are sensitive to objects at closer distances.

[4] For bodies in the inner solar system one should transform to the heliocentric frame, though in practice the distinction makes little difference.

[5] At this point, one can make cuts on the objects to disallow certain kinds of orbits as a function of Keplerian elements. For example, making a cut to consider only prograde orbits will drastically reduce the size of the grid.

point. As a result, we achieve more even coverage of physically possible rates, while minimizing the computational cost and opportunities for false positives. Although the process is random by nature, a sufficient number of random samples makes it nearly deterministic. We show an example of a grid for a 4-hour exposure sequence with 1" seeing in Figure 7.2.

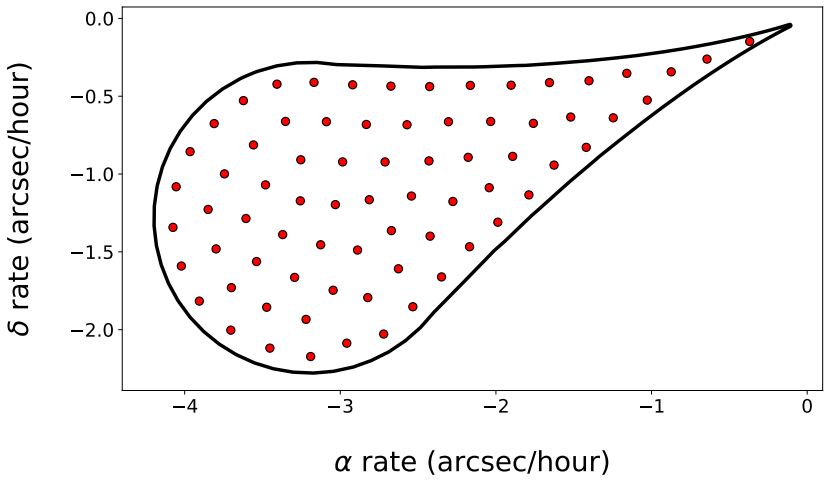


Figure 7.2: Sample grid of shift-and-stack rates corresponding to bound KBOs for a 4-hour exposure sequence with 1" seeing. The black teardrop-shaped boundary encompasses the space of possible bound orbits with topocentric distance between 35 and 1000 au. The red points are the sample rates computed using the procedure described above.

7.4.2 The Shift-and-Stack Procedure

After computing the shift-and-stack rates, we proceed with the stacking. We designate the first exposure in an exposure sequence as a reference image. We then compute the RA and Dec at the center of the reference image, and use the RA rate and Dec rate to compute the amount by which we have to shift each image to match with the reference image's center. In other words, we stack the pixels along the path taken by the center of the reference image for given RA and Dec rates. We do not consider variations in focal plane geometry across the chip, as the solid angle of a single chip is rather small (we thus assume that the chips are locally flat), and all of the images are `SWarped` (Bertin et al., 2002) for the difference imaging process. We perform

separate stacks for both the signal and weight images, and then obtain the full stack by dividing the stacked signal image by the stacked weight image.

After each stack we use the Python package `sep` (Barbary, 2016; Bertin & Arnouts, 1996) to extract all sources having at least 3 pixels with values above 1.5σ.[6] Each stack is contaminated with on the order of a few $\times 10^3$ spurious sources mostly consisting of cosmic rays, dead pixels, over-saturated pixels close to bright stars, and residuals from poorly subtracted stars and galaxies. Since we use on the order of 100 stacks per chip, the total number of spurious sources per chip is close to 10^5. Because the vast majority of spurious sources are not PSF-like, we have trained convolutional neural nets (CNNs) using `tensorflow` (Abadi et al., 2015) to reject them automatically. We trained one CNN on synthetic sources superimposed on background from DEEP difference images, and another on the `autoscan` training set that was used to train a random forest algorithm for background rejection in the Dark Energy Survey (DES) supernova search (Goldstein et al., 2015). Both CNNs retain nearly all of the signal and fail to reject different types of background, thus enabling a significant performance gain by requiring a source to be classified as *good* by both CNNs. This procedure cuts the number of sources per stack by three orders of magnitude, down to a few $\times 10^2$. After all of the stacks are completed for a given chip, we consider the complication that most objects are bright enough to be detected in adjacent stack rates. To eliminate this redundancy, we employ a `DBSCAN` clustering algorithm in pixel space to group detections associated with the same object.

The grid spacing in the initial shift-and-stack is good enough for source detection, but is too coarse to provide the best values for the position and rate for a given source. To refine the parameters, we use a Markov Chain Monte Carlo (MCMC) approach in which we perform targeted stacks on our candidates to maximize S/N. These targeted stacks are still restricted to the parameter space of bound orbits, but are now continuous in RA and Dec rates. This procedure enables us to obtain refined RA and Dec rates with uncertainties, while simultaneously optimizing the measured RA and Dec of the source. After refining the parameters of our detections, we

[6] For completeness, we note that we used the `matched` filter for our detections. However, the details are of little consequence, as even a simple peak finding algorithm works well for this purpose.

discard all sources with rates slower than 3 pixels per hour (distance $\gtrsim$ 150 au), as such slow rates tend to accumulate false positives and poorly subtracted stars. We feed our remaining candidates through a final CNN that reduces the number of sources per chip to $\sim$ 10. The images we show to the CNN are similar to the right panel in Figure 7.3. Good sources tend to show a characteristic radial pattern, while false sources do not.

Once we have refined our sources' rates and positions, we compute their flux and flux uncertainty using sep. We use these values to calibrate the magnitudes of our detections against the known magnitudes of our implanted sources, as well as obtaining a magnitude uncertainty.

Finally we do a reverse stack on our data, in which we repeat the above procedure with negated RA and Dec rates. Because no physical KBO would appear as a point source when stacked at these rates, all sources that result from this stack are false positives. This reverse stack enables the critical step of accounting for false positives in our detections.

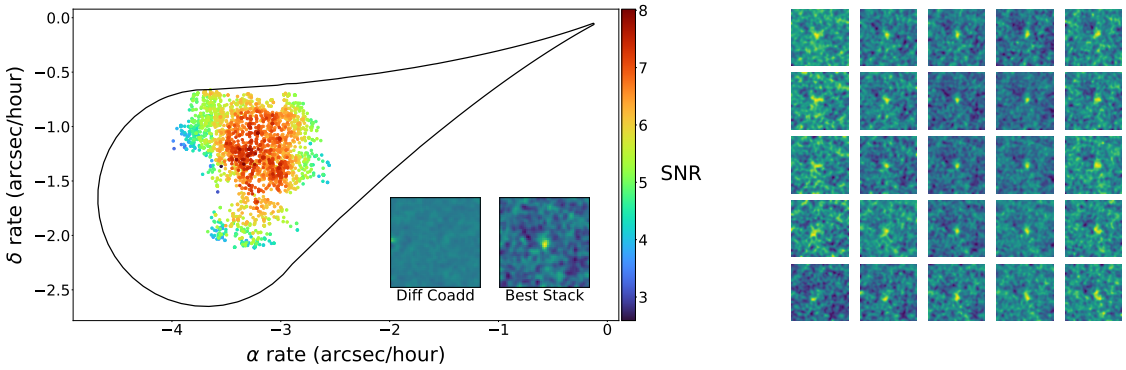


Figure 7.3: Vetting image for a synthetic source with an r-band magnitude of 27.0. This image format is used for visual inspection of our candidates. The left panel is what we call an MCMC plot. The black teardrop-shaped region is the boundary of the space of possible KBO rates of motion. The points are sampled from our MCMC, with colors corresponding the S/N of the object in the stack. Each point represents a targeted stack at the given rate. As a stack approaches the correct rate, causing the source to appear optimally point-like, the S/N increases in this characteristic manner. The inset labeled *Diff Coadd* is the stationary stack; the image shows no discernible signal. The inset labeled *Best Stack* is the stack at the best rate, as determined by the MCMC. In this image the KBO is quite apparent. The right panel shows a grid of stacks centered at the best rate, and offset in increments of 1 pixel per hour in RA rate and Dec rate.

7.4.3 Candidate Vetting

The final step in the discovery pipeline is the visual inspection of the grids that were labeled as *good* by the final CNN. This step includes the implanted synthetic sources and the sources from the reverse shift-and-stack. To do this visual inspection we developed a website where users vote `yes`, `no`, or `maybe` on a candidate and have their vote recorded to a database. The images are presented to the vetters in a blind manner, meaning that the vetter has no indication of whether an object is an implanted source, a source from the reverse shift-and-stack, or a true candidate. By blindly vetting all sources, we can reliably compute a voter's true and false positive rates for `yes`, `no`, and `maybe` votes. We require votes from three unique vetters for each object, and then combine the votes into a probability that a source is "real" using the following framework.

The odds (i.e., betting odds) of a source being good given a vote are

$$O(+|\text{vote}) = O(+)\frac{P(\text{vote}|+)}{P(\text{vote}|-)} \tag{7.1}$$

where O represents odds and P represents probability. The $+$ symbol means that the object is truly a good source, and the $-$ symbol means that the object is truly a bad source. We calculate the prior $O(+)$ using the excess in the number of sources in the forward shift-and-stack over the number of sources in the reverse shift-and-stack. The fraction on the right-hand-side of Equation (7.1) is the Bayes factor,

$$B \equiv \frac{P(\text{vote}|+)}{P(\text{vote}|-)}. \tag{7.2}$$

The quantity $P(\text{vote}|+)$ is calculated as the probability of assigning a given vote to an implanted source. Similarly, the quantity $P(\text{vote}|-)$ is calculated as the probability assigning a given vote to a source from the reverse shift-and-stack. Given multiple votes, we can simply update the information by taking the product of Bayes factors as

$$O(+|\text{votes}) = O(+)\prod_i B_i \tag{7.3}$$

where B_i is the Bayes factor of the ith voter. We can then convert the odds from Equation (7.3)

into the probability that a source is real as

$$P(+|\text{votes}) = \frac{O(+|\text{votes})}{1 + O(+|\text{votes})}. \tag{7.4}$$

We assign the values calculated by Equation (7.4) as a weight (w) for each of our detections. This treatment allows us to take a probabilistic approach in studying our detections in Sections 7.6-7.10.

7.5 Detections

In this section we qualitatively analyze our detections; we do more thorough quantitative analyses in Sections 7.6–7.10. In our 20 nights of data we detected a weighted sum of 2297.9 objects with weight greater than 0.01, corresponding to 2896 unique sources. We have elected to omit 3698 sources with weight less than 0.01, as such detections are rather unlikely to be real, and their omission does not change the results of our analysis. While the majority of our remaining sources have weight close to 1, there are some more ambiguous cases. We show a mosaic of all detections with weight ≥ 0.4 in Figure 7.4.

Based on our distribution of observed magnitudes (shown in Figure 7.5), it appears that our detection efficiency begins to fall off at magnitudes fainter than $r \sim 26$ (though we explore this in detail in Section 7.6).

It is also informative to examine the sky moving rates of our detections, which we display in Figure 7.6. In this figure, the size of each marker is proportional to its weight. The most apparent feature here is a large population of objects moving at approximately 3" per hour, mostly corresponding to CCs (see Section 7.8). Based on the increased density of small points at slow rates of motion, we note a propensity for slow-moving false positives to pass our CNNs, but to be later given low weights after human inspection. At faster rates, points with low weights appear to be evenly spread. The presence of such features provides further evidence that a great deal of care in avoiding false positives is required when making statistical use of single-night

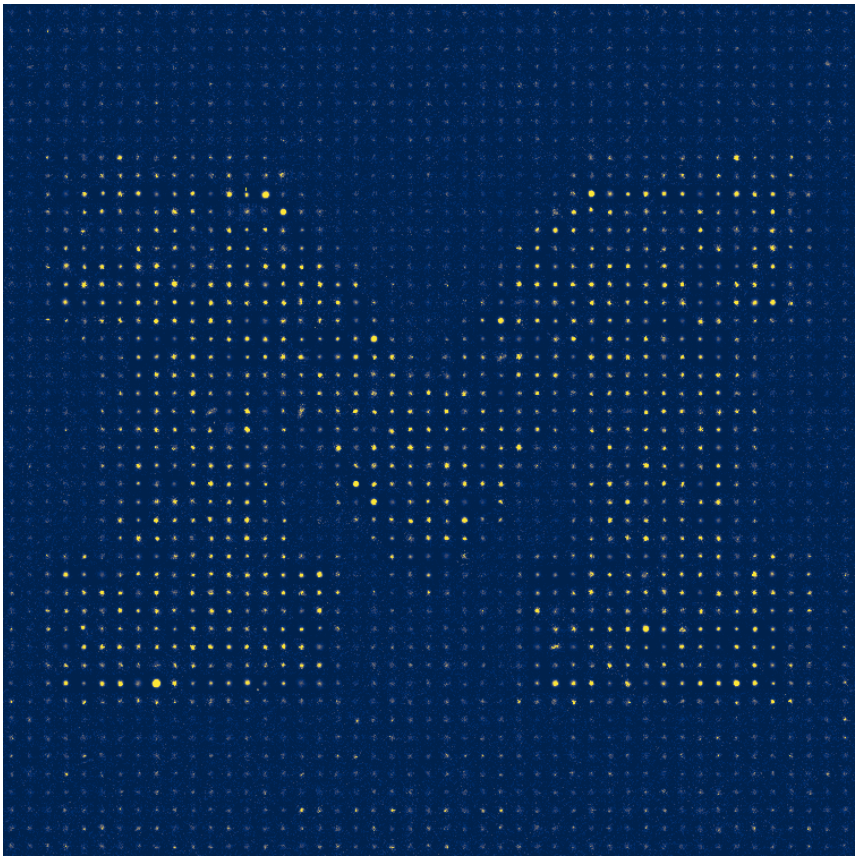


Figure 7.4: Mosaic of DEEP B1 detections with weight ≥ 0.4.

detections. When such detections can be linked to several epochs, single-night false positives will be less problematic.

7.6 Detection Efficiency

In order to make use of our detections, we must understand the efficiency with which we recover our implanted synthetic sources. We parameterize the detection efficiency as a function of apparent magnitude using a single hyperbolic tangent function, given the following equation

$$\eta(m) = \frac{\eta_0}{2}\left(1 - \tanh\left(\frac{m - m_{50}}{\sigma}\right)\right), \tag{7.5}$$

where η_0 is the peak detection efficiency, m_{50} is the magnitude at which the detection efficiency drops to $\eta_0/2$, and σ is the width of the hyperbolic tangent function.[7] We weight each of the

[7] More complicated functional forms do not improve the fit, and do not change the results of our analysis.

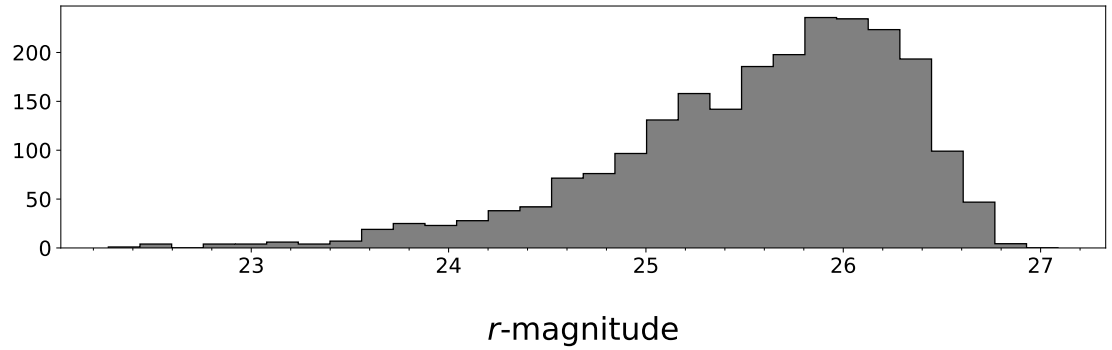


Figure 7.5: Weighted distribution of the apparent magnitudes of our detections. The weights are given by Equation [7.4].

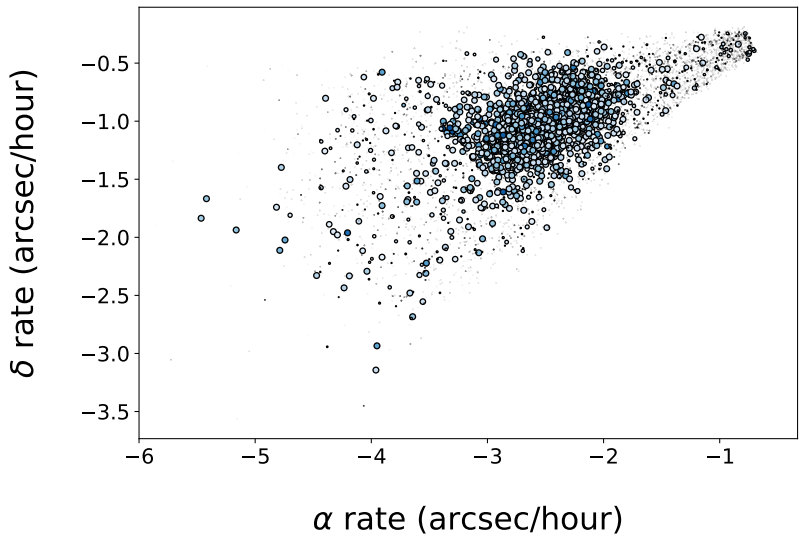


Figure 7.6: Sky moving rates of our candidate detections. The size of each marker is proportional to its weight, as calculated by Equation (7.4). The dense cloud of points corresponds mostly to CC detections (see Section 7.8).

detections in our fit using Equation (7.4), and maximize the likelihood given by

$$\ln \mathcal{L}(\theta) = \sum_i w_i \ln\left(\eta(m_i|\theta)\right) + (1 - w_i) \ln\left(1 - \eta(m_i|\theta)\right), \tag{7.6}$$

where θ is the vector of function parameters, and undetected fakes receive $w = 0$.

We display the best-fit for each night in Figure 7.7, and list the fit parameters in Appendix 7.A. Note that when fitting our data, we truncate our detection efficiency at m_{50}, and ignore all fainter detections.

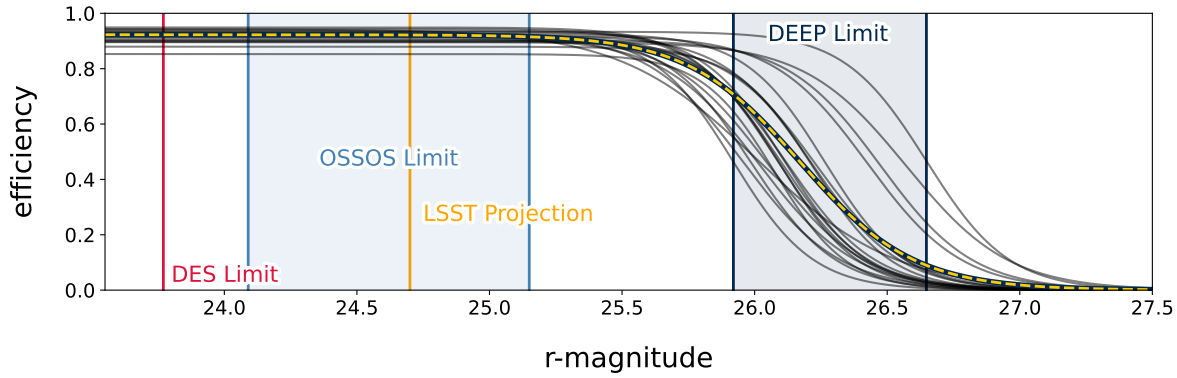


Figure 7.7: Recovery efficiency for implanted sources as a function of r-band magnitude. The combined efficiency for all 20 nights of data is given by the dashed line, while individual nights' efficiencies are given by the grey lines. The average m_{50} for our entire survey is $m_r \sim 26.2$, with individual nights ranging from 25.92–26.65. Our average peak efficiency, η_0, is $\gtrsim 0.92$, with individual nights ranging from 0.85–0.95. For reference, we also show limits from DES (Bernardinelli et al., 2022), OSSOS (Bannister et al., 2018) and LSST (Ivezić et al., 2019).

7.7 The Luminosity Function

We use our characterized detections to compute the differential sky density Σ of the Kuiper Belt as a function of apparent magnitude.[8] We reiterate that for this analysis we are using only single-epoch data (i.e., we have not linked these detections across multiple epochs), so we must treat each night as an independent survey.

For a given probability distribution $\Sigma(m)$, the expected number of detections by a survey is given by

$$\bar{N} = \Omega \int \eta(m)\Sigma(m|\theta)dm \tag{7.7}$$

where Ω is the survey's areal coverage and $\eta(m)$ is its detection efficiency. Next, the probability of randomly drawing an object with magnitude m from $\Sigma(m)$ is given by

$$P(m|\theta) = \int \Sigma(m'|\theta)\epsilon(m')\,dm' \tag{7.8}$$

where ϵ is a functional representation of the magnitude uncertainty for which we have adopted a Gaussian centered at m, with a width of δm.

[8] This quantity is colloquially referred to as the luminosity function.

We calculate the underlying luminosity function (Σ) of the Kuiper Belt by maximizing the likelihood ($\mathcal{L}$) given by

$$\mathcal{L}(\theta) = \prod_{k=1}^{n} e^{-\bar{N}_k} \prod_{j=1}^{N_k} P(m_{j,k}|\theta)^{w_{j,k}} \tag{7.9}$$

(see, e.g., Loredo 2004, Fraser et al. 2014) where the index k runs through each night of data, and the index j runs through a survey's detections. The variable θ is a vector of function parameters. The value $w_{j,k}$ denotes the weight of the jth object detected by the kth survey, as calculated by Equation (7.4).[9]

Several previous works have studied the form of the luminosity function for KBOs (Bernstein et al., 2004; Petit et al., 2006; Fraser et al., 2008; Fraser & Kavelaars, 2009; Fuentes et al., 2009). Following the example of these studies, we fit our data with functional forms of varying complexity. We first try a single power law given by

$$\Sigma_{single}(m) = 10^{\alpha(m-m0)}, \tag{7.10}$$

where m_0 is the magnitude at which the density of objects is one per square degree, and α is the power law slope. We next try a rolling power law given by

$$\Sigma_{rolling}(m) = \Sigma_{23} 10^{\alpha_1(m-23)+\alpha_2(m-23)^2}, \tag{7.11}$$

where Σ_{23} is the number of objects with $m_r = 23$ per square degree, while α_1 and α_2 control the shape of the function. We fit a broken power law given by

$$\Sigma_{broken}(m) = \begin{cases} 10^{\alpha_1(m-m_0)} & m < m_B \\ 10^{\alpha_2(m-m_0)+(\alpha_1-\alpha_2)(m_B-m_0)} & m \geq m_B \end{cases} \tag{7.12}$$

were m_0 is a normalization parameter, m_B is the magnitude at which the break occurs, and α_1 and α_2 are the bright-and-faint slopes, respectively. Finally we fit an exponentially tapered power law with the functional form

$$\Sigma_{taper}(m) = \exp\left[-10^{\beta(m_B-m)}\right] 10^{(\alpha-\beta)m-(m_0\alpha)} \left[\alpha 10^{m\beta} + \beta 10^{m_B\beta}\right] \tag{7.13}$$

[9] By applying a weight to each of our candidate detections, we are properly accounting for false positives in our data. This weight term, which is not present in other KBO analyses of this nature, was first proposed by Manski & Lerman (1977).

where α is the faint-end power law slope, β is the strength of the exponential taper, m_0 is a normalization parameter, and m_B is the absolute magnitude at which the exponential taper begins to dominate.

In each case, we obtain a best-fit using an optimizer, and then estimate our uncertainties by running an MCMC. We then use a survey simulation technique to test the quality of our fits. We randomly sample a population of objects from our best-fit magnitude distribution, and then impose the detection criteria of our survey. After we simulate our detections, we construct a simulated empirical distribution. By repeating this process many times, we end up with a distribution of simulated empirical distribution functions that reflects the distribution of possible outcomes for our survey if our best-fit is the underlying truth. Then for each simulated survey, we calculate a test statistic by computing its quantile among the set of simulated surveys as a function of magnitude, and calculating the fraction of the parameter space for which the survey was outside of the 95th percentile. We then calculate the same test statistic for our actual detections, and compute its quantile among the set of simulated test statistics, which we call Q_{outlier}. Values of $Q_{\text{outlier}} > 0.95$ indicate a poor fit. Finally, we compute the Bayes Information Criterion (BIC) for each of our fits. In general, lower values of BIC indicate a preferred model. We demonstrate our results in Figure 7.8, and summarize them in Table 7.1.

Functional Form	Best-Fit Parameters	BIC	Q_{outlier}
Single Power Law	$\langle \alpha, m_0 \rangle = \langle 0.59^{+0.01}_{-0.01}, 23.20^{+0.05}_{-0.05} \rangle$	18.4	0.998
Rolling Power Law	$\langle \alpha_1, \alpha_2, \Sigma_{23} \rangle = \langle 0.78^{+0.05}_{-0.05}, -0.05^{+0.01}_{-0.01}, 0.56^{+0.07}_{-0.06} \rangle$	4.0	0.74
Broken Power Law	$\langle \alpha_1, \alpha_2, m_0, m_B \rangle = \langle 1.02^{+0.17}_{-0.14}, 0.54^{+0.02}_{-0.02}, 23.36^{+0.06}_{-0.06}, 23.80^{+0.25}_{-0.19} \rangle$	0	0.49
Tapered Power Law	$\langle \alpha, \beta, m_0, m_B \rangle = \langle 0.22^{+0.06}_{-0.05}, 0.11^{+0.03}_{-0.02}, 13.12^{+4.11}_{-6.39}, 29.97^{+2.30}_{-1.89} \rangle$	13.1	0.94

Table 7.1: Best-fit parameters and statistics for each of the distributions we tested on the full KBO sample luminosity function. The BICs have been normalized such that the minimum value among the distributions is 0.

First we note that we strongly rule out the single power law. The rolling, broken, and exponentially tapered power laws all provide acceptable fits (though the exponentially tapered power law fit is marginal). Judging by the BIC, the broken power law is marginally preferable to the

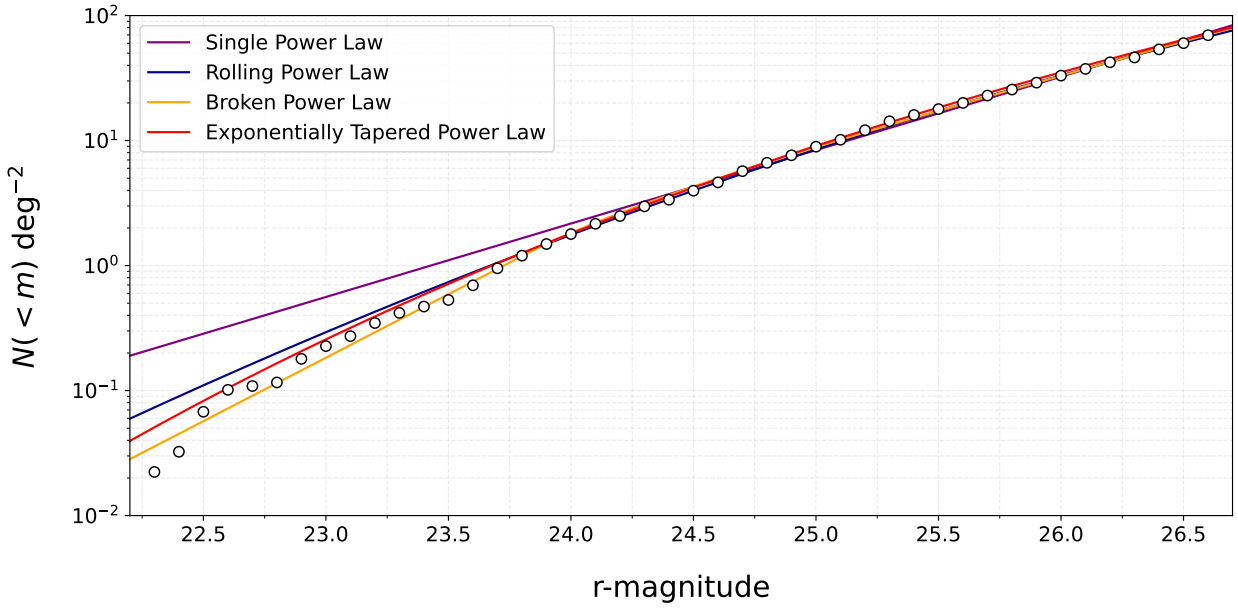


Figure 7.8: Luminosity function for the full DEEP KBO sample. Best-fit cumulative distributions for the single, rolling, broken, and exponentially tapered power laws are shown in purple, blue, yellow, and red respectively. The points and 1-σ error bars represent the cumulative distribution of our detections, corrected for efficiency and weight. Note that the points are meant only as a visual aid; we always use the maximum likelihood technique to fit our data.

rolling and exponentially tapered power laws, but we contend that a dearth of statistics faintward of $m_r = 24$ limits the usefulness of such comparisons. There is also no overwhelming physical motivation for choosing between the distributions. While Kavelaars et al. (2021) showed that the absolute magnitude distribution of the CCs is well-described by an exponentially tapered power law, there is no reason *a priori* that it should fit the full Kuiper Belt luminosity function. In fact, we might expect the full Kuiper Belt luminosity function to be fit poorly by an exponentially tapered power law, because it is likely a mix of multiple distributions. However, as we show in Section 7.8, our detections are dominated by CCs. It is therefore unsurprising that the exponentially tapered power law yields a reasonable fit. Given these considerations, we opt not to choose a preferred model, and instead claim for the time being that all three distributions provide acceptable fits to the DEEP detections. Future DEEP studies will be able to make more definitive statements.

7.8 Isolating a Sample of Cold Classicals

While our single-night detections do not provide nearly enough information for secure dynamical classification, we can use their apparent rates of motion to make assumptions about their dynamical classes. Consider the space of allowable sky motions for bound orbits shown in Figure 7.9. As before, the large black outline contains the space of possible rates of motion for objects on bound orbits at topocentric distances of 35–1000 au. The blue region is the allowable parameter space for CCs (42.4 au $< a <$ 47.7 au, $0 < e < 0.2$, $0^\circ < i < 4^\circ$), and the red region is the allowable parameter space for a somewhat arbitrary definition of Hot Classicals (HCs) that we are only using for the purpose of demonstration (42.4 au $< a <$ 47.7 au, $0 < e < 0.2$, $4^\circ < i < 45^\circ$). Note that the exact shape and orientation of these regions varies as a function of sky position and epoch.

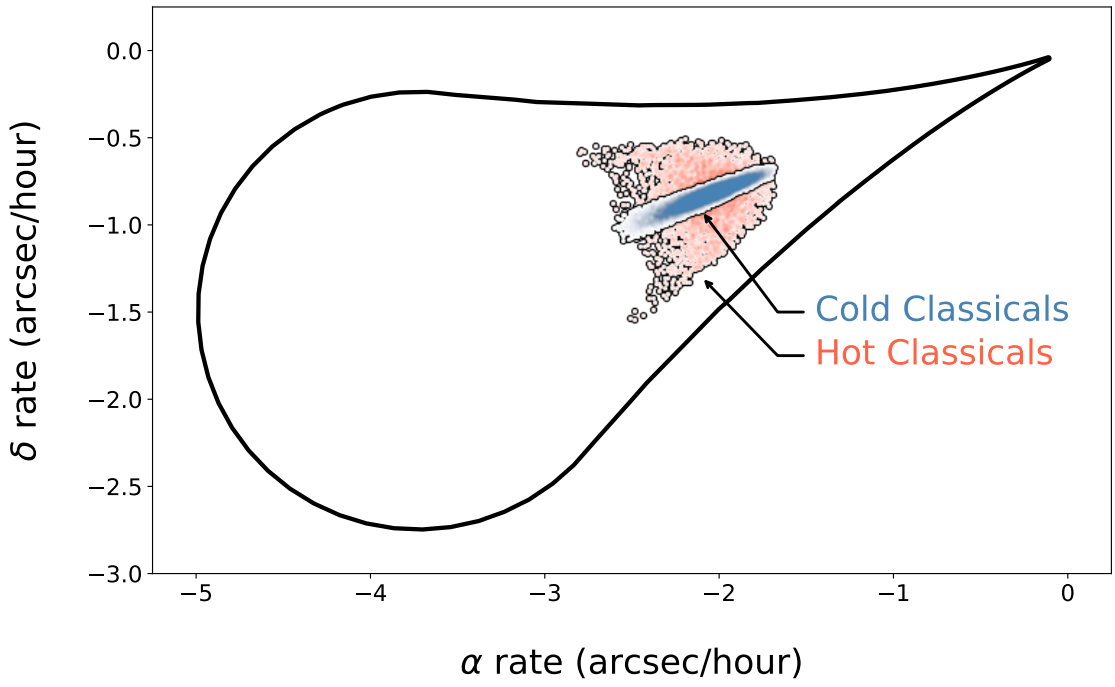


Figure 7.9: Sky motion parameter space for simulated CCs and HCs in B1a on 2020-10-18.

In practice, there is some overlap in the rates of motion between different dynamical classes, which is accentuated by the uncertainty in our rate measurements. We use survey simulation to

isolate the CCs in our detections. Using the OSSOS++ Kuiper Belt model[10], we calculate which objects our survey would have detected, and apply a smear in the RA and Dec rates consistent with our real detections. For each simulation we use a kernel density estimator to draw a region containing 95% of CCs. We find that we can isolate a fairly pure sample (on average $\sim$ 70% purity; see Table 7.2) of CCs from our single-night data. This purity is somewhat serendipitous, as the DEEP B1 fields happen to be in a patch of sky that is relatively uncontaminated by objects in resonance with Neptune. We speculate that the uncertainties inherent to the distribution of objects in the OSSOS++ model are larger than the purely statistical uncertainties listed in Table 7.2. Nevertheless, these values should provide a realistic estimate of the purity of the CC samples that we attempt to isolate for the analysis in Sections 7.9 and 7.10.

Finally, we obtain a distance estimate and uncertainty for each of our CC detections. This estimate relies on the fact that in the regime of the CCs, the relationship between an object's heliocentric distance and the inverse of its apparent rate of motion can be well-approximated as linear. For each night of data, we project our CC population model (obtained from the OSSOS++ model) into the space of RA rate and Dec rate, and fit a line relating distance to the inverse of the apparent sky motion. We then use the resulting relationship to compute the heliocentric distance of each of our detections. By sampling from the covariance of our detections' rates, we obtain Gaussian distance uncertainties of at most a few au, which end up being precise enough to fit an absolute magnitude distribution (Section 7.10).

7.9 The Luminosity Function of the Cold Classicals

We now consider the subset of our detections that have rates consistent with being CCs (see Section 7.8). For these fits, we follow the same general procedure as Section 7.7, but we must now account for the contamination in our sample. Note that we omit both occurrences of our B1f field due to the projected low purity of the isolated CC sample. We do this by sampling from each night's detections that have rates consistent with being a CC, accepting each detection with

[10]The OSSOS++ model is designed to be a realistic model of the Kuiper Belt, including correct relative population sizes of the different Kuiper Belt dynamical classes (Kavelaars et al., 2021; Crompvoets et al., 2022).

Night	Field	Total Detections	Potential CCs	CC Fraction	HC Fraction	Resonant Fraction
2019-08-27	B1c	89.94	69.91	0.68 ± 0.06	0.19 ± 0.04	0.12 ± 0.05
2019-08-28	B1a	78.39	64.56	0.74 ± 0.05	0.15 ± 0.03	0.11 ± 0.03
2019-08-29	B1b	160.83	131.18	0.74 ± 0.04	0.15 ± 0.03	0.10 ± 0.02
2019-09-26	B1a	81.58	65.10	0.75 ± 0.05	0.14 ± 0.04	0.11 ± 0.04
2019-09-27	B1b	77.85	56.96	0.74 ± 0.06	0.16 ± 0.05	0.09 ± 0.03
2019-09-28	B1c	118.20	88.98	0.68 ± 0.06	0.18 ± 0.04	0.13 ± 0.04
2020-10-15	B1b	68.72	53.86	0.71 ± 0.04	0.18 ± 0.04	0.10 ± 0.03
2020-10-16	B1c	106.54	85.13	0.66 ± 0.05	0.19 ± 0.05	0.13 ± 0.04
2020-10-17	B1e	120.69	103.62	0.75 ± 0.05	0.15 ± 0.04	0.09 ± 0.03
2020-10-18	B1a	124.90	102.37	0.71 ± 0.04	0.17 ± 0.04	0.11 ± 0.02
2020-10-19	B1d	54.50	42.09	0.66 ± 0.07	0.21 ± 0.05	0.12 ± 0.04
2020-10-20	B1f	56.72	38.71	0.57 ± 0.08	0.23 ± 0.05	0.18 ± 0.06
2020-10-21	B1d	76.85	53.24	0.66 ± 0.06	0.21 ± 0.05	0.11 ± 0.04
2021-09-27	B1d	68.32	58.31	0.68 ± 0.06	0.19 ± 0.05	0.12 ± 0.04
2021-10-01	B1b	110.33	85.49	0.71 ± 0.05	0.18 ± 0.05	0.10 ± 0.02
2021-10-02	B1f	80.02	56.98	0.55 ± 0.07	0.26 ± 0.07	0.18 ± 0.06
2021-10-03	B1i	114.10	92.96	0.69 ± 0.05	0.18 ± 0.03	0.12 ± 0.04
2021-10-04	B1c	84.95	61.39	0.68 ± 0.06	0.18 ± 0.03	0.13 ± 0.05
2021-10-05	B1h	109.67	90.02	0.74 ± 0.04	0.15 ± 0.04	0.10 ± 0.03
2021-10-06	B1e	103.83	81.73	0.73 ± 0.06	0.16 ± 0.05	0.10 ± 0.04

Table 7.2: Dynamical classification purity for DEEP detections. (Center) Weighted number of detections in each of our 20 fields, along with the weighted number of objects consistent with being CCs. Note that we have ignored all objects fainter than the m_{50} of the night in which they were detected. (Right) Fraction by dynamical class of the objects with rates consistent with being CCs, as determined by survey simulation.

a probability given by the CC fraction column in Table 7.2. We then do the maximum likelihood fit, and compute our uncertainties as in the previous sections. We repeat this procedure 100 times, in order to obtain a realistic estimate of the uncertainties caused by the contamination in our sample. For the fitting statistics, we now report the mean of BIC and $Q_{\rm outlier}$ from our simulations as $\langle {\rm BIC} \rangle$ and $\langle Q_{\rm outlier} \rangle$ respectively. We summarize our results in Table 7.3 and show the fits in Figure 7.10.

We again strongly rule out a single power law. As with the full set of KBOs, we cannot rule

Functional Form	Best-Fit Parameters	$\langle$BIC$\rangle$	$\langle Q_{\rm outlier} \rangle$
Single Power Law	$\langle \alpha, m_0 \rangle = \langle 0.58^{+0.02}_{-0.02}, 23.61^{+0.07}_{-0.07} \rangle$	14.0	0.998
Rolling Power Law	$\langle \alpha_1, \alpha_2, \Sigma_{23} \rangle = \langle 0.83^{+0.08}_{-0.08}, -0.07^{+0.02}_{-0.02}, 0.29^{+0.06}_{-0.05} \rangle$	2.6	0.49
Broken Power Law	$\langle \alpha_1, \alpha_2, m_0, m_B \rangle = \langle 1.19^{+0.39}_{-0.30}, 0.51^{+0.03}_{-0.03}, 23.62^{+0.11}_{-0.10}, 23.80^{+0.50}_{-0.20} \rangle$	0	0.44
Tapered Power Law	$\langle \alpha, \beta, m_0, m_B \rangle = \langle 0.17^{+0.05}_{-0.04}, 0.13^{+0.03}_{-0.03}, 12.01^{+3.98}_{-4.96}, 29.06^{+1.85}_{-1.34} \rangle$	11.4	0.77

Table 7.3: Best-fit parameters and statistics for each of the distributions tested for the luminosity function of the CC subsample of our detections. The BIC values have each been rescaled such that the minimum value among the distributions is 0.

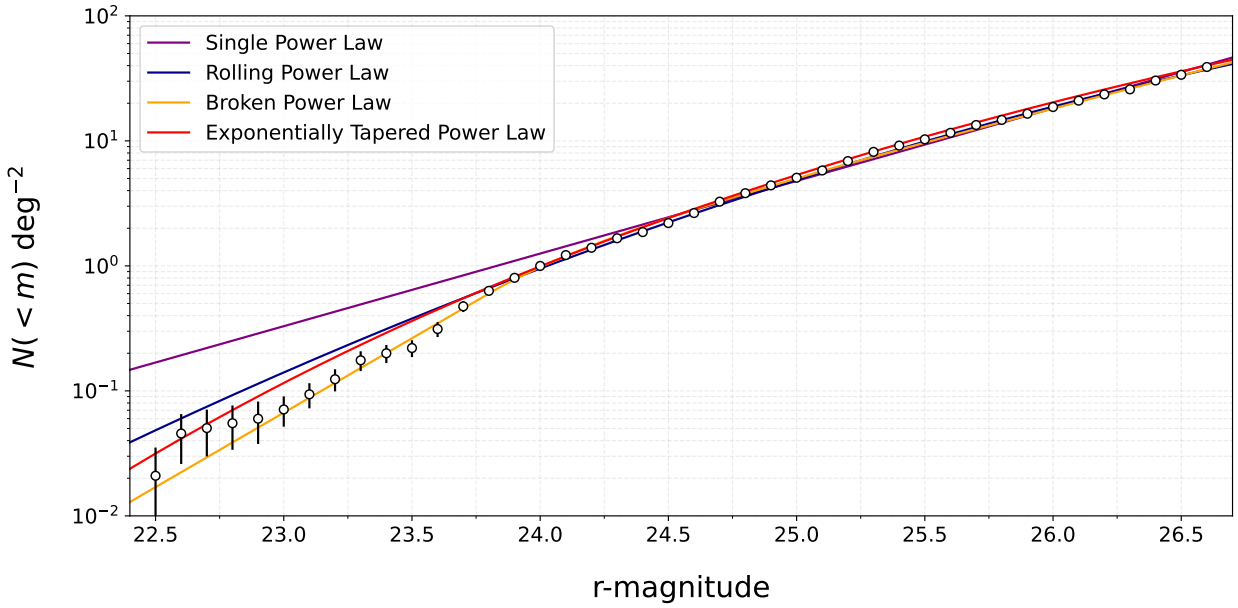


Figure 7.10: Best-fit cumulative distributions for the luminosity function of the cold classicals. The single, rolling, broken, and exponentially tapered power laws are shown in purple, blue, yellow, and red, respectively. The points and 1-σ error bars represent the cumulative distribution of our detections, corrected for efficiency and weight. Note that the points are meant only as a visual aid; we always use the maximum likelihood technique to fit our data.

out the rolling, broken, or exponentially tapered power laws. It is worth noting (though not at all conclusive) that the exponentially tapered power law provides a better fit to the luminosity function of the CC subsample than it did to the full KBO sample.

7.10 The Absolute Magnitude Distribution of the Cold Classicals

Determining the absolute magnitude (H) of an object requires simultaneous knowledge of its apparent magnitude and its heliocentric distance. While our detections have reasonably well-constrained apparent magnitudes ($\sim \pm 0.1$), their distances are not as well-constrained from single-night detections. However, as we found in Section 7.8, enforcing that a detection is a CC places a strong prior on its heliocentric distance, r, yielding distance estimates with tolerable uncertainty.

For a given probability distribution $\Sigma(H)$, the expected number of detections by a survey is given by

$$\bar{N} = \Omega \iint \eta(m)\Sigma(H(r, m)|\theta)\Gamma(r)\, dmdr, \tag{7.14}$$

where Ω is the survey's areal coverage and Γ is the underlying radial distribution of objects in the survey's field of view. Note that we use the OSSOS++ Kuiper Belt model (Kavelaars et al., 2021) to compute a kernel density for Γ. Since our fields are all observed near opposition, it is a good approximation to use

$$H = m - 5\log_{10}(r(r-1)). \tag{7.15}$$

Next, the probability of randomly drawing an object with magnitude m and heliocentric distance r from the distribution $\Sigma(H)$ is given by

$$P(m, r|\theta) = \iint \Sigma(H(r', m')|\theta)\epsilon(m')\Gamma(r')\gamma(r')\, dm'dr' \tag{7.16}$$

where ϵ is a functional representation of the magnitude uncertainty, and γ is a functional representation of the uncertainty in heliocentric distance. Note that we have taken ϵ to be a Gaussian centered at m, with a width of δm, and γ to be a Gaussian centered at r, with a width of δr.

We again use a modified version of the method described by Fraser et al. (2014) in which we maximize the likelihood given by

$$\mathcal{L}(\theta) = \prod_{k=1}^{n} e^{-\bar{N}_k} \prod_{j=1}^{N_k} P(m_{j,k}, r_{j,k}|\theta)^{w_{j,k}} \tag{7.17}$$

where the index k runs through the surveys (in our case individual nights of data), and the index j runs through a survey's detections. Note that we do the fit and treat the uncertainties using the same techniques as in Section 7.9.

We again fit single, rolling, broken, and exponentially tapered power laws. Note, however, that we have changed the definition of the rolling power law to

$$\Sigma_{rolling}(H) = \Sigma_8 10^{\alpha_1(H-8)+\alpha_2(H-8)^2} \tag{7.18}$$

where Σ_8 is the number of objects with $H_r = 8$ per square degree. We show our fits in Figure 7.11, and present our results in tabular form in Table 7.4.

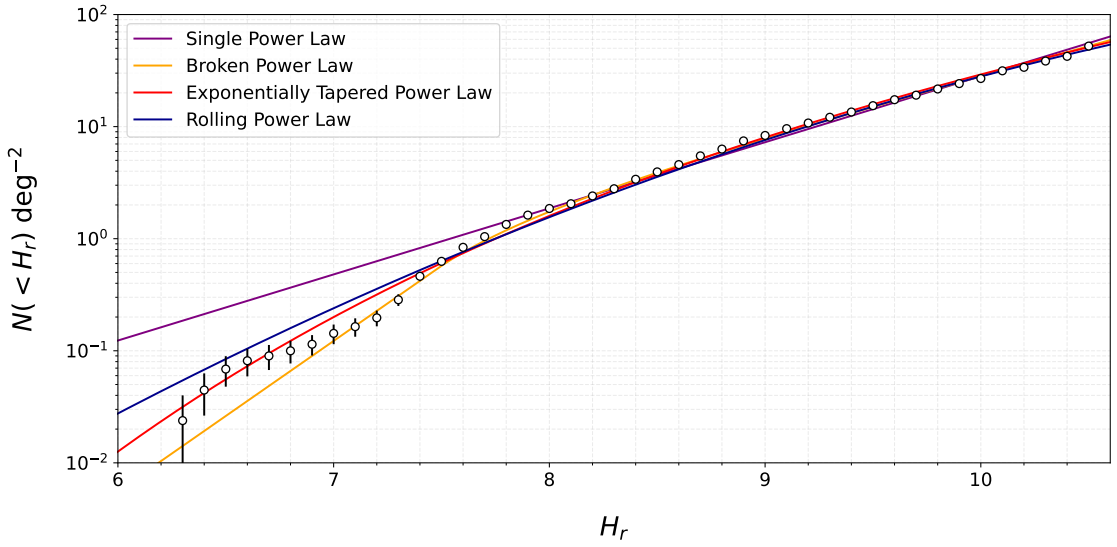


Figure 7.11: Best-fit cumulative distributions for the absolute magnitude distribution of the cold classicals. The single, rolling, broken, and exponentially tapered power laws are shown in purple, blue, yellow, and red, respectively. The points and 1-σ error bars represent the cumulative distribution of our detections, corrected for efficiency and weight.

While the BIC favors the broken power law, we don't believe that we can give preference to any distribution. This lack of constraining power is largely due to our dearth of statistics at the bright-end of the H distribution (we have no CC detections brighter than $H_r \sim 6.3$). In contrast, Kavelaars et al. 2021 found that the OSSOS++ detections were inconsistent with rolling and broken power laws due to a taper at the bright-end of the H distribution. The OSSOS++ model sensitivity was only possible due to the high purity of their sample and its well-characterized orbits.

Functional Form	Best-Fit Parameters	$\langle$BIC$\rangle$	$\langle Q_{\rm outlier}\rangle$
Single Power Law	$\langle \alpha, H_0\rangle = \langle 0.59^{+0.02}_{-0.02}, 7.32^{+0.07}_{-0.07}\rangle$	15.4	0.986
Rolling Power Law	$\langle \alpha_1, \alpha_2, \Sigma_8\rangle = \langle 0.68^{+0.04}_{-0.03}, -0.07^{+0.02}_{-0.02}, 2.69^{+0.20}_{-0.19}\rangle$	4.6	0.78
Broken Power Law	$\langle \alpha_1, \alpha_2, H_0, H_B\rangle = \langle 1.34^{+0.42}_{-0.28}, 0.53^{+0.03}_{-0.03}, 7.32^{+0.08}_{-0.10}, 7.48^{+0.21}_{-0.20}\rangle$	0	0.55
Tapered Power Law	$\langle \alpha, \beta, H_0, H_B\rangle = \langle 0.27^{+0.05}_{-0.04}, 0.17^{+0.04}_{-0.04}, 2.11^{+1.75}_{-2.09}, 10.98^{+1.28}_{-1.03}\rangle$	9.5	0.79

Table 7.4: Best-fit parameters and statistics for each of the distributions tested for the absolute magnitude distribution function of the CC subsample of our detections. The BIC values have each been rescaled such that the minimum value among the distributions is 0.

Our inability to rule out the rolling and broken power laws may be due to some combination of a relatively small survey area and contamination in our CC sample. On the other hand, the OSSOS++ sample only went as deep as $H_r \sim 8.3$ (at least two magnitudes brighter than DEEP), requiring them to assume a faint-end power law. In contrast, we constrain the faint-end of the H distribution rather well. Our measured faint-end slope, $\alpha = 0.27^{+0.05}_{-0.04}$, is consistent with SI simulations (Abod et al., 2019). In Figure 7.12 we show a comparison of our exponentially tapered power law fit with that of Kavelaars et al. (2021). The two results are consistent (i.e., the OSSOS result is within our 95% confidence region), and both surveys are consistent with the result of Bernstein et al. (2004), which is the deepest KBO survey to date.

The consistency of our detections with a rolling power law warrants careful consideration. First, we note that for appropriate parameters, a rolling power law is functionally equivalent to a Gaussian. All rolling power law fits we have presented in this work satisfy these criteria, and can therefore be equally well-represented as normal distributions. If the CCs are normally distributed in H, it would imply that they have a characteristic size, after which the H distribution turns over. While we currently have no reason to suspect that the size distribution turns over, our data cannot rule it out as a possibility. Furthermore, the only survey in the literature with data beyond our limit (Bernstein et al., 2004) cannot resolve the tension, as it is consistent with extrapolations from both the rolling and exponentially tapered power laws. While more precise measurements of the CC H distribution from forthcoming DEEP data will help to bring the true distribution into clearer focus, an even deeper targeted CC survey (Stansberry et al., 2021) will be required

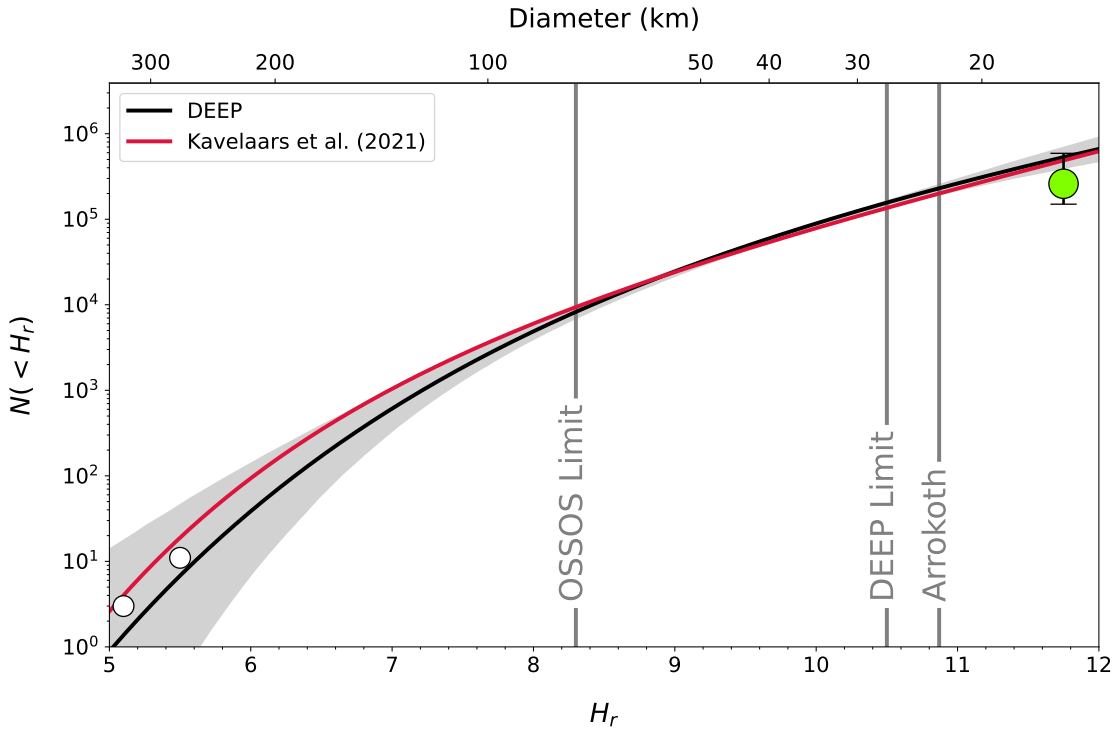


Figure 7.12: Comparison of the DEEP and OSSOS cold classical absolute magnitude distributions. The black line is our best-fit exponentially tapered power law, and the dark grey region represents a 95% confidence interval. Note that we have no detections brighter than $H_r \sim 6.3$. The red line is the best-fit from Kavelaars et al. (2021), and the white and green circles were taken from the same work. The white circles represent where the inventory of the CCs is considered complete, and the green circle is computed using the detections from Bernstein et al. (2004).

to completely resolve the tension.

Given our absolute magnitude distribution, we can calculate the total mass of the CCs as

$$M_{CC} = \frac{1}{F} \int \Sigma(H|\theta) M(H|p, \rho)\, dH \tag{7.19}$$

where F is an estimate of the average fraction of the total CC population per sq. deg. in one of our fields (calculated using the OSSOS++ model), and M is the mass of a body as a function of H given a geometric albedo, p, and a mass density, ρ. Assuming typical values of $p = 0.15$ and $\rho = 1$ g cm^{-3}, our fit to an exponentially tapered power law implies that the total mass of the CCs is $M_{CC} \approx 0.0022^{+0.0003}_{-0.0002} M_{\oplus}$.

7.11 Consistency with Deeper surveys

The only survey in the literature that is deeper than the present work is that of Bernstein et al. (2004), which reached $m_{50} = 29.02$[11] over a search area of 0.019 deg^2. Although the survey area is quite small, we can use it as a powerful lever arm to determine whether our fits remain valid down to $H_r \sim 12$. To do so, we simulate the survey of Bernstein et al. (2004) (B04 hereafter), assuming that our fits represent the true underlying CC H distribution. For each H distribution, we simulate the survey 10^5 times, and then ask whether the true number of detections by the survey ($N = 3$) is commensurate with the suite of simulations. In particular, we calculate $P(\leq N)$, the probability that the survey would have made fewer than or exactly 3 detections. For the broken power law we find $P(<= 3) < 0.001$, for the exponentially tapered power law we find $P(<= 3) = 0.03$, and for the rolling power law we find $P(<= 3) = 0.33$.

Thus, this test strongly rules out that our best-fit broken power law extends beyond the realm of our DEEP detections, into the realm of the B04 detections. Our best-fit exponentially tapered power law is only marginally consistent with the B04 detections, and our best-fit rolling power law is consistent with the B04 detections. We are hesitant to rule out the exponentially tapered power law due to uncertainties in our fit, and because of a curiosity of the B04 survey. Although the B04 survey had sensitivity as faint as $r \sim 29.02$, its faintest detection was $r \sim 28.23$. We believe there are two possible explanations for this issue. First, it is possible (though it would be impossible to determine without redoing a 20-year-old analysis) that B04 made a mistake when estimating their efficiency. The second, more interesting explanation is that the H distribution of the CCs flattens out (and possibly turns over) somewhere between the limits of DEEP and B04. While DEEP will not be able to resolve this issue, a very deep targeted CC survey should be able to answer the question quite easily.

[11] We are using the correction $r - m_{F606W} = 0.15$.

7.12 Discussion and Conclusions

In this paper we have presented our single-night detections from 20 nights of data in the DEEP B1 field. By using a shift-and-stack technique we were able to achieve an r-band depth of ~ 26.2 over approximately 60 square degrees of sky. Our data yielded 2297.9 single-epoch candidate detections, including 1849.8 objects fainter than $m_r \sim 25$—the most such objects ever reported in a single survey by more than an order of magnitude.

Our claim of fractional discoveries is a first for KBO science, and as such may seem peculiar. However, as we have shown in Section 7.4.2, a weighted treatment of our detections allows us to properly account for false positives. By accounting for false positives, we are able to make full use of the data from even our deepest nights, whose faintest detections cannot be recovered in another epoch. Additionally, our statistics remain reliable near the detection limit where false positives tend to accumulate.

Using our carefully-treated single-night detections, we computed the luminosity function of the Kuiper Belt as a whole, as well as the CC population, down to $m_r \gtrsim 26.5$. In both cases, we were able to rule out a single power law as the underlying distribution. We found that rolling, broken, and exponentially tapered power laws yielded acceptable fits, though low statistics at the bright end prevented us from giving overwhelming preference to any model.

The most significant scientific result of this work is a measurement of the absolute magnitude distribution of the CCs down to $H_r \sim 10.5$. While a dearth of bright objects limits our constraining power at the bright end, our plethora of faint detections enable us to tightly constrain the faint end of the distribution. Our detections are consistent with an exponentially tapered power law with a faint-end slope $\alpha = 0.27^{+0.05}_{-0.04}$. This faint-end slope is marginally shallower than in previous literature, implying the presence of somewhat fewer small objects than was previously expected. This shallow faint-end slope is consistent with simulations of planetesimal formation via the streaming instability. This is of particular interest because the theory predicts a physically motivated size distribution (as opposed to other size distributions which have simply been

engineered to adequately describe observations) of CCs that matches observations over the full range of sizes probed to date. However, we urge that these conclusions must be approached with caution. Of particular consequence to the conclusions from this work, limited resolution in SI simulations leaves the theoretical faint-end slope somewhat uncertain (see Kavelaars et al. (2021) for an in-depth discussion of the outstanding problems with the streaming instability as a complete theory of planetesimal formation). While the exponentially tapered power law H distribution is a good fit to our CC detections, we cannot rule out rolling or broken power laws. On the other hand, our broken power law fit is in severe tension with Bernstein et al. (2004). Both the exponentially tapered and rolling power law distributions are consistent with the results of Bernstein et al. (2004), implying that the H distribution continues to flatten out beyond the DEEP limit. Interestingly, our rolling power law fit would imply a characteristic size for the CCs, beyond which the probability density rolls over.

Finally, we note some limitations of this work that can be improved upon in future studies. The most obvious limitation is our use of single-night detections. Although we developed robust new techniques to account for false positives, linked detections with well-constrained orbits are still preferable. Since we used single-night detections in this work, our selection of a CC subsample of our detections was rather rough, relying heavily on the OSSOS++ solar system model and the serendipity that Neptune was near these fields, meaning that they were necessarily relatively devoid of resonant objects. Linked orbits will enable proper dynamical classification, thus reducing the uncertainties in our studies of individual dynamical classes. In forthcoming work, we will analyze three additional fields similar to the B1 field. We will link our detections, yielding well-determined orbits for the smallest known KBOs. Our catalog of discoveries will enable studies of Kuiper Belt populations to unprecedented depth, providing deep insight to the formation and evolution of our planetary system.

7.A Pointings

Table 7.A.1 gives the pointing for each night of data, as well as the detection efficiency parameters.

Night	Field	RA (J2000)	DEC (J2000)	N_{exps}	η_0	m_{50}	σ
2019-08-27	B1c	352.92015	-3.62314	103	0.94	26.11	0.31
2019-08-28	B1a	351.87745	-5.03572	102	0.95	26.08	0.32
2019-08-29	B1b	353.61898	-5.29744	101	0.93	26.65	0.31
2019-09-26	B1a	351.38115	-5.24000	95	0.91	26.02	0.36
2019-09-27	B1b	353.12082	-5.50267	97	0.94	26.09	0.33
2019-09-28	B1c	352.42444	-3.82797	96	0.92	26.42	0.36
2020-10-15	B1b	352.86350	-5.60964	103	0.88	26.22	0.29
2020-10-16	B1c	352.16593	-3.93469	96	0.90	26.47	0.35
2020-10-17	B1e	353.90520	-4.19542	96	0.90	26.23	0.34
2020-10-18	B1a	351.12145	-5.34617	91	0.90	26.57	0.39
2020-10-19	B1d	354.60670	-5.86869	99	0.92	25.99	0.32
2020-10-20	B1f	353.20607	-2.52153	91	0.93	25.92	0.30
2020-10-21	B1d	354.60635	-5.86972	99	0.92	26.05	0.36
2021-09-27	B1d	354.80112	-5.78794	89	0.93	25.96	0.36
2021-10-01	B1b	353.05852	-5.52967	81	0.90	26.10	0.29
2021-10-02	B1f	353.40057	-2.44119	89	0.94	26.01	0.52
2021-10-03	B1i	355.13703	-2.69892	100	0.93	26.26	0.30
2021-10-04	B1c	352.36053	-3.85436	102	0.92	26.12	0.33
2021-10-05	B1h	355.83970	-4.37136	97	0.92	26.21	0.33
2021-10-06	B1e	354.09928	-4.11514	97	0.85	26.08	0.27

Table 7.A.1: Pointings and efficiency statistics for the DEEP B1 fields. The positions of the fields at each epoch aim to track as many objects as possible by accounting for the effects of Earth reflex motion and TNO shear. Each exposure is 120 seconds long and is taken in the VR band. The final three columns show the best-fit of each night to Equation 7.5.

Chapter 8

A Serendipitous Search for Jupiter Trojans

8.1 Introduction

Although DEEP was designed as a Kuiper Belt survey, its observing strategy makes its data an excellent resource for studying other solar system populations. For example, it turns out that some of the fields discussed in Chapter 7 coincide with the sky position of a high-density region of Jupiter's L4 Trojan population (see Figure 8.1). When observed from Earth near opposition, Jupiter Trojans have sky moving rates of approximately 15"/hour. Given DECam's pixel scale of 0.2637", this rate of motion corresponds to approximately 1 pixel per minute, or 2 pixels during a 120-second DEEP exposure. Since our images have a PSF width of approximately 1" (or 4 pixels), the smear from a Trojan's motion during the exposure is small enough that *it is possible* to search for Jupiter Trojans in the DEEP images with only minimal loss in sensitivity compared to slower-moving objects such as KBOs.

Jupiter Trojans are one of the most populous dynamical classes in the solar system, with more than 10^4 multi-opposition objects currently listed in the Minor Planet Center. Some solar system formation models predict that Jupiter's Trojan regions have been populated in part by planetesimals scattered inward from the outer solar system during Neptune's outward migration (Tsiganis et al., 2005; Morbidelli et al., 2005). If it is true that the Jupiter Trojans formed in a similar region of the solar system to the KBOs, despite having been exposed to very different environments for the past few billion years, the two populations should share some properties such as composition, color, and size distribution.

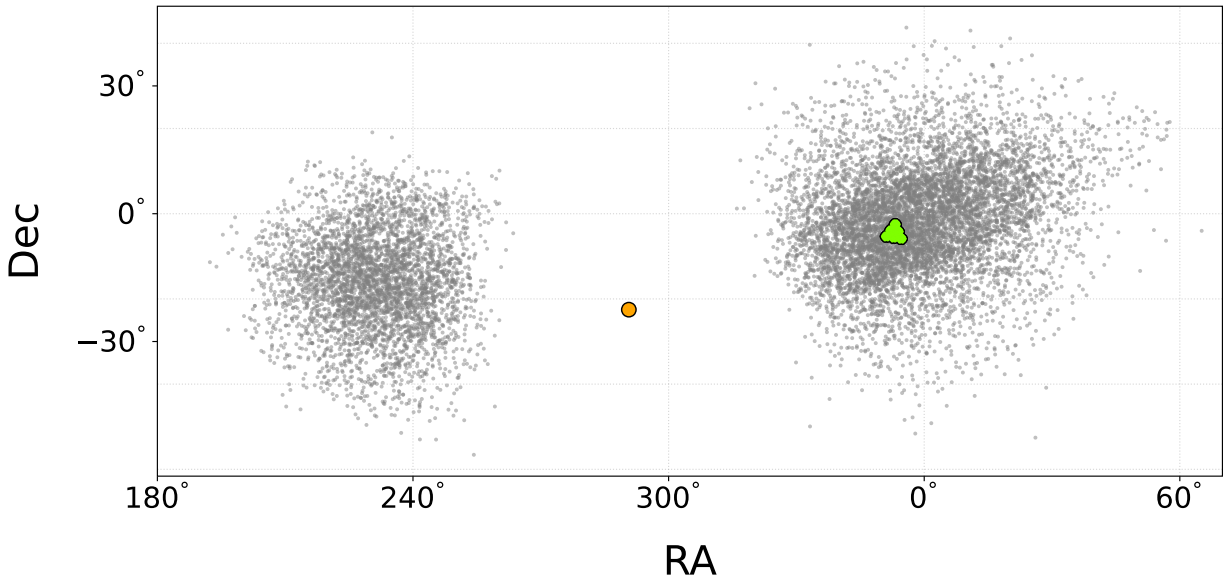


Figure 8.1: Sky positions of Jupiter and its Trojan clouds on October 18 2020. The grey points represent the sky positions of all known Jupiter Trojans. The group of points on the left is Jupiter's L5 Trojan cloud, and the group on the right is the L4 Trojan cloud. The large orange point represents the sky position of Jupiter on the same night. The green points represent DEEP's B1 pointings from 2020. By chance, the pointings are directly in the center of Jupiter's L4 Trojan cloud.

While the census of Jupiter Trojans is quite large, we still need more data to test theories of our solar system's formation. NASA's ongoing Lucy mission will make close flybys of 7 Jupiter Trojans of various sizes, shapes, and colors over the next decade, providing invaluable data to understand the origin of the Jupiter Trojans, and thus the formation of our solar system (Levison et al., 2021). The mission still has room to add more observation targets if any appropriate targets can be found (personal correspondence with Lucy PI Hal Levison). We believe that the scientific payoff from finding another observation target justifies the effort of developing techniques for finding fainter Jupiter Trojans. By finding fainter objects, we will be able to find more objects, thereby increasing the likelihood of finding a new observation target.

In this chapter we present a proof-of-concept shift-and-stack search for Jupiter Trojans in our DEEP B1a field. Note that our goal with this work was simply to *find* Jupiter Trojans; we leave all analysis of their properties for future work. We observed the field on October 18 2020 at

a solar elongation of approximately 145.5 degrees, (approximately one month after opposition); see Figure 8.2 for a demonstration of the observing geometry.[1] In Section 8.2 we describe the steps we took to pre-process our data, and carry out the search for Jupiter Trojans. We present preliminary results of our search in Section 8.3. Finally we conclude in Section 8.4 by discussing of the implications of our preliminary results and future directions for this work.

8.2 Data Processing

Our search strategy was nearly identical to the procedure described in Chapter 7, so we omit most of the details here, and instead give a broad overview.

We began by generating a synthetic population of Myr-stable Jupiter Trojans in our field, using the method described in Chapter 2. Altogether, we implanted 877 Trojans (approximately 15 per CCD), with *r*-band magnitudes between 23.0 and 27.0. We did not design the orbital distribution of the synthetic bodies to be realistic, but rather to cover the parameter space of sky motion possible for real Jupiter Trojans.

After we generated synthetic sources, we ran our images through the pre-processing pipeline described in Chapter 7. We did not generate the "bright" images for this analysis (recall that for the "bright" images we did not mask pixels brighter than 2σ), as the point of this experiment was simply to test the shift-and-stack technique on faint Jupiter Trojans.

We next computed the shift-and-stack rates of interest. We decided to search for all prograde orbits at barycentric distances beyond 4 au. Opening up our search to closer distances resulted in an enormous parameter space compared to the KBO search in Chapter 7, necessitating 13087 stack rates to keep the maximum trailing under 1", with an additional 13087 stack rates required for the reverse shift-and-stack. Since there were so many stack rates, we opted to fill the space of bound orbits with a hexagonal grid, rather than using the K-means clustering algorithm described in Chapter 7 which would have required a large amount of computation for a relatively small reduction in the number of grid points.

[1] As an aside, if we had chosen to observe a field directly at opposition, we would have been looking at a region of the sky with a much lower density of Jupiter Trojans.

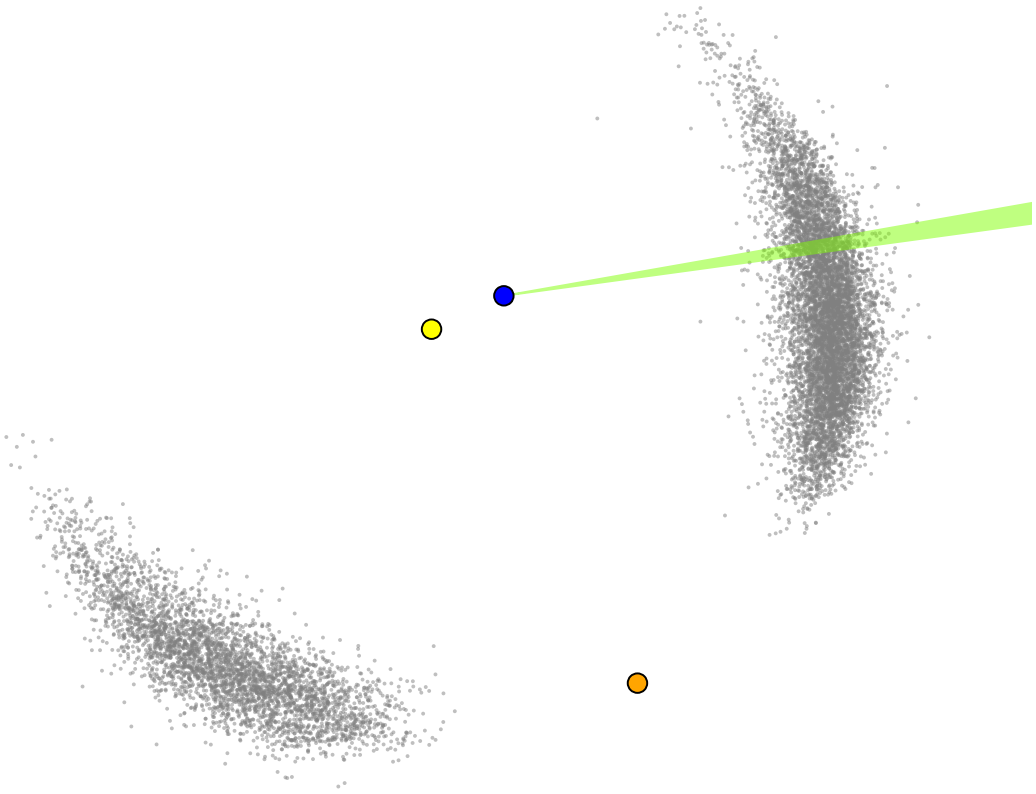


Figure 8.2: Positions of Jupiter and its Trojan clouds projected into the ecliptic plane. The grey points represent the positions of all known Jupiter Trojans on October 18 2020. The group of points leading Jupiter's orbit (Jupiter orbits counterclockwise in this orientation) is Jupiter's L4 Trojan cloud, and the trailing group the L5 Trojan cloud. The large orange point represents Jupiter's position on the same night, while the yellow point represents the Sun, and the blue point represents Earth. The green region represents the region of space subtended by DEEP's B1a pointing on October 18 2020.

We then stacked our data, and for each rate we extracted all sources brighter than 2σ, scored each source with a CNN, and wrote sources that passed the CNN to a catalog. Finally in order to avoid repeat detections of the same object, we used a `DBSCAN` algorithm to cluster sources in the space of RA, Dec, RA rate and Dec rate. Our final product was a catalog of 9250 candidate

sources including synthetic sources and sources from the reverse shift-and-stack (down from 2×10^7 sources before any filtering).

Since the goal of this experiment was simply to show that shift-and-stack is an effective technique for discovering faint Jupiter Trojans, we decided to forego the vetting process described in Chapter 7, and move on with examining our data.

8.3 Results

8.3.1 Detection Efficiency

We calculated our detection efficiency using the same functional form as in Chapter 7. Since we did not go through the human-vetting process for our candidate detections, we could not use the weighing scheme developed in Chapter 7. Rather, we used a binary scheme in which we gave synthetic sources with an associated detection a weight of 1, and synthetic sources without an associated detection have a weight of 0. As a result, the detection efficiency may be slightly inflated. On the other hand, the analysis in Section 8.3.2 suggests that the false positive rate is fairly low for fast-moving bodies such as Jupiter Trojans, so the calculated efficiency is probably not too far from the truth. Note that we ignored objects in our catalog of synthetic sources brighter than $m_r = 24.5$, as all such sources were at least partially masked during our image preparation. Our best fit gave $m_{50} = 26.51$, $\eta_0 = 0.77$, and $\sigma = 0.29$. We display our efficiency function in Figure 8.3.

We found that our peak detection efficiency was somewhat lower for Jupiter Trojans than it was for the KBOs in Chapter 7. Since the discrepancy exists even for our brightest detections, it cannot be purely due to Jupiter Trojans smearing more than KBOs during the 120-second exposures. Rather, it seems that the Jupiter Trojans move into chip gaps more frequently than KBOs due to their faster moving rates. A typical Jupiter Trojan moves at approximately 15"/hour, or approximately 1' over the course of our 4-hour exposure sequence. Each DECam CCD subtends a width of approximately 17.6' and a height of approximately 9.7', meaning that we would naively

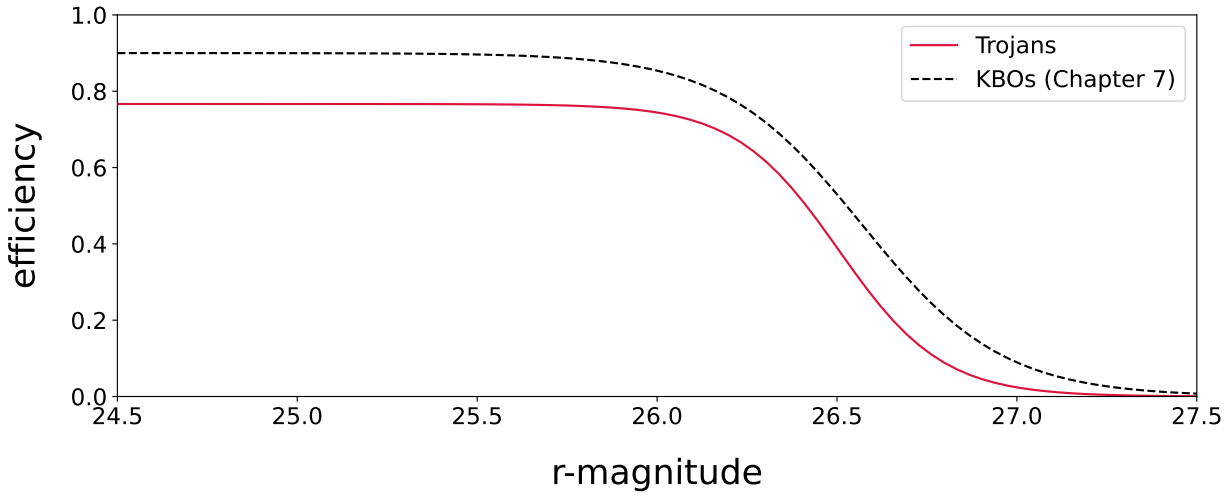


Figure 8.3: Detection efficiency for Jupiter Trojans (red) and KBOs (black, taken from Chapter 7) in DEEP B1a.

expect to lose $\sqrt{(1/17.6)^2 + (1/9.7)^2} \sim 12\%$ of our peak sensitivity. For the analysis in Chapter 7, B1a on October 18 2020 had $\eta_0 = 0.9$, meaning that we would expect $\eta_0 \sim 0.9 \times 0.88 \sim 0.79$ for Jupiter Trojans. This prediction is similar to our best-fit value of $\eta_0 = 0.77$.

8.3.2 Detections

Although we did not visually inspect all of our detections, we can learn a lot by just examining their rates of motion on the sky. We plot all of our detections' moving rates in Figure 8.4. In this figure, detections with S/N < 8 are plotted as grey points, and detections with S/N ≥ 8 are plotted as red points.

First, we note by looking at the sources from the reverse shift-and-stack that the false positive rate sharply decreases as the rate of motion increases to more than a few arcseconds per hour. Next we note some features present in the detections from the forward shift-and-stack that are absent from the reverse shift-and-stack. We identify these features as the KBOs, the Jupiter Trojans, and the Hildas.[2] It appears that there is also a more amorphous overdensity of objects in the forward shift-and-stack that likely corresponds to main-belt asteroids.

[2]The Hildas are in an internal 2:3 mean-motion resonance with Jupiter.

The fact that these dynamical features are so apparent without any human vetting is encouraging for the feasibility of future searches. In fact, we estimate that our true positive rate in the region of the Jupiter Trojans is over 90%. It would not have been possible to make out these structures without the use of our CNN for filtering out false positives.

8.4 Future Work

Unfortunately, the cadence of the DEEP data is unsuitable for linking Jupiter Trojan detections between epochs; the bodies are simply moving much faster on the sky than the KBOs that DEEP was designed to track. It is still possible, however, to study the properties of our single-night discoveries using the techniques we developed for the analysis in Chapter 7. We plan to search the rest of our B1 fields for Jupiter Trojans, and then do an analysis similar to the one presented in Chapter 7. We will attempt to isolate the Jupiter Trojans from our single-night detections using their rates of motion, and measure their absolute magnitude distribution. If successful, this measurement will be two magnitudes deeper than the current state-of-the-art (Uehata et al., 2022; Yoshida et al., 2019), pushing into the sub-kilometer size regime.

Finally, given the success of this project, we have begun the process of planning a dedicated search for a new Lucy target. Toward this end, we were awarded 2.5 nights on DECam in 2023 to search for L5 Jupiter Trojans.

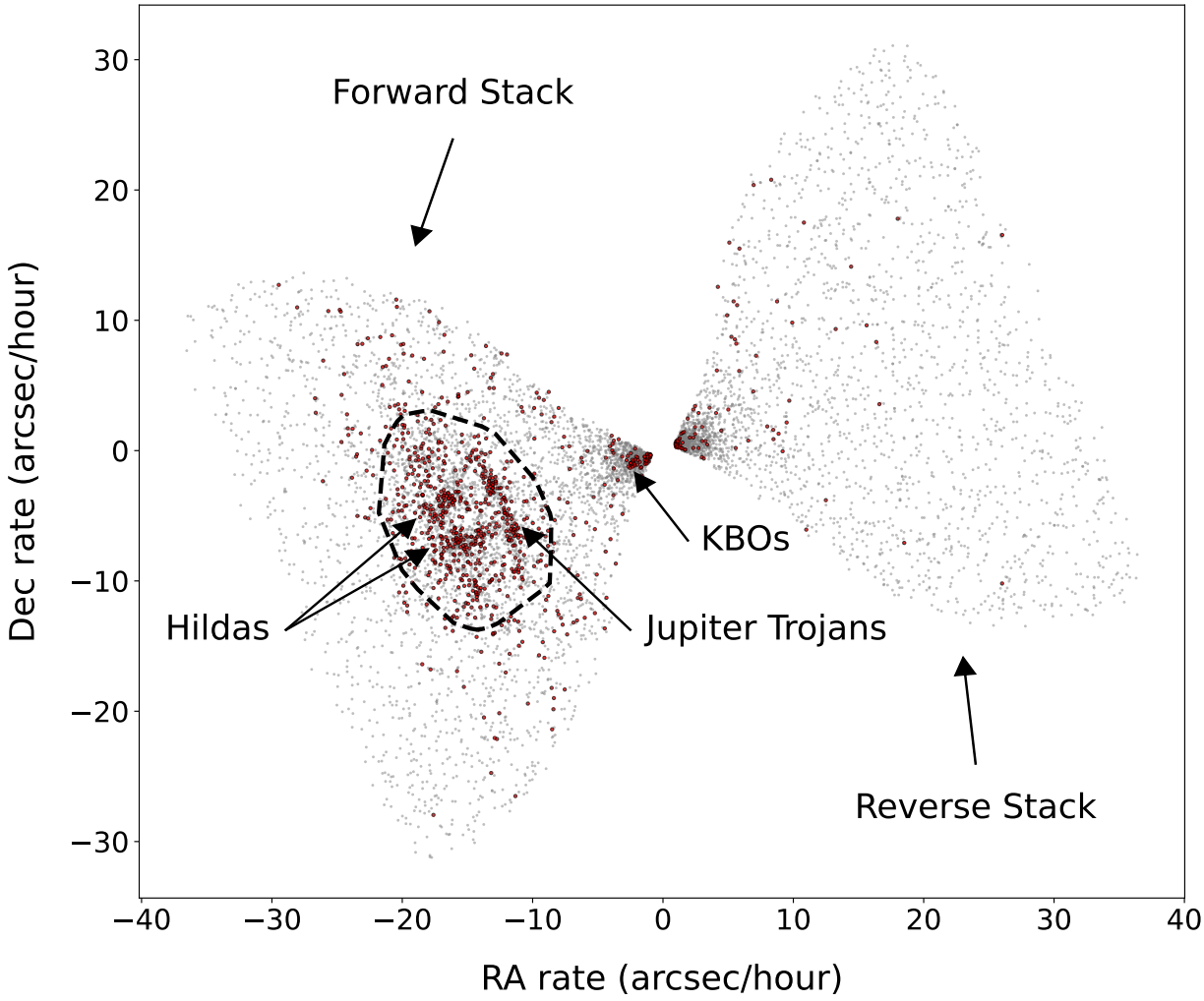


Figure 8.4: Sky moving rates of all detections. The grey points represent detections with S/N < 8, and the red points represent detections with S/N ≥ 8. The region on the left corresponds to the forward shift-and-stack, while the region on the right corresponds to the reverse shift-and-stack. The dynamical groups of KBOs, Jupiter Trojans, and Hildas are all visible. The black dotted line roughly encloses a region of high density.

Chapter 9

Project Hail Mary: A Ten Trillion Yard Touchdown Pass

It ain't over 'till it's over.

— Yogi Berra

9.1 Introduction

NASA launched its New Horizons (NH) spacecraft in 2006 with a mission to study Pluto and its moon Charon. When the spacecraft arrived at the system in 2015, it returned images that transformed our understanding of the icy world and all those like it (Stern et al., 2015). In 2019, as NH continued on its way out of the solar system, it executed a flyby of the Kuiper Belt Object (KBO) Arrokoth (Spencer et al., 2020). The images it sent back after the encounter have offered us our deepest insight into planetesimal formation and the nature of our primordial solar system. Now at a distance of 55 au, well beyond the Kuiper Belt and traversing the farthest reaches of the observed solar system, the spacecraft has enough fuel remaining to make one final divert maneuver to observe an object—if one can be found. Even if no encounter target can be found, there is still scientific value in finding KBOs that can be observed at a distance from the spacecraft at a high phase angle using its LORRI camera. Such observations have enabled us to study KBO shapes and surface properties in a way that is not possible from Earth (Porter et al., 2016; Verbiscer et al., 2022).

9.2 Data

To search for a flyby target, we obtained four half-nights of data on DECam in the vicinity of NH's sky position.[1] Much like the DEEP data presented in Chapter 7, each half-night consists of approximately 100 exposures of the same field, each with a duration of 120 seconds. We chose a cadence of two nights separated by 48 hours in June of 2021 and then two more nights separated by 48 hours in August of 2021. The reasoning for our observing cadence was that temporally close nights would help to disambiguate false positives from real moving objects, and the two-month gap between pairs of nights would be long enough to get a rough orbit determination for any new discoveries, but not so long that the process of linking detections would suffer from confusion. We list each night's pointings and metadata in Table 9.1.

Night	RA (J2000)	DEC (J2000)	N_{exps}
2021-06-11	287.2205	-19.8092	132
2021-06-13	287.2203	-19.8101	128
2021-08-10	285.8163	-19.9140	94
2021-08-12	285.7823	-19.9181	122

Table 9.1: Pointings and metadata for the NH fields.

We pre-processed the data using the same pipeline described in Chapter 7, with the key difference that we only implanted synthetic sources at barycentric distances between 60 and 100 au, with r-band magnitudes between 26.0 and 28.0. We chose such a limited parameter space for our synthetic objects because we were only concerned with quickly finding any objects that would be encounterable by the spacecraft, rather than doing comprehensive population studies.

We note that the data we obtained was particularly challenging to analyze, as our pointings' low galactic latitude and proximity to the galactic center resulted in a high density of stars in our images. For a demonstration of the stellar crowding, see Figure 9.1 for a template of one of

[1] Our search was part of a broader effort led by the New Horizons team, which used more than a dozen nights of time on the Subaru telescope.

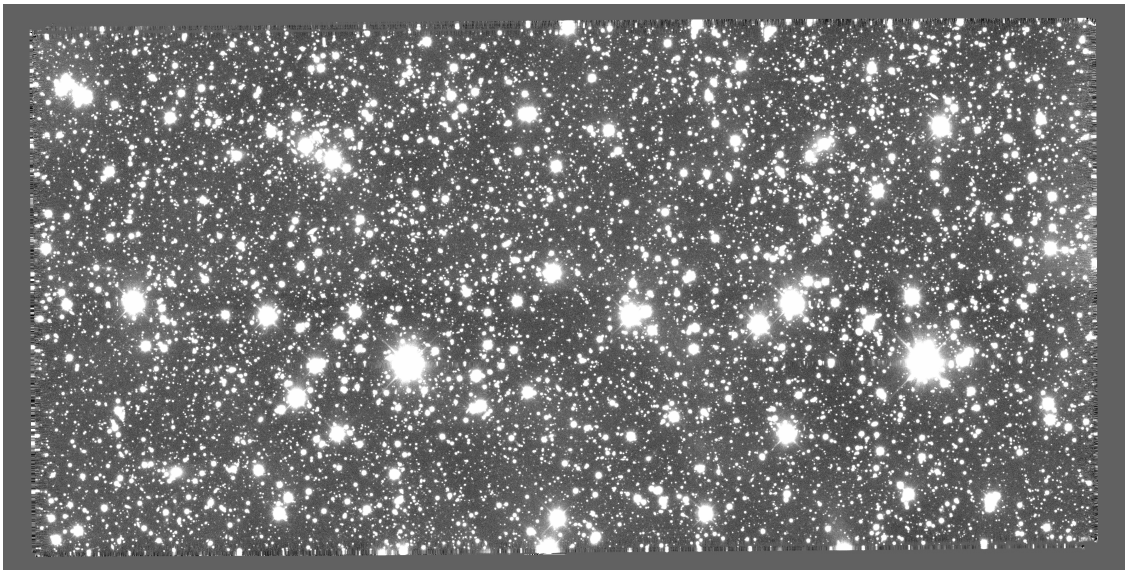


Figure 9.1: Template image for one CCD in the New Horizons target search. Notice the extreme crowding in the field by bright stars.

our CCDs. To make matters worse, the seeing conditions under which we obtained our images were not good, even by DECam's relatively low seeing standards. As a result, we had to mask approximately half of our pixels.

9.3 Analysis

9.3.1 Generating a Source Catalog

With our data pre-processed, we continued with the shift-and-stack procedure as described in Chapter 7. Since we were only interested in finding objects encounterable or observable by the NH spacecraft, we created a grid of shift-and-stack rates covering the parameter space of prograde orbits beyond 50 au. Instead of filling our grid such that the maximum trailing was 1 PSF width as described in Chapter 7, we sampled a denser grid such that the maximum trailing was one DECam pixel, or approximately 0.2637 arcseconds.[2] Then for each rate in our grid, we stacked our images, extracted all peaks brighter than 3σ, scored each peak with a CNN, and wrote the full catalog of sources to a file. After we finished with the stacking and extraction, we used

[2]Note that we computed a unique grid for each night.

the scores from our CNN to pare our catalog such that we retained most of our fakes, while eliminating as many false positives as possible.[3] Finally we used a `DBSCAN` algorithm to cluster the remaining sources in sky position on a night-by-night basis, so as to prevent sources from appearing multiple times in adjacent stack rates.

At this point we realized that it was a mistake to generate our synthetic sources such that the brightest object had an r-band magnitude of 26.0, as our detection efficiency in each night fell to $\sim 10-40\%$ at $m_r = 26.0$. For this reason, it was not particularly useful for us to study our recovery efficiency.[4] On the other hand, the presence of the implanted synthetic sources we did detect gave us confidence that our algorithms were working as intended. Since our goal was to find objects encounterable or observable by NH rather than doing a population study, we chose to continue with our analysis in spite of a dearth of useful fakes.

Our final source catalog contained 88,225 total sources, with 55,579, 7537, 15,553, and 9556 detections from the four nights respectively. For each detection in the catalog, we measured α, δ, $\dot{\alpha}$, $\dot{\delta}$, epoch, and S/N. Since we extracted sources down to 3σ, the catalog was composed almost entirely of false positives that were indistinguishable from real sources in single-epoch data.

9.3.2 Linking Sources

To disambiguate false positives from real detections while to maintaining sensitivity to faint objects, we turned to a technique called linking. The goal of linking is simply to determine which detections are actually repeat detections of the same object. In principle, it is possible to link sources using the brute force approach of generating all unique permutations of detections in our catalog (plus the null detection), and then trying to fit each triplet or quadruplet to a Keplerian orbit. Candidate objects that are moving on Keplerian orbits will result in good fits, while any given candidate composed of false positives is unlikely to result in a good fit (though with a large enough number of single-epoch false positives, false positive linkages certainly happen). Doing a

[3] Due to time pressure, we completed this step by hand in a rather ad-hoc manner.

[4] In principle, we could generate a brighter set of synthetic objects and redo the analysis, but the relatively poor quality of our data and the results of Chapter 9.4 suggest that it is not worthwhile to do so.

brute force search in our catalog would have required us to check more than 10^{16} orbits, which is not computationally reasonable. Instead, we took an approach inspired by the `HelioLinC` algorithm (Holman et al., 2018).

Consider a system with an origin, an observer, and an object that is being observed. Let the vector $\mathbf{R}$ denote the position of the observer with respect to the system's origin in some fixed frame. For example, in a common scenario the origin is the solar system barycenter, the fixed frame is the ICRS, and $\mathbf{R}$ points to some observatory on the surface of Earth. Let the distance between the object and the observer be given by ρ, and let the direction from the observer to the object be given by the unit vector $\mathbf{p}$. If we choose to work in equatorial coordinates, $\mathbf{p}$ is given by Equation (9.1) and its derivative $\dot{\mathbf{p}}$ is given by Equation (9.2). Note that $\dot{\mathbf{p}}$ is *not* necessarily a unit vector.

$$\mathbf{p} = \langle \cos\delta \cos\alpha, \cos\delta \sin\alpha, \sin\delta \rangle \tag{9.1}$$

$$\dot{\mathbf{p}} = \langle -\dot{\alpha} \cos\delta \sin\alpha - \dot{\delta} \sin\delta \cos\alpha, \dot{\alpha} \cos\delta \cos\alpha - \dot{\delta} \sin\delta \sin\alpha, \dot{\delta} \cos\delta \rangle \tag{9.2}$$

Then one can write down the state vector of the object as follows.

$$\mathbf{x} = \mathbf{R} + \rho\mathbf{p} \tag{9.3}$$

$$\dot{\mathbf{x}} = \dot{\mathbf{R}} + \dot{\rho}\mathbf{p} + \rho\dot{\mathbf{p}} \tag{9.4}$$

An important subtlety is that Equations (9.3) and (9.4) give the state of the rock at the epoch when it reflected the photons that were observed. Thus to get the true state of the rock at the epoch of observation, one must correct for light travel time.

Notice that we have measured four of the six parameters (α, δ, $\dot{\alpha}$, and $\dot{\delta}$) in the formalism described above. Then, to have a fully determined orbit, each detection requires only ρ and $\dot{\rho}$. Following the example of the `HelioLinC` algorithm, we assert that each of our detections will have a given value of ρ and $\dot{\rho}$ at a reference epoch, t_0. Then for each of our detections, we know six pieces of information: $\langle \rho, \dot{\rho} \rangle$ at t_0 and $\langle \alpha, \delta, \dot{\alpha}, \dot{\delta} \rangle$ at the observation epoch. Then for each of our detections we solve iteratively for the values of ρ and $\dot{\rho}$ at the epoch of observation,

thus fully defining the orbit.[5] Then with the orbit of each detection fully determined, one can propagate all orbits to t_0, and compute their sky positions as observed from the solar system barycenter. Repeated detections of the same object cluster on the sky at t_0 as one approaches the true values of ρ and $\dot{\rho}$.

In practice, we find that we need to scan through only a few values of ρ, and a few values of $\dot{\rho}$ to exhaustively link all detections between 50 and 100 au. For each $\langle \rho, \dot{\rho} \rangle$ pair, we calculate the orbit of each of our detections, and propagate the orbits to t_0.[6] We then used a k-d tree to find all groups of detections within some angular radius, which we took to be 120 arcseconds. For each group, we calculated all possible permutations of three detections at unique epochs, resulting in more than 2×10^8 candidate orbits.

We then further pared our list of candidate orbits by using Gauss' method of initial orbit determination to get a rough, computationally inexpensive orbit fit for each triplet. We retained all triplets that fit to orbits with $e < 3$ leaving 978 unique candidates.[7] For each of the remaining triplets, we did a full orbit fit using the method described by Bernstein & Khushalani (2000), and retained all triplets that resulted in bound, prograde orbits. For each such candidate, we calculated the minimum approach distance between NH's nominal trajectory and the best-fit orbit, and the epoch at which the approach will occur. This step yielded 83 candidates that will make an approach within 5 au of NH at some epoch after June 2023. Of these 83 candidates, only 8 triplets that did not contain any synthetic sources survived visual inspection. We show an example of a linkage of a synthetic source in Figure 9.2, and an example of a real candidate in Figure 9.3. Given their current best-fit orbits, none of the 8 candidates will pass close enough to the spacecraft for a flyby (at most 10^5 km). However, all 8 candidate detections are potential targets for imaging by NH's LORRI camera over the next few years. For example, the best-fit orbit to the triplet shown in Figure 9.3 passes 0.98 au from the nominal NH trajectory on October 8,

[5]Some objects must be on unbound orbits in order to have the specified values of ρ and $\dot{\rho}$ at t_0. To simplify our analysis, we throw out such detections.

[6]We used a variety of reference epochs, including June 12 2021, June 27 2021, July 12 2021, July 27 2021, August 11 2021.

[7]We chose the $e < 3$ cutoff empirically using simulations to maximize the rejection of false linkages, while minimizing loss of good linkages.

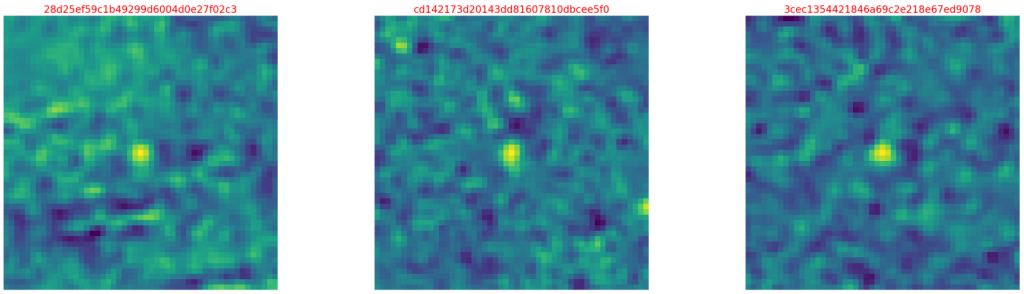


Figure 9.2: Linked recovery of a synthetic source.

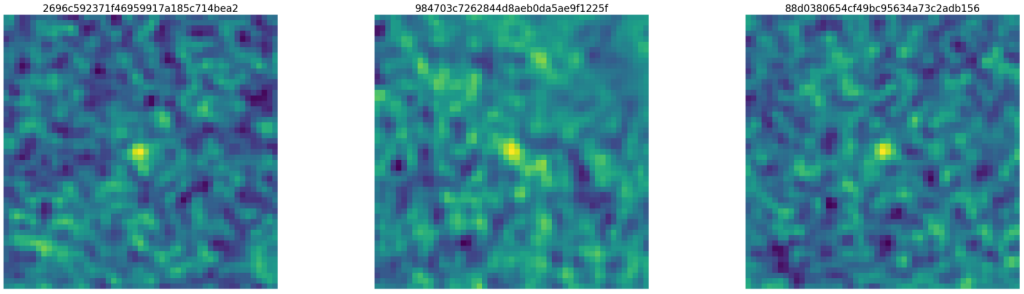


Figure 9.3: Linked detections of a faint candidate. The best-fit orbit for this triplet will pass within approximately 1 au NH on October 8 2025.

2025, making it an excellent candidate to be imaged from the spacecraft. Unfortunately, we were unable to find a convincing fourth detection of any of our 8 candidates, but that is unsurprising given the relatively low quality of our images, and the fact that some of the candidates are likely false positives. More observations are required to confirm any of the candidates as targets.

9.4 Prospects for the Existence of a Flyby Target

We now take a step back to evaluate the probability that an object suitable for a flyby exists in the first place. NH has enough fuel onboard for a boost (essentially an instantaneous change in velocity) of $\Delta v \sim 30$ m/s, and it will be able to communicate with Earth until at least January 2042.[8].

[8]These numbers come from internal discussions with the NH science team

To estimate the probability that an encounter target exists, we used the OSSOS++ solar system model described in Chapter 7. Recall that the model provides a 20-times oversampling of the outer solar system, and strives to provide a realistic representation of the relative population sizes of different dynamical classes. There is some uncertainty in the model, especially at the outer reaches of the solar system that are relevant for the present application, but it is nevertheless our best option.

For each of the 530 million objects in the OSSOS++ model, we used a numerical optimizer to calculate the epoch (t_{approach}) and distance (d_{approach}) of closest approach between the object and the nominal NH trajectory. We recorded any object that came within 5 au of NH after June 2023, and before January 2042. For each such object, one can approximate the required Δv for NH to make an encounter as

$$\Delta v_{\text{encounter}} = \frac{d_{\text{approach}}}{t_{\text{approach}} - t_{\text{boost}}}, \qquad t_{\text{approach}} > t_{\text{boost}}. \tag{9.5}$$

In Figure 9.4 we show $\Delta v_{\text{encounter}}$ vs apparent magnitude (as observed from Earth at the present epoch) for all objects in the model that NH can reach with $\Delta v \leq 100$ m/s for boosts on June 29 2023 (left) and June 1 2030 (right). The green region in each panel represents the parameter space for objects that are both reachable by the spacecraft, and bright enough to be observed by the Hubble Space Telescope (HST). Note that as the boost moves farther into the future, there are fewer encounterable objects. Since there are zero points in the green region, and the OSSOS++ model is oversampled by a factor of 20, the probability of there being an observable flyby candidate is less than 5%. Unless the distant populations in the OSSOS++ model are drastically undersampled, the prognosis for finding a flyby target is grim.

To make matters worse, the objects with $\Delta v_{\text{encounter}} < 100$ m/s are currently spread over more than 70 square degrees of sky. So even if such an object does exist, it could take as many as 10^4 HST fields reaching a depth of 29th magnitude to find it. While ground-based instruments can cover the required sky area to complete the search, none are currently capable of achieving depths much beyond 27th magnitude using state-of-the-art techniques for moving object detection.

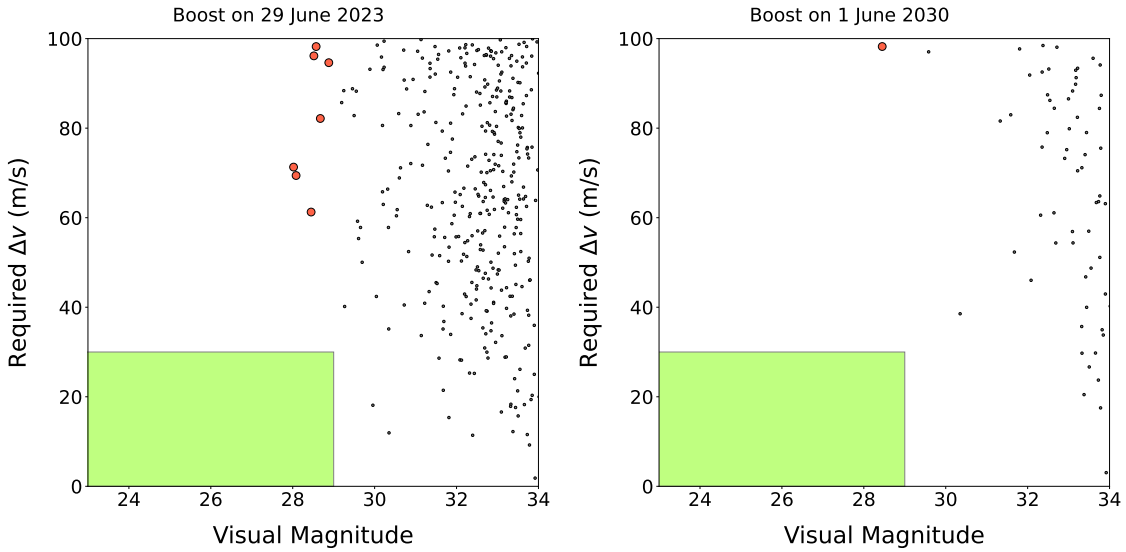


Figure 9.4: $\Delta v_{\text{encounter}}$ vs apparent magnitude for all objects in the model that NH can reach with $\Delta v \leq 100$ m/s. The figure on the left corresponds to a boost on June 29 2023, and the figure on the right corresponds to a boost on June 1 2030. The green region in each panel represents the parameter space for objects that are both reachable by the spacecraft, and bright enough to be observed by HST, assuming an albedo of 0.15. The large red points are brighter than magnitude 29.

9.5 Future Work

While it is extremely unlikely that we will be able to find another flyby target for NH, it is not impossible. NASA's Roman Space Telescope, which is slated to launch in 2027, will be able to achieve depths similar to HST with a 100-times larger field of view. It may be the right tool for the job, but with each passing day the parameter space of encounterable objects is shrinking. To have any chance of finding a flyby target, we must develop techniques that use software to combine images not just from the same night, but taken days, weeks, or even months apart. It will be critical for the NH mission to have such techniques ready by the time Roman launches.

Chapter 10

`HelioStack`: A Novel Approach to the Minor Planet Detection Problem

Out of intense complexities, intense simplicities emerge.

— Winston Churchill

10.1 Introduction

The problem of finding moving solar system bodies in astronomical images has been thoroughly studied for more than two hundred years. The traditional approach is to create a catalog of transient sources, and then use linking techniques such as orbit fitting to associate sources that are consistent with being repeated observations of the same object. While the traditional approach has been incredibly successful (it has yielded the discovery of more than a million solar system bodies), it is limited to finding objects with brightness above the single-image detection threshold.

The standard procedure for detecting moving sources fainter than the single-image detection threshold is a technique called shift-and-stack (or digital tracking), in which one stacks images along the orbit of a moving object. When the signal from the object is added coherently in enough images, the object becomes detectable (see Figure 10.1). The shift-and-stack technique has enabled surveys to achieve depths up to three magnitudes fainter than their single-exposure

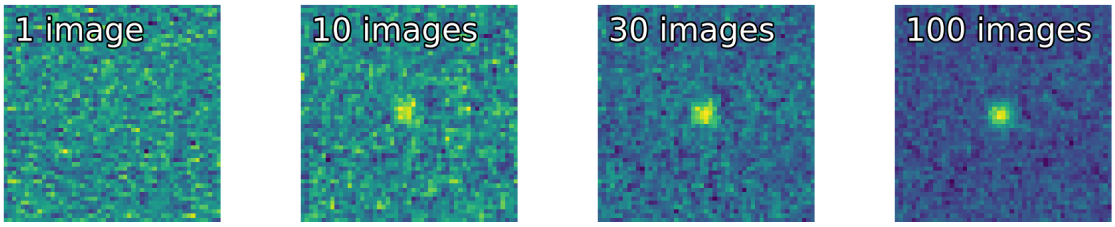


Figure 10.1: Images of a sub-threshold KBO in stacks of 1, 10, 30, and 100 images.

limits (see Chapter 7 and the references therein). To date, all shift-and-stack efforts (except for Bernstein et al. (2004), which searched an extremely small field of 0.019 $\deg^2$) have been limited to stacking images taken in a single night. When the duration of an exposure sequence exceeds one night, the proper motion of even the most distant solar system bodies becomes nonlinear in the topocentric frame, drastically increasing the complexity of the problem. Limitation to a single night of data has relegated the shift-and-stack technique to be useful only for dedicated shift-and-stack surveys.

If we can stack data taken days, months, or even years apart, we will be able to achieve the depth of a shift-and-stack search for surveys with more traditional cadences such as SDSS, DES, or the upcoming LSST. The increased yield of solar system objects by gaining a few magnitudes of sensitivity over such large survey areas would be enormous.

Here we lay the groundwork for an algorithm to enable such searches. The algorithm, called `HelioStack`, is a general solution to the problem of finding moving solar system bodies in any given set of images. `HelioStack` takes inspiration from its eponymous solution to the asteroid linking problem, `HelioLinC` (Holman et al., 2018). The motivation for both is as follows: when a body is observed from a moving reference frame, its apparent motion depends on both the body's own motion and the motion of the observer. If, however, the observer could be stationary with respect to the system's origin, an observed body's apparent motion would be entirely due to its own movement. Thus with full knowledge of a body's orbit, one can simplify the description of the apparent motion of minor planets by correcting for the reflex motion caused by observing

in a moving frame. Although we do not possess such knowledge for the objects in our images *a priori*, we can search the space of all orbits of interest.

We begin in Section 10.2 by introducing a new orbit parameterization that simplifies the description of the sky motion of bodies on Kepler orbits. We continue in Section 10.3 with a description of how to apply our new parameterization to the asteroid linking problem. In Section 10.4 we briefly describe how we intend to extend our formalism to the minor planet detection problem. Finally we conclude in Section 10.5 with an identification of the outstanding questions for the `HelioStack` algorithm, and a discussion of the next steps in its development and implementation.

10.2 A New Orbit Parameterization

Suppose for a moment that we want to take the set of all astronomical images ever recorded, and search it for bodies on all possible orbits. How do we even define the set of all possible orbits that may pass through our images? The problem is complicated, as the images are taken at different times and from different locations. One solution would be to search for all possible orbits at all possible times, but doing so would result in degeneracy in the set of orbits being checked. There are two tricks we can use to avoid such degeneracy. The first trick is to define the orbits at some common reference epoch, t_0. The second is to parameterize orbits in the heliocentric/barycentric frame rather than the topocentric frame. In principle, one can then choose all orbits of interest using any valid parametrization, but we propose the following as it has some particularly nice properties.

For this parameterization we work in heliocentric/barycentric equatorial coordinates. First we define a unit vector that specifies the sky position as

$$\hat{\mathbf{r}} = \langle \cos\delta \cos\alpha, \cos\delta \sin\alpha, \sin\delta \rangle. \tag{10.1}$$

We then define two more unit vectors that are mutually orthogonal to $\hat{\mathbf{r}}$, specifying a complete

right-handed orthogonal basis.

$$\hat{\mathbf{A}} = \langle -\sin\alpha, \cos\alpha, 0 \rangle \tag{10.2}$$

$$\hat{\mathbf{D}} = \langle -\sin\delta\cos\alpha, -\sin\delta\sin\alpha, \cos\delta \rangle. \tag{10.3}$$

Now let the distance and its rate of change be given by r and v_r, respectively. Finally, let the body's speed in the plane defined by $\hat{\mathbf{A}}$ and $\hat{\mathbf{D}}$ be given by v_Ω, and let its direction be given by the angle ψ, which is measured with respect to $\hat{\mathbf{A}}$. The state vector of the test orbit can then be obtained as follows.

$$\mathbf{r} = r\hat{\mathbf{r}} \tag{10.4}$$

$$\mathbf{v} = v_r\hat{\mathbf{r}} + v_\Omega\left(\cos\psi\hat{\mathbf{A}} + \sin\psi\hat{\mathbf{D}}\right) \tag{10.5}$$

Thus we have a complete specification of an orbit using the parameters $\langle \alpha, \delta, r, v_r, v_\Omega, \psi, t_0 \rangle$.

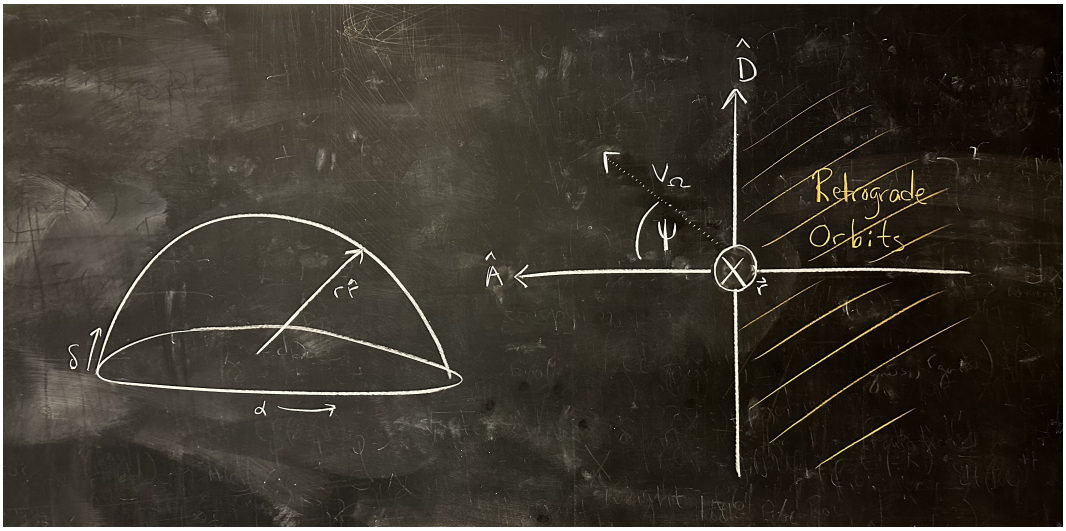


Figure 10.2: Schematic of the `HelioStack` coordinate system. The panel on the left shows the vector describing a body's position in the equatorial frame. The panel on the right shows the coordinate plane defined by the $\hat{\mathbf{A}}$ and $\hat{\mathbf{D}}$ vectors. Here the $\hat{\mathbf{r}}$ vector is going into the page at the coordinate origin. The angle ψ is measured with respect to the $\hat{\mathbf{A}}$ vector, and v_Ω gives the body's speed in the plane. Note that for values of $\psi \in [\pi/2, 3\pi/2]$, all orbits are retrograde with respect to the equatorial plane.

In this coordinate system, the orbit's specific energy and angular momentum are given by

$$\epsilon = \frac{v^2}{2} - \frac{\mu}{r} \qquad h = rv_\Omega \tag{10.6}$$

respectively, where $v^2 = v_r^2 + v_\Omega^2$. We can use these relations to write expressions for the orbit's semi-major axis (a) and eccentricity (e) as follows.

$$a = -\left(\frac{v^2}{\mu} - \frac{2}{r}\right)^{-1} \tag{10.7}$$

$$e = \left[1 + \frac{r^2 v_\Omega^2}{\mu}\left(\frac{v^2}{\mu} - \frac{2}{r}\right)\right]^{1/2} \tag{10.8}$$

Since we have r, a, and e specified at time t_0, we can calculate the orbit's universal anomaly, χ (in the case of an elliptical orbit, this is the eccentric anomaly, E). Thus for a given set of $\langle r, v_r, v_\Omega \rangle$, all points on the sky have identical values of $\langle a, e, \chi \rangle$.

With a page of algebra, one can show that in the above parameterization, an orbit's inclination is a function of only δ and ψ, given by $\cos i = \cos\delta \cos\psi$. It is sometimes more convenient to specify the orbit in terms of $\langle \alpha, \delta, r, v_r, v_\Omega, i, t_0 \rangle$. This parameterization has the advantage that it fixes the orbital inclination independently of δ, which is useful for solving the problem posed in Section 10.3. It comes with the drawback that a given set of $\langle \alpha, r, v_r, v_\Omega, i, t_0 \rangle$ is not necessarily valid for all δ. Luckily, it is trivial to identify invalid orbits using the constraint $i \geq |\delta|$.

10.3 A New Method of Linking Detections

By design, the orbit parameterization described in Section 10.2 is extremely convenient for linking detections of solar system bodies. Typically when linking detections, one has a catalog containing information about each detection's sky position and epoch. Sometimes the catalog will contain information about the object's proper motion, in which case the detections are called streaks or tracklets.

In the `HelioLinC` algorithm, Holman et al. (2018) introduced the idea of asserting the unknown parameters for each detection in a catalog in order to fully specify an orbit. With each detection's orbit fully specified, it is simple to propagate all orbits to a common epoch. Detections that are actually repeated observations of the same object tend to cluster in parameter space at the common epoch as the asserted parameters get close to their true values. This approach

drastically simplifies the algorithmic complexity of the linking problem in comparison to more traditional linking techniques.

The effectiveness of the `HelioLinC` algorithm is dependent on the coordinate system. For example the use of Cartesian coordinates in this algorithm would be a nightmare, as all six of the orbit parameters would be varying as a function of time. On the other hand, our new parameterization using $\langle \alpha, \delta, r, v_r, v_\Omega, i, t_0 \rangle$ is ideally suited to the problem, as it can be re-hashed as $\langle \alpha, \delta, a, e, \chi, i, t_0 \rangle$. Each of the parameters a, e, and i are invariant in the approximation of the two-body problem. Furthermore, since a is invariant, the mean motion n is also invariant. Then, since we know e, we can solve for χ at all times.

Now imagine that we want to search our catalog for orbits with a given set of initial conditions, $\langle r, v_r, v_\Omega, i \rangle$ (or alternatively $\langle a, e, \chi, i \rangle$), at some reference epoch t_0. First, we note that a, e, and i are almost always invariant to good approximation, and can be assigned to each detection in the catalog.[1] Then one can calculate χ at the epoch of each detection by solving Kepler's equation. Finally since each detection's α and δ are measured in the topocentric frame at the epoch of observation, it is necessary to transform the detections into the heliocentric/barycentric frame. Since we know the position of the observer at the epoch of the observation, we have all of the information required to make the transformation. At this point, all detections in the catalog have a fully defined orbit parameterized by $\langle \alpha, \delta, a, e, \chi, i, t \rangle$.[2] Then it is simple to analytically propagate all orbits to the reference epoch and search for clusters in $\langle \alpha, \delta \rangle$. One can repeat this process for different sets of initial conditions until they have exhaustively checked the parameter space of interest.

10.4 A New Method for Detecting Minor Planets

The linking procedure we outlined in Section 10.3 can be easily generalized as a solution to the problem of minor planet detection. We will not go into great detail here, but the basic idea is that rather than registering a catalog of detections to a common epoch, one can register all

[1] This approach is invalid for objects undergoing rapid orbital evolution.

[2] Note that once the orbit has been fully specified, one should correct the orbit for light travel time.

pixels in a set of images to a common epoch. Let us imagine for a moment that we have some arbitrary set of images, and we want to search them for some arbitrary set of orbits.

We begin by defining a blank image of the sky at a reference epoch, t_0. We need to define some sort of a grid on the reference image, but since it is a sphere (and is therefore non-euclidean) it cannot be tiled uniformly using square pixels. Instead we generate a roughly uniform covering of points on the sphere to define a mosaic of sectors. In principle the sectors can be any shape, but triangles are a natural choice and allow for reasonably good control over the sector size. In general, the size of a sector is a tunable parameter, but it should be roughly commensurate with the PSF width of the images being considered. Note that by partitioning the reference image into a mosaic rather than registering each image onto a common grid, the algorithm becomes agnostic of the instrument used to take the images.

With the reference mosaic defined, we can move on to the process of accumulating signal. First we choose an initial condition given by $\langle r, v_r, v_\Omega, i \rangle$. Then we register every pixel back to the reference epoch using the technique described in Section 10.3, and accumulate signal in each sector.[3] Finally we identify sectors with high signal-to-noise ratios as candidate moving objects. We can then repeat the process until we have exhaustively searched the parameter space of interest.

10.5 Future Work

While the orbit parameterization that we have introduced in this chapter has a sound theoretical footing, we have not yet explored details of its implementation for the problems of minor planet linking and detection. There are a number of open questions such as

— What is the best way to sample the initial condition space of $\langle r, v_r, v_\Omega, i \rangle$ to exhaustively search for all orbits of interest without any redundant calculation?

[3] Alternatively, if an orbit's a, e, or i may evolve significantly over the time span of the images, it is preferable to take a more complicated approach using N-body integration. In such a case, one can define an orbit for each sector in the reference mosaic at t_0 using $\langle \alpha, \delta, r, v_r, v_\Omega, i \rangle$, and then use N-body integration to propagate each vertex to each epoch for which there is data, accumulating the signal that lands inside the sector. This approach is much more computationally intensive, and is probably unnecessary in most cases.

— How can we minimize the number of false positives, both for the linking problem and the detection problem?

— What kind of computational resources are required to implement the detection algorithm at the scale of a large survey?

The full development of the ideas introduced in this chapter represents a significant amount of work, and will be my primary focus over the next two years as a Schmidt AI in Science Postdoctoral Fellow at the University of Michigan. In particular, I will need to leverage state-of-the-art techniques in computer vision (and probably other AI techniques) to avoid being inundated by false positive detections. While the convolutional neural nets discussed in Chapter 7 worked extremely well for DEEP, they were relatively rudimentary, and were engineered to be *good enough* to enable the DEEP analysis. The AI applications to the `HelioStack` algorithm will have to be much more carefully crafted and tested. Through my fellowship I will be able to access the resources in large-scale computing and computer vision required to make `HelioStack` work.

When we have worked out all of the implementation details, `HelioStack` will be the first algorithm to enable minor planet searches over arbitrarily long time baselines. We intend to test the algorithm on archival data from surveys such as DEEP and DES, searching for particular classes of solar system bodies such as Jupiter Trojans, Neptune Trojans, or extremely distant minor planets. We expect that `HelioStack` will be an integral tool for sub-threshold minor planet detection in future astronomical surveys such as LSST. It will facilitate more flexibility in the observing cadence for solar system surveys, and will enable general astronomical surveys to achieve minor planet search depths that are currently only achievable by dedicated shift-and-stack surveys.

Chapter 11

Conclusion

It's almost over; it's just begun.

— Bo Burnham, *All Eyes on Me*

I began this thesis by identifying a handful of key technological innovations that transformed the way we study our solar system. First the telescope enabled the detection of objects fainter than could be seen by the unaided eye. Combining telescopes with photography enabled astronomers to find even fainter objects, *and* to directly record their observations for later study. Then astronomers developed the modern astronomical survey, using telescopes and photographic plates to systematically scan the sky for moving bodies. Finally as we entered the digital age, the CCD and the computer enabled much more ambitious surveys, yielding solar system detections at an incredible rate.

We are now in a transition to a new scientific paradigm defined by big data, high-performance computing clusters, and AI. Incredible advances in computation have enabled the study of problems that were out of reach only a few years ago. If we can take full advantage of these technological innovations, we have the opportunity to push into a new regime in Figure 1.1, and in turn revolutionize our understanding of the solar system. In this thesis, I developed techniques to do just that. The text was split into two distinct parts, roughly corresponding to *dynamics* and *discovery*. In Part I, I investigated problems in solar system dynamics using a semi-analytic approach in which I combined pure analytic theory with highly detailed simulations. In Part II, I

developed new techniques for discovering faint solar system bodies.

To prepare for solving the problems lying at the intersection of observation, theory, and computation, I began my PhD research with the development of a software package for solar system dynamics, called `spacerocks` (see Chapter 2). `spacerocks` proved to be an invaluable tool throughout my thesis, making it trivial to do everything from quickly testing new ideas, to setting up full-scale numerical experiments. I used it for *every* chapter in this thesis, and I will continue to develop and maintain the software so it can serve as a tool for the community.

Dynamical Studies

For the first portion of my PhD, I used a combination of analytic and numerical techniques to work on the dynamical problems comprising Part I of this thesis.

First in Chapter 3 I presented the work published in Napier et al. (2021c). We examined the claim that the so-called ETNOs are having their orbits aligned by the gravitational influence of the hypothesized Planet Nine. By carefully simulating the three most productive Kuiper Belt surveys of the past decade, we were able to show that when the surveys' observational biases were fully accounted for, their ETNO detections were consistent with being drawn from an isotropic population. In other words, it is possible that the apparent clustering is an artifact of survey bias—the ETNOs only appear to be clustered because surveys have only been sensitive to ETNOs in the apparent cluster. Of course we cannot rule out the possibility that the ETNOs truly are clustered, and our surveys just happen to only be sensitive to the parameter space of the cluster. However, the hypothesis of an isotropic underlying distribution offers a simpler explanation.

Next in Chapters 4 and 5, I presented the work published in Napier et al. (2021a) and Napier et al. (2021b). Inspired by the recent detections of the first two interstellar objects 'Oumuamua (Meech et al., 2017) and Borisov (Jewitt & Luu, 2019), we investigated the capture of interstellar objects by our solar system. Using a suite of 500 million numerical simulations, we calculated the capture cross section for interstellar objects as a function of hyperbolic excess velocity, as well as the characteristic dynamical lifetime of captured objects. The numerical results

for the capture cross section closely matched our analytic theory, implying that our theoretical formulation captures all of the essential physics of the capture process. We estimate that the steady-state mass of captured interstellar objects in our solar system is less than 10^{19} grams—1000 times less massive than Earth's atmosphere.

Finally in Chapter 6, I presented the work published in Napier et al. (2022). Inspired by the continued lack of any known dynamically stable Earth Trojans, we proposed a collisional mechanism by which such objects may have been destabilized. We used a suite of numerical simulations to show that an initially-stable population of primordial Earth Trojans would have been severely disrupted (almost certainly decimated) by the asteroid impacts that gave Earth its so-called Late Veneer. Our results provide one satisfactory explanation for the lack of observed Earth Trojans, though it is possible that other solutions exist. Although we specifically investigated the process for Earth in this study, the process is applicable for any Trojan-hosting planet.

Novel Methods of Minor Planet Detection

For the second portion of my PhD, I developed software to discover faint minor planets in large datasets of astronomical images.

In Chapter 7 I presented a search for Kuiper Belt objects in 20 pointings from DEEP. By using a shift-and-stack technique to extract sub-threshold sources, and convolutional neural networks to reject false positives, we were able to reach an average depth of $m_r \sim 26.2$ with a relatively small amount of human vetting. We then used the results of the human vetting process to calculate the probability that each of our true detections was a true positive, enabling proper quantification of a false positive rate. Altogether, we ended up with ~ 2300 single-epoch detections. The main scientific result of this work was a measurement of the absolute magnitude distribution of the Cold Classical Kuiper Belt. Interestingly, our data is well-fit by an exponentially-tapered power law, which is the functional form predicted by streaming instability simulations (Abod et al., 2019; Kavelaars et al., 2021). The work in Chapter 7 utilized only one quarter of the images from the DEEP survey. Future analyses of the full dataset will reveal large populations of many more of

the Kuiper Belt's dynamical structures, enabling detailed theoretical studies of the solar system's formation. By its conclusion, DEEP will have doubled the number of known Kuiper Belt objects.

In Chapter 8 I presented a proof-of-concept shift-and-stack search for Jupiter Trojans in one of the DEEP fields. While the search required us to investigate a parameter space two orders of magnitude larger than the parameter space for KBOs, the algorithms in Chapter 7 transferred incredibly well. At a glance, it appears that we detected well over 100 new Jupiter Trojans (not to mention hundreds of new Hildas and asteroids), with an estimated false positive rate under 10% prior to any human inspection. Given the success of this project, we have begun the process of planning a dedicated search for a new target for NASA's Lucy spacecraft.

In Chapter 9 I presented a search for a flyby target for NASA's NH spacecraft. We performed shift-and-stack searches of four DECam fields observed over a 2-month period, extracting sources down to 3σ. We used a novel combination of convolutional neural networks and orbit linking to reject false positive detections, and found that the technique worked extremely well. However, despite finding some candidate objects that may be observable at a distance from the spacecraft, we did not find any objects suitable for a flyby. The absence of a flyby target is unsurprising, given our estimate that a survey covering 70 square degrees of sky in the vicinity of NH to 29th magnitude, would have less than a 5% probability of finding a flyby target.

Finally in Chapter 10 I introduced a novel approach to the minor planet detection problem, called `HelioStack`. We showed that by shifting perspective from the topocentric frame to the heliocentric/barycentric frame, and then choosing an appropriate orbit parameterization, one can drastically simplify the description of the apparent motion of bodies on Keplerian orbits. This project is still in its early stages, but we anticipate that it will be critical for enabling sub-threshold minor planet detection in future astronomical surveys such as LSST (Ivezić et al., 2019).

Closing Remarks

The work in this thesis has demonstrated how recent innovations in computing have enabled the study of our solar system in unprecedented detail, both on observational and theoretical fronts.

Despite these successes we have still only scratched the surface of what is possible. There are a number of outstanding questions about the nature of our solar system. For example,

— Are there minor body populations that have not yet been discovered?

— Do the outer reaches of our solar system harbor any yet-undetected planets?

— What are the details of planetesimal formation?

— What major dynamical events occurred during our solar system's formation and evolution?

Each of these questions can only be resolved, either directly or indirectly, by discovering and studying more solar system bodies. With each new minor planet discovery, especially those from undersampled populations, we inch closer to being able to piece together the history of our solar system's formation and evolution.

With ongoing advancements in theory, AI, and astronomical surveys (such as LSST) we are on the precipice of a new revolution in the discovery of solar system bodies. By using software to combine images taken not just over the span of a few hours, but rather over observing campaigns spanning days, weeks, or even months, we will be able to achieve sensitivity to faint solar system bodies in general-purpose astronomical surveys that is currently only possible in shift-and-stack surveys. In order to use the immense datasets generated by modern astronomical surveys to their full potential, it will be necessary to enlist the help of AI; there is simply too much data to be handled by humans or traditional computing approaches. The success of long time-baseline shift-and-stack will mark a new era in the discovery of solar system bodies.

Bibliography

Abadi, M., et al. 2015, TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems. https://www.tensorflow.org/

Abod, C. P., Simon, J. B., Li, R., et al. 2019, ApJ, 883, 192, doi: 10.3847/1538-4357/ab40a3

Abramowitz, M., & Stegun, I. A. 1972, Handbook of Mathematical Functions (New York: Dover)

Adams, F. C. 2010, ARA&A, 48, 47, doi: 10.1146/annurev-astro-081309-130830

Adams, F. C., & Laughlin, G. 2001, Icarus, 150, 151, doi: 10.1006/icar.2000.6567

Adams, F. C., Proszkow, E. M., Fatuzzo, M., & Myers, P. C. 2006, ApJ, 641, 504, doi: 10.1086/500393

Adams, F. C., & Spergel, D. N. 2005, Astrobiology, 5, 497, doi: 10.1089/ast.2005.5.497

Adams, J. C. 1846, MNRAS, 7, 149, doi: 10.1093/mnras/7.9.149

Alard, C., & Lupton, R. H. 1998, ApJ, 503, 325, doi: 10.1086/305984

Allen, R. L., Bernstein, G. M., & Malhotra, R. 2001, The Astrophysical Journal Letters, 549

Andrews, S. M., Huang, J., Pérez, L. M., et al. 2018, ApJL, 869, L41, doi: 10.3847/2041-8213/aaf741

Annex, A., Pearson, B., Seignovert, B., et al. 2020, The Journal of Open Source Software, 5, 2050, doi: 10.21105/joss.02050

Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33, doi: 10.1051/0004-6361/201322068

Bailey, B. L., & Malhotra, R. 2009, Icarus, 203, 155, doi: 10.1016/j.icarus.2009.03.044

Bannister, M. T., Kavelaars, J. J., Petit, J.-M., et al. 2016, AJ, 152, 70, doi: 10.3847/0004-6256/152/3/70

Bannister, M. T., Gladman, B. J., Kavelaars, J. J., et al. 2018, ApJS, 236, 18, doi: 10.3847/1538-4365/aab77a

Barbary, K. 2016, The Journal of Open Source Software, 1, doi: 10.21105/joss.00058

Bate, R. R., Mueller, D. D., & White, J. E. 1971, Fundamentals of astrodynamics. (New York: Dover)

Batygin, K., Adams, F. C., Batygin, Y. K., & Petigura, E. A. 2020, AJ, 159, 101, doi: 10.3847/1538-3881/ab665d

Batygin, K., Adams, F. C., Brown, M. E., & Becker, J. C. 2019, PhR, 805, 1, doi: 10.1016/j.physrep.2019.01.009

Batygin, K., & Brown, M. E. 2016, AJ, 151, 22, doi: 10.3847/0004-6256/151/2/22

—. 2021, ApJL, 910, L20, doi: 10.3847/2041-8213/abee1f

Becker, A. 2015, HOTPANTS: High Order Transform of PSF ANd Template Subtraction. http://ascl.net/1504.004

Becker, J. C., Khain, T., Hamilton, S. J., et al. 2018, The Astronomical Journal, 156, 81, doi: 10.3847/1538-3881/aad042

Becker, J. C., Khain, T., Hamilton, S. J., et al. 2018, AJ, 156, 81, doi: 10.3847/1538-3881/aad042

Bernardinelli, P. H., Bernstein, G. M., Sako, M., et al. 2020a, The Planetary Science Journal, 1, 28, doi: 10.3847/PSJ/ab9d80

—. 2020b, ApJS, 247, 32, doi: 10.3847/1538-4365/ab6bd8

—. 2022, ApJS, 258, 41, doi: 10.3847/1538-4365/ac3914

Bernstein, G., & Khushalani, B. 2000, AJ, 120, 3323, doi: 10.1086/316868

Bernstein, G. M., Trilling, D. E., Allen, R. L., et al. 2004, AJ, 128, 1364, doi: 10.1086/422919

Bertin, E., & Arnouts, S. 1996, Astronomy and Astrophysics, Supplement, 117, doi: 10.1051/aas:1996164

Bertin, E., et al. 2002, in Astronomical Society of the Pacific Conference Series, Vol. 281, Astronomical Data Analysis Software and Systems XI, ed. D. A. Bohlender, D. Durand, & T. H. Handley, 228

Binney, J., & Tremaine, S. 2008, Galactic Dynamics: Second Edition (Princeton, NJ: Princeton Univ. Press)

Borisov, G. 2019, MPEC 2019-R106 (September 11)

Bottke, William F., J., Vokrouhlický, D., Rubincam, D. P., & Nesvorný, D. 2006, Annual Review of Earth and Planetary Sciences, 34, 157, doi: 10.1146/annurev.earth.34.031405.125154

Bottke, W., Vokrouhlicky, D., Marshall, R., et al. 2023, arXiv e-prints, arXiv:2307.07089, doi: 10.48550/arXiv.2307.07089

Bottke, W. F., Jedicke, R., Morbidelli, A., Petit, J.-M., & Gladman, B. 2000, Science, 288, 2190, doi: 10.1126/science.288.5474.2190

Brasser, R., Duncan, M. J., & Levison, H. F. 2006, Icarus, 184, 59, doi: 10.1016/j.icarus.2006.04.010

—. 2007, Icarus, 191, 413, doi: 10.1016/j.icarus.2007.05.003

Brasser, R., Mojzsis, S. J., Werner, S. C., Matsumura, S., & Ida, S. 2016, Earth and Planetary Science Letters, 455, 85, doi: 10.1016/j.epsl.2016.09.013

Brown, M. E. 2001, The Astronomical Journal, 121, 2804. http://stacks.iop.org/1538-3881/121/i=5/a=2804

—. 2017, The Astronomical Journal, 154, doi: 10.3847/1538-3881/aa79f4

Brown, M. E., & Batygin, K. 2019, The Astronomical Journal, 157, 62, doi: 10.3847/1538-3881/aaf051

Brown, M. E., Trujillo, C., & Rabinowitz, D. 2004, ApJ, 617, 645, doi: 10.1086/422095

Buie, M. W., et al. 2020, The Astronomical Journal, 159, doi: 10.3847/1538-3881/ab6ced

Cambioni, S., Malhotra, R., Hergenrother, C. W., et al. 2018, in Lunar and Planetary Science Conference, Lunar and Planetary Science Conference, 1149

Campbell, W. W. 1892, PASP, 4, 264, doi: 10.1086/120521

Canup, R. M. 2012, Science, 338, 1052, doi: 10.1126/science.1226073

Christou, A. A., & Georgakarakos, N. 2021, MNRAS, 507, 1640, doi: 10.1093/mnras/stab2223

Connors, M., Wiegert, P., & Veillet, C. 2011, Nature, 475, 481, doi: 10.1038/nature10233

Copernicus, N. 1543, De revolutionibus orbium coelestium (Nuremberg: Johannes Petreius)

Crompvoets, B. L., Lawler, S. M., Volk, K., et al. 2022, PSJ, 3, 113, doi: 10.3847/PSJ/ac67e0

Ćuk, M., Hamilton, D. P., & Holman, M. J. 2012, MNRAS, 426, 3051, doi: 10.1111/j.1365-2966.2012.21964.x

Ćuk, M., & Stewart, S. T. 2012, Science, 338, 1047, doi: 10.1126/science.1225542

Dark Energy Survey Collaboration, Abbott, T., Abdalla, F. B., et al. 2016, MNRAS, 460, 1270, doi: 10.1093/mnras/stw641

DES Collaboration. 2005, ArXiv e-prints. https://arxiv.org/abs/astro-ph/0510346

Dones, L., Levison, H. F., & Duncan, M. 1996, in Astronomical Society of the Pacific Conference Series, Vol. 107, Completing the Inventory of the Solar System, ed. T. Rettig & J. M. Hahn, 233–244

Duncan, M., Quinn, T., & Tremaine, S. 1987, AJ, 94, 1330, doi: 10.1086/114571

—. 1988, ApJL, 328, L69, doi: 10.1086/185162

Edgeworth, K. E. 1949, Monthly Notices of the Royal Astronomical Society, 109, 600, doi: 10.1093/mnras/109.5.600

Efron, B., & Tibshirani, R. 1986, Statistical Science, 1, 54 , doi: 10.1214/ss/1177013815

Fernandez, J. A., & Ip, W. H. 1984, Icarus, 58, 109, doi: 10.1016/0019-1035(84)90101-5

Flaugher, B., Diehl, H. T., Honscheid, K., et al. 2015, AJ, 150, 150, doi: 10.1088/0004-6256/150/5/150

Fraser, W. C., Brown, M. E., Morbidelli, A., Parker, A., & Batygin, K. 2014, ApJ, 782, 100, doi: 10.1088/0004-637X/782/2/100

Fraser, W. C., & Kavelaars, J. J. 2009, AJ, 137, 72, doi: 10.1088/0004-6256/137/1/72

Fraser, W. C., Kavelaars, J. J., Holman, M. J., et al. 2008, Icarus, 195, 827, doi: 10.1016/j.icarus.2008.01.014

Fuentes, C. I., George, M. R., & Holman, M. J. 2009, ApJ, 696, 91, doi: 10.1088/0004-637X/696/1/91

Galle, J. G. 1846, Monthly Notices of the Royal Astronomical Society, 7, 153, doi: 10.1093/mnras/7.9.153

Genda, H., Brasser, R., & Mojzsis, S. J. 2017, Earth and Planetary Science Letters, 480, 25, doi: 10.1016/j.epsl.2017.09.041

Gladman, B., & Kavelaars, J. J. 1997, Astronomy and Astrophysics, 317

Gladman, B., Marsden, B. G., & Vanlaerhoven, C. 2008, in The Solar System Beyond Neptune, ed. M. A. Barucci, H. Boehnhardt, D. P. Cruikshank, A. Morbidelli, & R. Dotson (The University of Arizona Press), 43

Gladman, B., & Volk, K. 2021, ARA&A, 59, 203, doi: 10.1146/annurev-astro-120920-010005

Goldstein, D. A., et al. 2015, The Astronomical Journal, 150, doi: 10.1088/0004-6256/150/3/82

Grundy, W. M., et al. 2020, Science, 367, doi: 10.1126/science.aay3705

Hamilton, S. 2019, PhD thesis, University of Michigan

Hands, T. O., & Dehnen, W. 2020, MNRAS, 493, L59, doi: 10.1093/mnrasl/slz186

Heggie, D. C. 1975, MNRAS, 173, 729, doi: 10.1093/mnras/173.3.729

Heinze, A. N., Metchev, S., & Trollo, J. 2015, The Astronomical Journal, 150

Herschel, W., & Watson, D. 1781, Philosophical Transactions of the Royal Society of London Series I, 71, 492

Holman, M. J., Akmal, A., Farnocchia, D., et al. 2023, PSJ, 4, 69, doi: 10.3847/PSJ/acc9a9

Holman, M. J., Payne, M. J., Blankley, P., Janssen, R., & Kuindersma, S. 2018, AJ, 156, 135, doi: 10.3847/1538-3881/aad69a

Holman, M. J., et al. 2004, Nature, 430

Hui, M.-T., Wiegert, P. A., Tholen, D. J., & Föhring, D. 2021, ApJL, 922, L25, doi: 10.3847/2041-8213/ac37bf

Ito, T., & Ohtsuka, K. 2019, Monographs on Environment, Earth and Planets, 7, 1, doi: 10.5047/meep.2019.00701.0001

Ivezić, Ž., Kahn, S. M., Tyson, J. A., et al. 2019, ApJ, 873, 111, doi: 10.3847/1538-4357/ab042c

Jeffers, S. V., Manley, S. P., Bailey, M. E., & Asher, D. J. 2001, MNRAS, 327, 126, doi: 10.1046/j.1365-8711.2001.04747.x

Jewitt, D., & Luu, J. 1993, Nature, 362, 730, doi: 10.1038/362730a0

—. 2019, ApJL, 886, L29, doi: 10.3847/2041-8213/ab530b

Kavelaars, J. J., et al. 2021, Astrophysical Journal Letters, 920, doi: 10.3847/2041-8213/ac2c72

Kepler, J. 1609, Astronomia nova, doi: 10.3931/e-rara-558

—. 1619, Harmonices mundi Libri V, doi: 10.3931/e-rara-8723

Kessler, R., Marriner, J., Childress, M., et al. 2015, AJ, 150, 172, doi: 10.1088/0004-6256/150/6/172

Khain, T., et al. 2020a, The Astronomical Journal, 159, doi: 10.3847/1538-3881/ab7002

Khain, T., Becker, J. C., Lin, H. W., et al. 2020b, AJ, 159, 133, doi: 10.3847/1538-3881/ab7002

Kleine, T. 2011, Nature, 477, 168, doi: 10.1038/477168a

Kouwenhoven, M. B. N., Goodwin, S. P., Parker, R. J., et al. 2010, MNRAS, 404, 1835, doi: 10.1111/j.1365-2966.2010.16399.x

Kozai, Y. 1962, AJ, 67, 591, doi: 10.1086/108790

Lada, C. J., & Lada, E. A. 2003, ARA&A, 41, 57, doi: 10.1146/annurev.astro.41.011802.094844

Lagrange, J.-L. 1776, Mémoires de l' Académie de Berlin, 199 Oeuves complétes VI 255 (Paris, Gauthier-Villars 1869)

Laplace, P. S. 1796, Exposition du système du monde (Cercle-Social, Paris), doi: 10.3931/e-rara-497

—. 1799-1825, Traité de Mécanie Céleste (Paris: Gauthier-Villa)

—. 1846, in Sur les Comètes, Connaissance des temps

Laskar, J. 1989, Nature, 338, 237, doi: 10.1038/338237a0

Lawler, S. M., Kavelaars, J., Alexandersen, M., et al. 2018, ArXiv e-prints. https://arxiv.org/abs/1802.00460

Lay, O. P., Carlstrom, J. E., Hills, R. E., & Phillips, T. G. 1994, ApJL, 434, L75, doi: 10.1086/187578

Le Verrier, U. J. 1846, Astronomische Nachrichten, 25, 65

Lehmann, B. V., Ross, O., Webber, A., & Profumo, S. 2021, MNRAS, doi: 10.1093/mnras/stab1121

Levison, H. F., & Duncan, M. J. 1994, Icarus, 108, 18, doi: 10.1006/icar.1994.1039

Levison, H. F., Olkin, C. B., Noll, K. S., et al. 2021, The Planetary Science Journal, 2, 171, doi: 10.3847/PSJ/abf840

Li, G., & Adams, F. C. 2015, MNRAS, 448, 344, doi: 10.1093/mnras/stv012

—. 2016, ApJL, 823, L3, doi: 10.3847/2041-8205/823/1/L3

Lidov, M. L. 1962, P&SS, 9, 719, doi: 10.1016/0032-0633(62)90129-0

Lifset, N., Golovich, N., Green, E., Armstrong, R., & Yeager, T. 2021, AJ, 161, 282, doi: 10.3847/1538-3881/abf7af

Lingam, M., & Loeb, A. 2018, AJ, 156, 193, doi: 10.3847/1538-3881/aae09a

Loredo, T. J. 2004, in American Institute of Physics Conference Series, Vol. 735, Bayesian Inference and Maximum Entropy Methods in Science and Engineering: 24th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering, ed. R. Fischer, R. Preuss, & U. V. Toussaint, 195–206, doi: 10.1063/1.1835214

Luu, J. X., & Jewitt, D. C. 2002, ARA&A, 40, 63, doi: 10.1146/annurev.astro.40.060401.093818

Magnier, E. A., Schlafly, E., Finkbeiner, D., et al. 2013, ApJS, 205, 20, doi: 10.1088/0067-0049/205/2/20

Malhotra, R. 1993, Nature, 365, 819, doi: 10.1038/365819a0

—. 1995, AJ, 110, 420, doi: 10.1086/117532

—. 2019, Nature Astronomy, 3, 193, doi: 10.1038/s41550-019-0697-z

Malyshkin, L., & Tremaine, S. 1999, Icarus, 141, 341, doi: 10.1006/icar.1999.6174

Manski, C., & Lerman, S. R. 1977, Econometrica, 45, 1977. https://EconPapers.repec.org/RePEc:ecm:emetrp:v:45:y:1977:i:8:p:1977-88

Markwardt, L., Adams, F., Gerdes, D., et al. 2021, The First Near-IR Spectroscopic Survey of Neptune Trojans, JWST Proposal. Cycle 1

Markwardt, L., Gerdes, D. W., Malhotra, R., et al. 2020, MNRAS, 492, 6105, doi: 10.1093/mnras/staa232

Marzari, F., & Scholl, H. 2013, Celestial Mechanics and Dynamical Astronomy, 117, 91, doi: 10.1007/s10569-013-9478-7

McKinnon, W. B., et al. 2020, Science, 367, doi: 10.1126/science.aay6620

Meech, K. J., Weryk, R., Micheli, M., et al. 2017, Nature, 552, 378, doi: 10.1038/nature25020

Melosh, H. J. 2003, Astrobiology, 3, 207, doi: 10.1089/153110703321632525

Millis, R. L., Buie, M. W., Wasserman, L. H., et al. 2002, AJ, 123, 2083, doi: 10.1086/339481

Moore, N. W. H., Li, G., & Adams, F. C. 2020, ApJ, 901, 92, doi: 10.3847/1538-4357/abb08f

Moorhead, A. V., & Adams, F. C. 2005, Icarus, 178, 517, doi: 10.1016/j.icarus.2005.05.005

Morbidelli, A., Baillié, K., Batygin, K., et al. 2022, Nature Astronomy, 6, 72, doi: 10.1038/s41550-021-01517-7

Morbidelli, A., Batygin, K., Brasser, R., & Raymond, S. N. 2020, MNRAS, 497, L46, doi: 10.1093/mnrasl/slaa111

Morbidelli, A., Levison, H. F., Tsiganis, K., & Gomes, R. 2005, Nature, 435, 462, doi: 10.1038/nature03540

Murray, C. A. 1989, A&A, 218, 325

Murray, C. D., & Dermott, S. F. 2000, Solar System Dynamics (Cambridge: Cambridge Univ. Press)

Mustill, A. J., Raymond, S. N., & Davies, M. B. 2016, MNRAS, 460, L109, doi: 10.1093/mnrasl/slw075

Namouni, F., & Morais, M. H. M. 2020, MNRAS, 494, 2191, doi: 10.1093/mnras/staa712

Napier, K. 2020, SpaceRocks, https://github.com/kjnapier/spacerocks

Napier, K. J., Adams, F. C., & Batygin, K. 2021a, PSJ, 2, 53, doi: 10.3847/PSJ/abe76e/53

—. 2021b, PSJ, 2, 217, doi: 10.3847/PSJ/ac29bb

Napier, K. J., Markwardt, L., Adams, F. C., Gerdes, D. W., & Wen Lin, H. 2022, PSJ, 3, 121, doi: 10.3847/PSJ/ac6958

Napier, K. J., Gerdes, D. W., Lin, H. W., et al. 2021c, PSJ, 2, 59, doi: 10.3847/PSJ/abe53e

Nesvorný, D. 2015, The Astronomical Journal, 150, doi: 10.1088/0004-6256/150/3/73

—. 2018, ARA&A, 56, 137, doi: 10.1146/annurev-astro-081817-052028

Nesvorný, D. 2020, Research Notes of the AAS, 4, doi: 10.3847/2515-5172/abceb0

Newton, I. 1687, Philosophiae naturalis principia mathematica (J. Societatis Regiae ac Typis J. Streater), doi: 10.5479/sil.52126.39088015628399

Park, R. S., Folkner, W. M., Williams, J. G., & Boggs, D. H. 2021, AJ, 161, 105, doi: 10.3847/1538-3881/abd414

Parker, A. H., & Kavelaars, J. J. 2010, Publications of the Astronomical Society of the Pacific, 122, doi: 10.1086/652424

Parker, R. J. 2020, Royal Society Open Science, 7, 201271, doi: 10.1098/rsos.201271

Petit, J. M., Holman, M. J., Gladman, B. J., et al. 2006, MNRAS, 365, 429, doi: 10.1111/j.1365-2966.2005.09661.x

Petit, J.-M., Kavelaars, J. J., Gladman, B. J., et al. 2011, AJ, 142, 131, doi: 10.1088/0004-6256/142/4/131

Pfalzner, S. 2013, A&A, 549, A82, doi: 10.1051/0004-6361/201218792

Pineault, S., & Duquet, J.-R. 1993, MNRAS, 261, 246, doi: 10.1093/mnras/261.2.246

Polishook, D., Jacobson, S. A., Morbidelli, A., & Aharonson, O. 2017, Nature Astronomy, 1, 0179, doi: 10.1038/s41550-017-0179

Portegies Zwart, S. F. 2009, ApJL, 696, L13, doi: 10.1088/0004-637X/696/1/L13

Porter, S. B., Spencer, J. R., Benecchi, S., et al. 2016, ApJL, 828, L15, doi: 10.3847/2041-8205/828/2/L15

Raymond, S. N., Schlichting, H. E., Hersant, F., & Selsis, F. 2013, Icarus, 226, 671, doi: 10.1016/j.icarus.2013.06.019

Read, J. I. 2014, Journal of Physics G Nuclear Physics, 41, 063101, doi: 10.1088/0954-3899/41/6/063101

Rein, H., & Spiegel, D. S. 2015, MNRAS, 446, 1424, doi: 10.1093/mnras/stu2164

Rein, H., & Tamayo, D. 2015, MNRAS, 452, 376, doi: 10.1093/mnras/stv1257

Rice, M., & Laughlin, G. 2019, ApJL, 884, L22, doi: 10.3847/2041-8213/ab4422

Rombach, R., Blattmann, A., Lorenz, D., Esser, P., & Ommer, B. 2021, High-Resolution Image Synthesis with Latent Diffusion Models. https://arxiv.org/abs/2112.10752

Safronov, V. S. 1972, Evolution of the Protoplanetary Cloud and Formation of the Earth and the Planets (Israel Program for Scientific Translations)

Salucci, P., Nesti, F., Gentile, G., & Frigerio Martins, C. 2010, A&A, 523, A83, doi: 10.1051/0004-6361/201014385

Scholtz, J., & Unwin, J. 2020, Phys. Rev. Lett., 125, 051103, doi: 10.1103/PhysRevLett.125.051103

Schwamb, M. E., Jones, R. L., Chesley, S. R., et al. 2018, arXiv e-prints, arXiv:1802.01783. https://arxiv.org/abs/1802.01783

Shankman, C., Kavelaars, J. J., Bannister, M. T., et al. 2017, AJ, 154, 50, doi: 10.3847/1538-3881/aa7aed

Sheppard, S. S., & Trujillo, C. 2016, AJ, 152, 221, doi: 10.3847/1538-3881/152/6/221

Sheppard, S. S., Trujillo, C., & Tholen, D. J. 2016, ApJL, 825, L13, doi: 10.3847/2041-8205/825/1/L13

Sheppard, S. S., Trujillo, C. A., Tholen, D. J., & Kaib, N. 2019, The Astronomical Journal, 157, 139, doi: 10.3847/1538-3881/ab0895

Shu, F. H., Adams, F. C., & Lizano, S. 1987, ARA&A, 25, 23, doi: 10.1146/annurev.aa.25.090187.000323

Siraj, A., & Loeb, A. 2019, ApJL, 872, L10, doi: 10.3847/2041-8213/ab042a

Souami, D., & Souchay, J. 2012, A&A, 543, A133, doi: 10.1051/0004-6361/201219011

Spencer, J. R., Stern, S. A., Moore, J. M., et al. 2020, Science, 367, aay3999, doi: 10.1126/science.aay3999

Stansberry, J. A., Fraser, W. C., Trilling, D. E., et al. 2021, An Ultra-Sensitive Pencil Beam Search for 10 km Trans-Neptunian Objects, JWST Proposal. Cycle 1, ID. #1568

Steel, D. 1998, P&SS, 46, 473, doi: 10.1016/S0032-0633(97)00232-8

Stern, S. A., Bagenal, F., Ennico, K., et al. 2015, Science, 350, aad1815, doi: 10.1126/science.aad1815

Tabachnik, S. A., & Evans, N. W. 2000, MNRAS, 319, 63, doi: 10.1046/j.1365-8711.2000.03760.x

Tamayo, D., Rein, H., Shi, P., & Hernandez, D. M. 2020, MNRAS, 491, 2885, doi: 10.1093/mnras/stz2870

Thomas, F., & Morbidelli, A. 1996, Celestial Mechanics and Dynamical Astronomy, 64, 209, doi: 10.1007/BF00728348

Tombaugh, C. W. 1946, Leaflet of the Astronomical Society of the Pacific, 5, 73

Tonry, J. L., Stubbs, C. W., Lykke, K. R., et al. 2012, ApJ, 750, 99, doi: 10.1088/0004-637X/750/2/99

Trilling, D., Gerdes, D., Trujillo, C., et al. 2019, in EPSC-DPS Joint Meeting 2019, Vol. 2019, EPSC–DPS2019–395

Trujillo, C. A., & Sheppard, S. S. 2014, Nature, 507

Tsiganis, K., Gomes, R., Morbidelli, A., & Levison, H. F. 2005, Nature, 435, 459, doi: 10.1038/nature03539

Tyson, J. A., Guhathakurta, P., Bernstein, G. M., & Hut, P. 1992, in American Astronomical Society Meeting Abstracts, Vol. 181, American Astronomical Society Meeting Abstracts, 06.10

Uehata, K., Terai, T., Ohtsuki, K., & Yoshida, F. 2022, AJ, 163, 213, doi: 10.3847/1538-3881/ac5b6d

Valdes, F., Gruendl, R., & DES Project. 2014, in Astronomical Society of the Pacific Conference Series, Vol. 485, Astronomical Data Analysis Software and Systems XXIII, ed. N. Manset & P. Forshay, 379

Valtonen, M. J. 1983, The Observatory, 103, 1

Valtonen, M. J., & Innanen, K. A. 1982, ApJ, 255, 307, doi: 10.1086/159830

Van Houten, C., Herget, P., & Marsden, B. 1984, Icarus, 59, 1, doi: https://doi.org/10.1016/0019-1035(84)90051-4

van Houten, C. J., van Houten-Groeneveld, I., Herget, P., & Gehrels, T. 1970, A&AS, 2, 339

van Houten-Groeneveld, I., van Houten, C. J., Wisse-Schouten, M., Bardwell, C., & Gehrels, T. 1989, A&A, 224, 299

Verbiscer, A. J., et al. 2022, The Planetary Science Journal, 3, doi: 10.3847/PSJ/ac63a6

Volk, K., & Malhotra, R. 2019, The Astronomical Journal, 158, doi: 10.3847/1538-3881/ab2639

von Zeipel, H. 1910, Astronomische Nachrichten, 183, 345, doi: 10.1002/asna.19091832202

Whidden, P. J., Bryce Kalmbach, J., Connolly, A. J., et al. 2019, AJ, 157, 119, doi: 10.3847/1538-3881/aafd2d

Whiteley, R. J., & Tholen, D. J. 1998, Icarus, 136, 154, doi: 10.1006/icar.1998.5995

Yabushita, S. 1980, A&A, 85, 77

Yoshida, F., Terai, T., Ito, T., et al. 2019, P&SS, 169, 78, doi: 10.1016/j.pss.2019.02.003

Yoshikawa, M., Tsuda, Y., Watanabe, S., et al. 2018, in Lunar and Planetary Science Conference, Lunar and Planetary Science Conference, 1771

Youdin, A. N., & Goodman, J. 2005, ApJ, 620, 459, doi: 10.1086/426895

Zhou, L., Xu, Y.-B., Zhou, L.-Y., Dvorak, R., & Li, J. 2019, A&A, 622, A97, doi: 10.1051/0004-6361/201834026