The Characterization and Discovery of Solar System Small Bodies in Modern Astronomical Surveys

Joachim Moeyens

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

Doctor of Philosophy

University of Washington

2022

Reading Committee:

Mario Jurić, Chair

Željko Ivezić

Andrew Connolly

Program Authorized to Offer Degree:

Astronomy

University of Washington

Abstract

The Characterization and Discovery of Solar System Small Bodies in Modern Astronomical Surveys

Joachim Moeyens

Chair of the Supervisory Committee:
Associate Professor of Astronomy Mario Jurić
Department of Astronomy

Starting in 2024, the Vera C. Rubin Observatory will conduct a comprehensive survey of the night sky named the Legacy Survey of Space and Time (LSST). LSST will observe the Southern sky every three nights continuously for ten years and is predicted to discover six million new minor planets. Claims in the literature have suggested that optically-derived diameters from such surveys are only accurate to within 50% compared to IR-based estimates. Using "Asteroid Thermal Modeling" (ATM), we show that optical measurements can be used to constrain asteroid diameters with 0.4% bias and scatter of only 17%. LSST will discover minor planets by observing "tracklets": intranight linkages of two or more observations. The requirement to observe tracklets places a strong constraint on cadence and makes datasets not constructed with such a cadence unsuitable for discovery searches. We present "Tracklet-less Heliocentric Orbit Recovery" (THOR), a cadence-independent discovery algorithm that does not require tracklets. We apply THOR to two weeks of observations from the Zwicky Transient Facility (ZTF) and show that it can recover 1.5-2x as many asteroids as its tracklet-based counterparts. We describe our efforts to deploy THOR on the Asteroid Discovery, Analysis, and Mapping (ADAM) platform – the beginnings of a cloud-based discovery service which we name ADAM::THOR. We use the prototype service to search 0.2% of the exposures contained in the NOIRLab Source Catalog and discover 104 new minor planets.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

GLOSSARY

APPLICATION PROGRAMMING INTERFACE (API): The method by which two computer or software systems communicate. In the cloud computing context, APIs are typically defined to expose services to the users both public and private of a cloud-based architecture.

ASTEROID, DISCOVERY, ANALYSIS, AND MAPPING (`ADAM`): A cloud-based astrodynamics platform designed to enable computationally intensive research in planetary science. Developed and maintained by the Asteroid Institute, a program of the B612 Foundation.

ASTEROID INSTITUTE: A virtual (remote) institute launched by the B612 Foundation in collaboration with the University of Washington that focuses primarily on Solar System small body research in the context of planetary defense.

B612 FOUNDATION: A private, charitable, non-profit organization leading research initiatives in planetary defense from asteroid impacts. B612 is named after the hypothetical asteroid home of the Little Prince in Antoine de Saint-Exupéry's Le Petit Prince. See: https://b612foundation.org/.

DATA INTENSIVE RESEARCH IN ASTROPHYSICS AND COSMOLOGY (DIRAC) INSTITUTE: A University of Washington-based center for advancing research in astronomy through the development of state-of-the-art algorithms and datasets that enable ground-breaking science. See: https://dirac.astro.washington.edu/.

GOOGLE CLOUD: Google's cloud infrastructure for services such as data analytics and analysis, storage, and computing. See: https://cloud.google.com/.

HELIOCENTRIC DISTANCE: The distance between the Sun and the target body, typically represented with the scalar, r.

HELIOCENTRIC POSITION VECTOR: The three-dimensional vector to a target body with its origin at the Sun, typically represented with the vector, $\vec{r}$. The magnitude of this vector is the heliocentric distance, r.

HIERARCHICAL EQUAL AREA ISOLATITUDE PIXELATION (HEALPIX): A scheme to visualize and divide data on a sphere into pixels of equal area. See Gorski et al. 2005 for algorithmic details.

SMALL BODY DATABASE BROWSER (SBDB): Jet Propulsion Laboratory's database of Solar System small bodies. The database and service maintains the orbits, physical characteristics, discovery circumstances, and a variety of other data productions of the known population of small bodies. See: https://ssd.jpl.nasa.gov/tools/sbdb_lookup.html/.

LEGACY SURVEY OF SPACE AND TIME (LSST): A ten-year multi-band photometric survey of the Southern sky to be conducted by the Vera C. Rubin Observatory starting in 2024. LSST will aim to address four primary science drivers: probing dark energy and dark matter, taking an inventory of the Solar System, exploring the transient optical sky, and mapping the Milky Way. See Ivezić et al. 2019 for survey details.

MINOR PLANET CENTER (MPC): The International Astronomical Union Minor Planet Center is the organization responsible for maintaining the official catalog of orbits and the observations of minor planets, comets, and irregular satellites of the planets in the Solar System. See: https://www.minorplanetcenter.net/.

NOIRLAB SOURCE CATALOG (NSC): A comprehensive catalog of nearly all of the public imaging data in NOIRLab's Astro Data Archive. The NSC catalogs 68 billion point-source meaurements from three observing facilities: CTIO (4m Blanco + DECam), KPNO (4m Mayall + Mosaic3), and Steward Observatory (2.3m Bok + 90Prime). See Nidever et al. 2021 for dataset details.

ORBIT DETERMINATION: Process by which the orbit of a target body is calculated from astronomical observations.

TOPOCENTRIC DISTANCE: The distance between an observer and the target body, typically represented with the scalar, Δ.

TOPOCENTRIC POSITION VECTOR: The three-dimensional vector to a target body with its origin at an observer, typically represented with the vector, $\vec{\Delta}$. The magnitude of this vector is the topocentric distance, Δ.

TRACKLET: A combination of two or more intranight observations that constrain the position and rate of motion of potential moving objects.

TRACKLET-LESS HELIOCENTRIC ORBIT RECOVERY (THOR): Small body discovery algorithm that does not require tracklets for discovery. Instead, THOR is able to discover Solar System small bodies using single detections linked across multiple nights.

VERA C. RUBIN OBSERVATORY: A next generation optical multi-band observatory under construction in Chile. The Vera C. Rubin Observatory will use the Simonyi Survey Telescope which features a 8.4 m primary mirror, a 3.5 degree diameter field of view, and a 3.2 gigapixel camera to conduct the Legacy Survey of Space and Time (LSST). See Ivezić et al. 2019 for survey details.

VIRTUAL MACHINE (VM): A computer emulated by software on a separate (remote) machine. Multiple virtual machines can be launched on the same hardware and will mimic the behavior of individual computing systems. This is the common use case in cloud-based infrastructure, where virtual machines of different types and configurations can be requested to run compute-intensive workloads.

WIDE-FIELD INFRARED SURVEY EXPLORER (WISE): A space-based infrared telescope launched in 2009 that conducted an all-sky infrared survey in four bands (3.4, 4.6, 12, and 22 μm) over a ten-month period. In 2013 the telescope was reactivated as the Near Earth Object Wide-Field Infrared Survey Explorer (NEOWISE).

ZWICKY TRANSIENT FACILITY (ZTF): A robotic time-domain survey of the night sky conducted at the Palomar Observatory using the 48 inch Samuel Oschin Telescope. ZTF features a 47 deg^2 field of view and 600 megapixel CCD. See Graham et al. 2019 for survey details.

ACKNOWLEDGMENTS

The author wishes to thank the University of Washington DiRAC Institute for fostering a world-class community of researchers, scientists, teachers, and mentors. This work would not have been possible without the countless suggestions and many meaningful contributions offered by the Institute's members, many of whom have been rightfully acknowledged as co-authors on the published articles summarized in this document.

Through their unparalleled generosity, the donors of the B612 Foundation funded the last two years of the author's doctoral studies. Without their generous support, the work detailed in the latter chapters of this work would not have been possible. Funding from the Asteroid Institute, a program of the B612 Foundation, made it so the author could continue to focus on the research he was passionate about. The author is eternally grateful and humbled by the opportunity he was gifted.

The author wishes to acknowledge the support of the Asteroid Institute and B612 Foundation. This research was made possible by leadership gifts to support the ADAM project provided by the W.K. Bowes Jr. Foundation, Steve Jurvetson, the McGregor Girand Charitable Endowment, The P. Rawls Family Fund, Tito's CHEERS, Yishan Wong and Kimberly Algeri-Wong, and two anonymous leadership donors, along with Founding Circle and Asteroid Circle members: B. Anders, M. Armstrong, R. Armstrong, G. Baehr, B. Burton, D. Carlson, S. Cerf, V. Cerf, C. Chapman, Y. Chapman, J. Chervenak, D. Corrigan, E. Corrigan, A. Denton, E. Dyson, A. Eustace, A. Fritz, L. Fritz, S. Galitsky, E. Gillum, L. Girand, Glaser Progress Foundation, D. Glasgow, A. Gleckler, J. Grimm, S. Grimm, G. Gruener,

V. K. Hsu Sons Foundation Ltd., J. Huang, J. D. Jameson, J. Jameson, Margaret Jonsson Family Foundation, D. Kaiser, S. Krausz, V. Lašas, J. Leszczenski, D. Liddle, S. Mak, G.McAdoo, S. McGregor, J. Mercer, J. Montrym, M. Mullenweg, D. Murphy, P. Norvig, S. Pishevar, R. Quindlen, N. Ramsey, R. Rothrock, E. Sahakian, R. Schweickart, A. Slater, T. Trueman, F. B. Vaughn, R. C. Vaughn, M. Welty, S. Welty, A. Werner, B. Wheeler, M. Wyndowe, and seven anonymous Founding and Asteroid Circle donors in addition to donors from over 46 countries around the world.

J. Moeyens wishes to acknowledge the support of the Washington Research Foundation Data Science Term Chair fund, as well as the University of Washington Provosts Initiative in Data-Intensive Discovery. The author acknowledges support from the University of Washington College of Arts and Sciences, the Department of Astronomy, and the DiRAC Institute. The DiRAC Institute is supported through generous gifts from the Charles and Lisa Simonyi Fund for Arts and Sciences, and the Washington Research Foundation.

J. Moeyens thanks the LSST Corporation Data Science Fellowship Program, his time as a Fellow has benefited this work.

This publication makes use of data products from the Wide-field Infrared Survey Explorer, which is a joint project of the University of California, Los Angeles, and the Jet Propulsion Laboratory/California Institute of Technology, funded by the National Aeronautics and Space Administration.

Based on observations obtained with the Samuel Oschin 48 inch Telescope and the 60 inch Telescope at the Palomar Observatory as part of the Zwicky Transient Facility project. ZTF is supported by the National Science Foundation under Grant No. AST-1440341 and a collaboration including Caltech, IPAC, the Weizmann Institute for Science, the Oskar Klein Center at Stockholm University, the University of Maryland, the University of Washington, Deutsches Elektronen-Synchrotron and Humboldt University, Los Alamos National Laboratories, the TANGO Consortium of Taiwan, the University of Wisconsin at Milwaukee, and

Lawrence Berkeley National Laboratories. Operations are conducted by COO, IPAC, and UW.

This research has made use of data and/or services provided by the International Astronomical Union's Minor Planet Center.

This work was facilitated through the use of advanced computational, storage, and networking infrastructure provided by the Hyak supercomputer system and funded by the STF at the University of Washington.

This work was supported by the Washington Research Foundation and by a Data Science Environments project award from the Gordon and Betty Moore Foundation (Award 2013-10-29) and the Alfred P. Sloan Foundation (Award 3835) to the University of Washington eScience Institute. This research has made use of data and/or services provided by the International Astronomical Union's Minor Planet Center.

We thank the organizers of an asteroid modeling workshop at the University of Washington (October 3–4, 2016), where the second chapter of this work was conceived: LSST Corporation, NASA NEO Office, and the B612 Foundation.

The author would like to extend a thank you to Mario Jurić, Željko Ivezić, Siegfried Eggl, Lynne Jones, Andy Connolly, Victoria Meadows, Matthew McQuinn, Scott Anderson, and Dino Bektešević at the University of Washington for their many helpful comments on this work.

The author is grateful to the anonymous referees for thorough reviews of the multiple revisions of the published works summarized in this document.

DEDICATION

To my parents who put me on this rock, to my sisters who taught me all about humility, and to my fiancée who has made the journey enjoyable.

Chapter 1

INTRODUCTION

The number of Solar System minor planet discoveries is growing rapidly, thanks to the continuation of present-day surveys such as the Panoramic Survey Telescope and Rapid Response System (Pan-STARRS) (Chambers et al., 2016), Catalina Sky Survey (CSS) (Larson et al., 2003), the Asteroid Terrestrial-impact Last Alert System (ATLAS) (Tonry et al., 2018), and the Zwicky Transient Facility (ZTF) (Graham et al., 2019). In the past decade, surveys such as Wide-field Infrared Survey Explorer (WISE) (Wright et al., 2010) and the Dark Energy Survey (DES) (Abbott et al., 2018) have contributed significant volumes of observations of small bodies in the Solar System. In the next decade, the Vera C. Rubin Observatory's Legacy Survey of Space and Time (LSST) (Ivezić et al., 2019) and NEO Surveyor (Sonnett et al., 2020) will increase the number of known objects from the currently known one million to about six million minor planets.

Such an increase in discoveries will present researchers with deep insight into the different populations of Solar System small bodies. Understanding their spatial distribution, their composition and physical properties will enable a higher-resolution look into the dynamical evolution of our Solar System and the material and processes that formed it. Among the parameters of interest is the characterization of asteroid diameters, or sizes, and parameters associated with thermal emission and reflection, namely emissivity and albedo as a function of wavelength.

Asteroid thermal flux modeling aims to estimate an asteroid's size (cross-section-equivalent diameter, hereafter diameter), its surface temperature distribution, and sometimes other physical properties such as emissivity, from measured infrared (IR) fluxes. The largest dataset of infrared flux measurements for asteroids was contributed by the WISE survey (Wright et al., 2010) and analyzed by the associated NEOWISE team (Mainzer et al. (2011), and references therein). Flux measurements in four WISE bands provide strong constraints on asteroid sizes and emissivities as a function of wavelength. A series of papers that produced size estimates for about 164,000 asteroids, as well as constraints on asteroid emissivity properties, was reviewed and summarized by Mainzer et al. (2015).

It appears that the pioneering analysis by the NEOWISE team can be improved in various

ways, as argued by Myhrvold (2018a,b). In particular, ignoring reflected sunlight can induce biases in estimated asteroid sizes and lead to underestimated size uncertainties. In addition, best-fit sizes can be biased due to assumptions on the priors for an asteroid's emissivity as a function of wavelength, which varies with the chemical composition of an asteroid's surface. Further biases are introduced by the use of fairly simplistic static thermal models to estimate diameters. The need for improvements in data analysis was recently acknowledged by Wright et al. (2018), but it appears that a number of data analysis issues remain open. In particular, the behavior of systematic and random (statistical) uncertainties for the best-fit parameters remains an active research topic (see for example Masiero et al. (2018a); Mommert et al. (2018); Wright (2007)).

Given that the WISE dataset is by far the largest of its kind and will not soon be surpassed, it is prudent to focus on improved data analysis. In Chapter 2, we present "Asteroid Thermal Modeling" (ATM), an open-source Python package that can model asteroid fluxes using the most common static thermal models. In addition to modeling fluxes, ATM can fit for diameter, characteristic temperature, and emissivity given observed fluxes with full Bayesian treatment of parameter priors and Markov Chain Monte Carlo (MCMC) sampling of the posterior probability density function.

Using ATM, we revisit the analysis done by the NEOWISE team and show that ATM can faithfully reproduce their published diameters for a well-observed subset of asteroids. To understand the role of optical surveys in asteroid characterization, we analyze correlations between WISE-based (IR) size measurements and SDSS-based (optical) colors and show that, counter to "popular" belief, optical measurements of asteroids can be used to constrain asteroid diameters. Lastly, using IR best-fit modeling parameters we identify 13 metallic asteroid candidates that could not have been identified solely using their optical colors.

Given that both optical and IR observations can be used to study and characterize a variety of properties of small bodies in the Solar System, the task remains to identify and discover asteroids and comets in astronomical surveys. Identifying minor planets in survey images and linking their detections into orbits continues to be a challenging problem. First,

linking asteroid detections across multiple nights is difficult due to the sheer number of possible linkages, and it is made even more challenging by the presence of false positives (Kubica et al., 2007; Denneau et al., 2013; Vereš & Chesley, 2017a,b; Jones et al., 2018). Second, the motion of the observer makes the linking problem nonlinear because minor planets will exhibit higher-order motion on the topocentric sky over the course of weeks (Holman et al., 2018). Finally, once potential linkages have been established, they need to be confirmed as possible orbits using computationally expensive orbit determination software. For example, the Vera C. Rubin Observatory estimates it will discover nearly six million Main Belt asteroids that will be observed hundreds of times over the course of its ten-year survey. Naively attempting to link hundreds of millions of asteroid detections over a ten-year period is not computationally feasible (Jones et al., 2018).

To make the linking problem more computationally tractable, surveys that aim to discover minor planets focus on constructing "tracklets": two-dimensional sky-plane motion vectors consisting of two or more detections spaced typically 20–90 minutes apart that constrain the direction and rate of motion of potential moving objects. Tracklets are constructed to reduce the number of possible linkages that could be formed by providing information on plausible direction and sky-plane angular velocity (Kubica et al., 2007). They are then linked into internight linkages known as "tracks": sky-plane paths of motion containing several tracklets spanning up to ∼15 nights, typically modeled with low-order polynomials. In the case of the LSST, for a moving object to be discoverable, it must be observed at least twice a night on at least three unique nights within a 15-day window to go through the tracklet-to-track creation process (Jones et al., 2018). In part due to the relative motion of the observer and the rate of motion of moving objects, both tracklets and tracks can exhibit high residuals relative to the fitted low-order polynomial, requiring relaxed fitting tolerances that can in turn lead to the creation of many spurious candidate linkages. Orbit determination (OD) algorithms are therefore required to run on each candidate track so spurious linkages can be identified and removed.

The Zwicky Transient Facility (ZTF), an optical time-domain survey scanning the entire

northern hemisphere of sky at a rate of more than 3700 deg^2 hr^{-1}, can be seen as a precursor to the LSST (Bellm et al., 2019; Graham et al., 2019; Masci et al., 2019). ZTF uses the ZTF Moving Object Discovery Engine (ZMODE; Masci et al. 2019) algorithm. Instead of linking tracklets directly into tracks, ZMODE first attempts to build a "stringlet." A stringlet forms an intermediate step between tracklets and tracks that allows for the linking of pairs of detections across nights before tracks are built. This approach was designed to accommodate ZTF's cadence during its main survey, where the cadence is frequently too sparse to form short intranight tracklets.

Recent work by Holman et al. (2018) has shown promising results by shifting the reference frame for linking detections to the heliocenter. By assuming a heliocentric distance and its rate of change, cleverly fitting tracklets for the remaining unknown parameters in inertial space, then propagating the resulting "arrows" to a common epoch, arrows corresponding to the same minor planets will form clusters. These clusters can then be extracted and subsequently validated by orbit fitting. As a testament to the effectiveness of HelioLinC, some 200,000 new minor planet orbits were recovered from the Minor Planet Center's Isolated Tracklet File (Holman et al., 2018).

Common in all of these approaches is the requirement to build tracklets, which in turn requires a telescope to perform multiple revisits to the same field in a night, then more revisits a few nights later, and so on. For a survey such as the LSST, which aims to balance four different science drivers, requiring such a cadence decreases the overall ease with which the other science drivers can be accommodated. Perhaps even more consequential is that any astronomical dataset not constructed with a tracklet cadence is not a dataset suited for minor planet discovery (the vast majority of astronomical datasets!). It is therefore prudent to investigate whether linking algorithms that are cadence-independent can be constructed and whether such algorithms can perform as well as – or better than – the current methods. An algorithm that does not demand a high revisit cadence could increase the efficiency of future surveys, as well as help multi-science missions such as the LSST.

In Chapter 3, we present "Tracklet-less Heliocentric Orbit Recovery" (THOR), a cadence-

and observer-independent discovery algorithm. We apply THOR to a short two-week slice of observations made by ZTF and show that it can recover 1.5-2 times as many known objects as the current state-of-the-art methods. We discuss how THOR would have performed had it been running as ZTF's linking algorithm when the observations were made. We conclude the analysis by identifying 10 discovery candidates that have remained undiscovered in the ZTF dataset for three years.

Finally, in Chapter 4, we discuss the ongoing work to prepare THOR for the Legacy Survey of Space and Time (LSST). In collaboration with the Asteroid Institute, a program of the B612 Foundation, we describe our efforts to deploy THOR on the Asteroid Discovery, Analysis, and Mapping (ADAM) platform: a scalable cloud-based astrodynamics research platform designed to enable compute-intensive research in planetary science and planetary defense. The deployment of THOR on ADAM, we name ADAM::THOR – the beginnings of a cloud-based asteroid discovery service.

We identify the second data release of the NOIRLab Source Catalog (Nidever et al., 2021) as an ideal precursor dataset while we await the first data from LSST. We describe in detail the ADAM::THOR infrastructure and the several enhancements made to THOR to enable searches of the NSC. We report on the initial results of these searches using ADAM::THOR. After processing just 0.2% of the exposures contained within the NSC, ADAM::THOR found 113 discovery candidates, of which 92 were observed for the first time, 12 were elevated out of the Isolated Tracklet File (ITF), and 9 were known objects that our attribution pipeline missed. We end by looking forward towards the LSST and describe the work that remains to make THOR a plausible successor to its tracklet-based counterparts.

Chapter 2

CHARACTERIZATION OF ASTEROIDS

2.1 Introduction

Asteroid thermal flux modeling aims to estimate an asteroid's size (cross-section-equivalent diameter, hereafter diameter), its surface temperature distribution, and sometimes other physical properties such as emissivity, from measured infrared (IR) fluxes. The largest dataset of infrared flux measurements for asteroids was recently contributed by the WISE survey (Wright et al., 2010) and analyzed by the associated NEOWISE team [Mainzer et al., 2011, and references therein]. Flux measurements in four WISE bands provide strong constraints on asteroid sizes and emissivities as a function of wavelength. A series of papers that produced size estimates for about 164,000 asteroids, as well as constraints on asteroid emissivity properties, was reviewed and summarized by Mainzer et al. (2015).

It appears that the pioneering analysis by the NEOWISE team can be improved in various ways, as argued by Myhrvold (2018a,b). In particular, ignoring reflected sunlight can induce biases in estimated asteroid sizes and lead to underestimated size uncertainties. In addition, best-fit sizes can be biased due to assumptions on the priors for an asteroid's emissivity as a function of wavelength, which varies with the chemical composition of an asteroid's surface. Further biases are introduced by the use of fairly simplistic static thermal models to calculate cross-section-equivalent diameters. The need for improvements in data analysis was recently acknowledged by Wright et al. (2018), but it appears that a number of data analysis issues remain open. In particular, the behavior of systematic and random (statistical) uncertainties for the best-fit parameters remains an active research topic (see for example Masiero et al. 2018a; Mommert et al. 2018; Wright 2007).

Given that the WISE dataset is by far the largest of its kind and will not soon be surpassed, it is prudent to focus on improved data analysis. We aim to contribute to such improvements by publicly releasing a new Python modeling tool, ATM (Asteroid Thermal Modeling)[1], designed to enable easy fitting of the most common static thermal models to WISE and other asteroid flux measurements. Data analysis software is often a crucial com-

[1]See https://github.com/moeyensj/atm.

ponent in delivering scientific results, as vividly exemplified by this case, and thus discussions of scientific reproducibility and transparency can be greatly enhanced by collaborative software development and code sharing. Thanks to rapidly developing tools and technologies, such as Jupyter Notebooks, Python, and GitHub, performing these steps in an open-source environment is now easier than ever. By releasing ATM we aim to increase reproducibility – a fundamental tenet of the scientific process.

In Section 2.2 we describe the mathematical and physical model underlying ATM, its Python implementation, and we discuss model validation using observational and modeled data for asteroid sizes from the literature. The capabilities of ATM, with emphasis on various treatments of Bayesian priors when also fitting emissivity are further illustrated in 2.2 using three well-observed asteroids. In Section 2.3 we apply ATM to a "gold" sample of ∼7000 best-observed asteroids from the NEOWISE dataset; in addition to best-fit diameters, we also obtain best-fit values for the emissivity of each asteroid across WISE bands W1 and W2. The best-fit sizes are compared to the values obtained by the NEOWISE team (the 2016 Planetary Data System version[2] (Mainzer et al., 2016)). In Section 2.4 we match the "gold" sample to optical data from the SDSS Moving Object Catalog, compute visual albedo p_V, and study its correlation with optical colors. We also discuss how optical colors can be used to estimate p_V and asteroid sizes when adequate infrared data are not available. Our results are summarized in Section 2.5.

As part of the ATM package release, we also include all data files used in this work including NEOWISE, Minor Planet Center, and SDSS Moving Object Catalog data. We provide Jupyter Notebooks with examples of how to select subsamples of objects, process and filter their flux measurements, and how to use ATM to estimate their diameters and infrared emissivities. All of the analysis presented here, including all the figures and tables, can be easily reproduced using the Notebooks released with the ATM package on GitHub. In particular, `https://github.com/moeyensj/atm_notebooks/blob/main/paper1/README.md` lists

[2]See https://sbn.psi.edu/pds/resource/neowisediam.html.

the Notebooks required to reproduce each figure in this paper. In addition to the Python code and Jupyter Notebooks, we also provide Docker containers tagged to coincide with the publication of this paper which allow the user to access and use all the components necessary (operating system, software dependencies, data files, the ATM Python code, and the analysis Notebooks) to reproduce every figure in this paper in a container-based environment.

2.2 ATM: An Open-source Tool For Asteroid Thermal Modeling

A detailed discussion of the relevant physics involved in asteroid thermal flux modeling, and a summary of models proposed in the literature, are presented in Myhrvold (2018a) and references therein. Here we only summarize the main results needed to understand how ATM works and what it computes. We also describe its Python implementation and discuss model validation using observational and model data for asteroid sizes from literature.

Static or instantaneous thermal models, like the ones discussed in this work, offer a relatively simple but effective method to estimate an asteroid's diameter. These models assume the asteroid is a non-rotating or simply rotating sphere with a well-defined surface temperature distribution which is often expressed as a function of angular distance from the subsolar point. Static thermal models differ significantly from more robust thermophysical models which seek to model conductivity, take into account spin axis, rotation, and orientation, consider the granularity and composition of the surface, and also how composition changes with depth. However, with added complexity comes an added increase in computational cost and an increased need for high quality and quantity of observations. Static thermal modeling, while more simple, offers a computationally feasible method to get estimates of diameters and other properties but with the bonus that they can be applied at scale. And thus, instantaneous thermal modeling is a natural choice for the WISE dataset. We now summarize the underlying physics that drive these models.

2.2.1 The Asteroid Flux Model Summary

We first introduce $F_\nu(\lambda)$: the specific flux (flux per unit frequency, ν) from an object. The SI units for F_ν are W m^{-2} Hz^{-1} ($= 10^3$ erg cm^{-2} s^{-1}). The specific flux can also be defined per unit wavelength, F_λ, using energy conservation $F_\nu|d_\nu| = F_\lambda|d_\lambda|$ and $\lambda\nu = c$. The choice of F_ν, as opposed to F_λ, is completely arbitrary. Similarly, the running variable can be either λ or ν, and the choice of λ is more convenient in this context.

The model flux from an asteroid, $F_\nu^{ast}(\lambda)$, corresponding to flux detected by the observer, $F_\nu^{obs}(\lambda)$, is the sum of the emitted thermal flux controlled by the asteroid's surface temperature distribution, and the portion of the incident solar flux reflected by the asteroid,

$$F_\nu^{ast}(\lambda) = F_\nu^{th}(\lambda) + F_\nu^{ref}(\lambda) \tag{2.1}$$

A given model spectrum $F_\nu^{ast}(\lambda)$, corresponding to observational quantity $F_\nu^{obs}(\lambda)$, is integrated over the bandpass (assumed known hereafter) to obtain observed in-band model fluxes for a given instrument. For example, Wright (2013) has derived simple quadrature formulae[3] that can be used to efficiently and accurately compute in-band fluxes for the four WISE bands from model flux $F_\nu^{ast}(\lambda)$.

Both $F_\nu^{th}(\lambda)$ and $F_\nu^{ref}(\lambda)$ depend on the relative positions of the Sun, the asteroid, and the observer, the asteroid's diameter, D, and its emissivity, $\epsilon(\lambda)$, which controls the balance between absorbed/emitted and reflected incident flux. The observing geometry is fully described by the heliocentric distance, r, the topocentric distance, Δ, and the angle subtended on the asteroid's surface by the lines of sight towards the Sun and the observer, the so-called phase angle, α. Hereafter we assume that asteroid orbital parameters are known and that r, Δ, and α can be easily computed using standard and readily-available tools (e.g., the JPL HORIZONS service[4].; `oorb`, Granvik et al. (2009)). To simplify nomenclature, we do not explicitly list these independent variables, unless necessary to avoid confusion.

---

[3]See Appendix A for a slight correction to the quadrature formula for the W3 band.

[4]See https://ssd.jpl.nasa.gov/horizons/.

2.2.1.1 The Emitted Flux

The observed thermal flux is obtained by integrating the emitted thermal flux per unit area over the visible surface of the asteroid,

$$F_{\nu}^{th}(\lambda) = \left(\frac{D}{2\Delta}\right)^2 \epsilon(\lambda) \int_{-\pi/2}^{\pi/2} \int_{-\pi/2-\alpha}^{\pi/2+\alpha} \pi B_{\nu}(T(\theta,\phi),\lambda) \cos(\phi-\alpha) \cos^2(\theta) d\phi d\theta \tag{2.2}$$

where θ, the angle of latitude, and ϕ, the angle of longitude, are the integration variables over the asteroid's surface (here we use geographic coordinates with $\theta = 0$ and $\phi = 0$ at the subsolar point, and both ranging from $-\pi$ to π; note that Myhrvold (2018a) used ISO coordinates with $0 < \theta < 2\pi$, and B_{ν} is the Planck function.

In this context of thermal emission, it is implied that asteroids are approximated as perfect spherical Lambertian emitters which follow Lambert's cosine rule. In reality, both scattered light and thermal emission may exhibit non-Lambertian properties. However, the effect in scattered light (which shows a strong opposition surge – a peaked reflectance for phase angles near zero) is stronger than the effect of non-Lambertian beaming in the case of thermal emission (see e.g. Rozitis & Green (2011)).

The temperature distribution across an asteroid's surface, $T(\theta,\phi)$, is model dependent. ATM implements the three most common instantaneous thermal models: the Standard Thermal Model (STM) (Lebofsky et al., 1986), the Fast Rotating Model (FRM) (Lebofsky & Spencer, 1989), and the Near-Earth Asteroid Thermal Model (NEATM) (Harris, 1998). For a comparison of these models and an analysis of their validity, please see Wright (2007); Mommert et al. (2018), and references therein. Common to all models is a temperature scale set by the so-called subsolar temperature, T_{ss}, which is the highest temperature on the asteroid's surface. Depending on the adopted model, the temperature variation across the surface can range in complexity from being constant to having a strong temperature gradient from the subsolar point (e.g., see Figure 1 in Mommert et al. (2018)). The temperature distribution for the STM and NEATM models is equivalent to

$$T(\theta,\phi) = \begin{cases} T_{ss}(\cos\theta\cos\phi)^{1/4} & \text{if } \theta \leq \pi/2, \phi \leq \pi/2 \\ 0 & \text{if } \theta > \pi/2, \phi > \pi/2 \end{cases} \tag{2.3}$$

In addition to differences in surface temperature distribution, the three implemented models may use different integration variables and different methods to account for phase angle. Equation 2.2 follows innovations introduced by NEATM, where the observer's direction cosine is directly integrated over the surface of a model asteroid.

To account for effects such as surface roughness, the subsolar temperature is modulated using the "beaming parameter", η. For a more detailed discussion of this parameter, please see Myhrvold (2018a). The η parameter is not constant and can vary with observing geometry, particularly, phase angle (Wright, 2007).

Given a model parameter for the temperature scale, T_1, defined as the subsolar temperature when $r = 1$ au, the subsolar temperature is simply

$$T_{ss} = \left(\frac{1 \text{ au}}{r}\right)^{1/2} T_1 \tag{2.4}$$

Unlike T_{ss}, which depends on the heliocentric distance, T_1 is nearly constant for a given asteroid (not exactly constant because of its slight dependence on η, see Equation 2.5 below).

The energy balance, discussed below, connects T_1 with the incident solar flux, an asteroid's physical properties and other model parameters. Myhrvold (2018a) made an important point that when fitting a model to data, all the other parameters are not directly relevant – it is only T_1 that controls the model fluxes and thus it is only T_1 that is directly constrained by the infrared flux data for a given observer's position.

2.2.1.2 The Energy balance and the Meaning of Best-fit Parameter T_1

Using the energy balance equation that equates the absorbed incident solar flux with the flux emitted by an asteroid, it can be shown (e.g., see Myhrvold 2018a and references therein) that

$$T_1 = \left(\frac{S(1-A)}{\sigma \eta \epsilon_B} \right)^{1/4} \tag{2.5}$$

where S is the solar constant at 1 au (S = 1360.8 W m^{-2}), $\sigma = 5.67 \times 10^{-8}$ W m^{-2} K^{-4} is the Stefan-Boltzmann constant, and A and ϵ_B are appropriately wavelength-averaged values of $1-\epsilon(\lambda)$ and $\epsilon(\lambda)$ over the incident and thermal flux distributions, respectively (for details, please see Section 3 in Myhrvold 2018a). It is usually assumed that $\epsilon_B = 0.9$, but the exact value is not crucial to modeling once the beaming parameter η is introduced – it is only the $\eta\epsilon_B$ product that can be constrained using best-fit T_1. Another conventional approximation is

$$A \approx A_V = p_V q, \tag{2.6}$$

where A_V is the Bond albedo (limited to the range 0–1), p_V is the geometric albedo in the visible band (can be larger than 1), and q is the empirically-derived phase integral. In the H-G magnitude system introduced by Bowell et al. (1989), $q(G) = 0.29 + 0.684G$, where G is the slope parameter of the phase function. A common assumption, when G is unknown, is $G = 0.15$. Therefore, the best-fit T_1 can be used to estimate p_V, given the value the $\eta\epsilon_B$ product.

The geometric albedo can also be constrained, when the asteroid's diameter, D, is known, using flux measurements at wavelengths sufficiently short for flux to be dominated by the reflected incident solar flux. For example, in the visual band

$$p_V = \left(\frac{1329 \text{ km}}{D} \right)^2 10^{-0.4H}, \tag{2.7}$$

where H is the asteroid's absolute magnitude in the visual band. When such measurements are available, the two constraints for p_V can be used to estimate the beaming parameter, η, or to simply check the model's internal consistency.

2.2.1.3 The Reflected Flux

Via Kirchoff's law, an asteroid's reflectivity, $p(\lambda)$ is

$$p(\lambda) = \frac{1 - \epsilon(\lambda)}{q(G)} \tag{2.8}$$

The reflected flux is then equivalent to

$$F_\nu^{ref}(\lambda) = \left(\frac{D}{2\Delta}\right)^2 \Psi(\alpha, G) \left[\frac{1 - \epsilon(\lambda)}{q(G)}\right] F_\nu^\odot(\lambda) \tag{2.9}$$

where the $H - G$ phase function, $\Psi(\alpha, G)$, and empirically derived phase integral, q(G), are purely geometric quantities that account for phase effects (Bowell et al., 1989).

The incident solar flux, $F_\nu^\odot(\lambda)$, at a distance r from the Sun, is given by

$$F_\nu^\odot(\lambda) = \left(\frac{R_\odot}{r}\right)^2 \pi B_\nu(T_\odot, \lambda) \tag{2.10}$$

where $T_\odot = 5778$ K and $R_\odot = 0.00465$ au. Therefore, when $\epsilon(\lambda)$ is known, the reflected flux is fully determined for a given observing geometry (note also that a fully known $\epsilon(\lambda)$ implies a given value of p_V). In Appendix D we examine the effect of using a more accurate solar spectrum on certain modeling scenarios presented in the latter sections of this work.

2.2.1.4 Fitting Model Fluxes to Data

The likelihood of obtaining N observed fluxes, F_i^{obs}, given model predictions, F_i^{ast}, can be written as (for details see, e.g., Chapter 4 in Ivezić et al. (2014))

$$L = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi(\sigma_i^2 + \Sigma^2)}} \exp\left(\frac{-(F_i^{obs} - F_i^{ast})^2}{2(\sigma_i^2 + \Sigma^2)}\right) \tag{2.11}$$

where σ_i is the flux measurement uncertainty and Σ accounts for variability. Assuming flat Bayesian priors for fitted model parameters, maximizing this likelihood function is equivalent to maximizing the Bayesian posterior probability density function. ATM also supports the Jeffreys' priors (flat distributions in $\log(D)$ and $\log(T_1)$, for details see Chapter 5 in Ivezić et al. (2014)).

The above expression assumes that the scatter of flux measurements around predicted model values follows a Gaussian distribution with standard deviation (measurement uncertainty) σ_i. In practice, most asteroids show variability with amplitudes ($\sim$1 mag) exceeding typical measurement uncertainties (≤ 0.2 mag for SNR ≥ 5). This variability is mostly due to non-spherical shapes, which are not captured by the model. For this reason, the likelihood expression includes the Σ^2 term. Given that $\Sigma \gg \sigma_i$ for reasonably high SNR, we have $\sigma_i^2 + \Sigma^2 \approx \Sigma^2$, which is presumed constant for a given asteroid. While it does not influence the values of the best-fit parameters, Σ does control their uncertainty. Following (Masiero et al., 2018b), we adopt $\Sigma = 0.2$ mag as typical uncertainty due to variability.

When emissivity $\epsilon(\lambda)$ is assumed known, the best-fit model parameters are effectively obtained by maximizing the log-likelihood in magnitude space, as a function of two free parameters, D and T_1 (for a discussion of emissivity fitting, see the next section),

$$\ln(L) = const. - \sum_{i=1}^{N} \left[m_i^{obs} - m_i^{ast}(D, T_1) \right]^2, \tag{2.12}$$

where m_i^{obs} are observed magnitudes and m_i^{ast} are their model predictions ($m = -2.5 \log_{10}(F) + const.$).

2.2.1.5 The Treatment of Emissivity, $\epsilon(\lambda)$

The assumed priors for $\epsilon(\lambda)$ play an important role in determining an asteroid's size. Technically, all fitted parameters require their prior probability distributions to be specified (for a discussion of Bayesian priors, see, for example, Chapter 5 in Ivezić et al. (2014)). When $\epsilon(\lambda)$, or any other parameter, is assumed to be known, their priors effectively become Dirac δ functions.

When $\epsilon(\lambda)$ is assumed known, in principle even just two infrared flux measurements can constrain an asteroid's diameter D and temperature T_1. In essence, the measured color (i.e. the ratio of the two measured fluxes) constrains T_1 and the overall flux level constrains D. When measurement errors are present, the constraints become more complex, as discussed

further below in the model validation section (Section 2.2.3). If more than two infrared flux measurements are available, if Σ is known from light curve analysis, and if measurement errors are sufficiently small, the model fluxes can be checked for internal consistency using statistical tests such as χ^2.

When $\epsilon(\lambda)$ is treated as a free parameter, care must be taken to avoid the $D\epsilon$ degeneracy when only thermal fluxes are measured: in Equation 2.2 parameters D and ϵ appear only as a product $D^2\epsilon$. This degeneracy can be broken when flux measurements at wavelengths where the reflected light component is non-negligible are available. Because in Equation 2.9 the reflected flux is proportional to $D^2(1-\epsilon)$ rather than proportional to $D^2\epsilon$, the $D-\epsilon$ degeneracy is broken and both ϵ and D can be estimated. The larger the contribution of the reflected flux to the total flux, the less correlated are the best-fit values of ϵ and D. We return to this point in Section 2.2.3.

In reality, a rather substantial variation in $\epsilon(\lambda)$ is observed for asteroids (e.g., see Figure 1 in Myhrvold (2018a)). If more than two flux measurements are available, as is the case for NEOWISE data with four bands, parameters from a judiciously chosen parameterization of $\epsilon(\lambda)$ can also be fit to data. In the case of four bandpasses, up to two such parameters can be well constrained[5] (in addition to D and T_1, for the total of four fitted parameters).

Unknown functions are often specified as power laws with two free parameters: the power-law index and its overall normalization. This common approximation would work poorly given the $\epsilon(\lambda)$ functions observed for asteroids. Instead, motivated by the observed behavior of the $\epsilon(\lambda)$ functions, we consider an ansatz ("a simplified procedure", or "a provisional mathematical assumption") where we fit ϵ_{W1W2}, the value of $\epsilon(\lambda)$ in the WISE bands 1 and 2, and ϵ_{W3W4}, the value of $\epsilon(\lambda)$ in the WISE bands 3 and 4. In other words, we approximate $\epsilon(\lambda)$ as a step function $\epsilon(\lambda) = \epsilon_{W1W2}$ for $\lambda < 6\ \mu m$, and $\epsilon(\lambda) = \epsilon_{W3W4}$ for $\lambda > 6\ \mu m$. Because the two shortest WISE bands include substantial contribution of reflected flux (e.g.,

[5]Here we ignore statistical techniques such as regularization, as well as hierarchical Bayesian modeling, which can be used to fit for more parameters than there are data points; for more details, please see Chapter 8 in Ivezić et al. (2014).

see Figure 1 in Myhrvold (2018a)), constraints can be placed on ϵ_{W1W2} and ϵ_{W3W4} for well-observed asteroids. We discuss these constraints in more detail in the next section. We note that asteroids with sizes estimated by other means (e.g., from occultation or radar measurements) represent an invaluable sample for constraining the behavior of $\epsilon(\lambda)$ among the asteroid population.

2.2.2 Model Implementation: ATM

We now briefly describe the ATM code architecture and highlight some aspects of its functionality. ATM is built on two bespoke Python classes: the 'observatory' and 'model' classes. The observatory class is used to describe an observatory's filter throughput curves as a function of wavelength. In the case of WISE, these are the modified quadrature formulae provided by Wright (2013). The model class has placeholder functions for how a thermal model describes the surface temperature distribution on a model asteroid, how the observing geometry is taken into account, and how the model integrates emitted flux over the surface of a model asteroid of unit diameter. The instances of the two classes, called objects, can then be passed as arguments with physical parameters (D, T_{ss}, ϵ, p, G) and observing parameters (α, r, Δ) to a series of functions. These functions calculate predicted in-band fluxes with and/ or without reflected sunlight for the given observatory. Additionally, the model class can be passed separately to a different series of functions to generate model spectral energy distributions (SEDs) at a range of desired wavelengths.

As indicated by Myhrvold (2018a), fitting for different desired parameters and integrating the emitted flux can be slow. To make modeling and subsequent fitting more computationally tractable, ATM comes with emitted flux lookup tables generated for each thermal model at each WISE quadrature wavelength and at a range of wavelengths at 1 μm intervals between 1 and 30 μm. The lookup tables are functions of a grid of T_{ss} values, ranging between 100 to 1200 K in steps of 0.5 K, and α values, ranging between 0 and π radians in steps of $\pi/360$. Each table therefore has ~800,000 saved integration evaluations. Flux values of phase angles and subsolar temperatures that do not fall on this grid are interpolated

between nearest neighbors using a bivariate spline approximation. Critically, interpolation is not conducted between wavelengths and so lookup tables need to be created for wavelengths not already included in the package.

To conduct MCMC sampling, ATM uses the `pymc3` (Salvatier et al., 2016) package. `pymc3` is a robust Python probabilistic programming and Bayesian inference package and it allows for multi-threaded sampling of the posterior; the ATM lookup table implementation in combination with `pymc3` allows 4000 samples to be extracted in approximately 40 s on a moderate CPU. Each sampling chain requires a single thread, hence, sampling the posterior with 10 chains would require 10 threads and, because `pymc3` supports parallelization, it takes the same amount of time to sample as a single chain. Given a set of asteroid flux measurements for either a single object or multiple objects, with assumptions on priors for fitting parameters, the user can fit for any combination of parameters as defined by Equation 2.2. Should the user decide to constrain emissivity as described in Section 2.2.1.5, this task too is made simple by the use of several keyword arguments in the fitting function that describe how emissivity and albedo should be constrained and calculated. We note that MCMC sampling is an inherently stochastic process and so best-fit values obtained may differ slightly between different runs (well within stated uncertainties) unless an explicit value is set for the random seed (which needs to be done for every chain or walker).

We refer the reader to the GitHub repository for further information on the modeling code. A list of the ATM's primary requirements[6] can be found in the repository. The Notebooks used to validate the code are located here: `https://github.com/moeyensj/atm_notebooks/tree/main/paper1/validation/`. The Notebooks used to analyze the WISE sample of data (presented in the following sections) are located here: `https://github.com/moeyensj/atm_notebooks/tree/main/paper1/analysis/`. A Notebook showing how different assumptions regarding emissivity and albedo affect the modeling and retrieval of physical parameters is located here: `https://github.com/moeyensj/atm_notebooks/`

[6]See https://github.com/moeyensj/atm/blob/main/requirements.txt.

tree/main/paper1/analysis/single_object_90367.ipynb.

2.2.3 ATM Validation

We validate and verify the ATM code in two steps. First, we generate synthetic fluxes from a hypothetical asteroid using given D, T_1, and $\epsilon(\lambda)$, and then use ATM to perform Markov Chain Monte Carlo (MCMC) fitting to recover the input parameters. This step serves as a validation of internal code consistency. Second, physical validation is performed using a few studies from the literature, where asteroid sizes are known independently from infrared flux measurements (e.g., from radar measurements and stellar occultation measurements).

2.2.3.1 Validation of the Internal Code Consistency

We generate synthetic fluxes in the four WISE bandpasses using an asteroid with $D = 1$ km, $T_1 = 422$ K, $\epsilon(\lambda) = 0.7$, and heliocentric distance of 3 *au*. The observatory is placed at 2 au from the asteroid. The synthetic fluxes generated using the NEATM model are shown in Figure 2.1. The four symbols with error bars correspond to the four in-band fluxes for the WISE observatory. As evident, the flux in the bluest band is dominated by the reflected flux component, while the two reddest bands are dominated by the emitted flux component. When fitting for D and T_1 using "observational" constraints in the four WISE bands, we add Gaussian scatter to each flux with a standard deviation of 20%. This scatter reflects the variability amplitude Σ from Equation 2.11.

The synthetic model flux, generated using NEATM, is fit using MCMC and three different thermal models implemented in ATM: NEATM, STM, and FRM. We use 20 chains, each with 3000 samples and 500 burn-in samples. Depending on the prior and the complexity of the posterior, it may take a number of steps or iterations before the MCMC sampler (or similar) approaches the vicinity of the maximum likelihood solution. To avoid biasing statistics with these samples, it is common practice to define a number of first samples to discard, these are known as burn-in samples. 3000 samples with 500 burn-in typically means 2500 effective samples of the posterior per chain, although it should be noted this depends

on implementation. The best-fit parameters obtained using the STM and FRM models are very similar, but not identical, to the parameters obtained with the correct NEATM model. The three best-fit SEDs for the fiducial validation case are shown in Figure 2.1.

The posterior probability density function (pdf) for the NEATM fitting case is shown in Figure 2.2. Note that the use of full multi-dimensional pdf for the analysis of constraints on the model parameters is superior to using the so-called "point estimates" obtained by maximum likelihood methods (e.g., the least squares method). A strong covariance between D and T_1 is clearly visible: a smaller D is compensated by higher T_1, with a bias (as well as scatter) in best-fit D twice as large as the corresponding T_1 bias (because the observed flux approximately scales with $D^2T_1^4$). The equivalent 1-σ uncertainties for D and T_1 are 9.1% and 5.4%, respectively. These uncertainty levels can be approximately explained from first principles: given the flux uncertainty of 20% and 4 observations (1 in each band), the overall flux normalization uncertainty is about 10% ($= 20\%/\sqrt{4}$). If T_1 were known, the best-fit uncertainty for D would be about 5%, and if D were known, the best-fit uncertainty for T_1 would be about 2.5%. Given the covariance between best-fit D and T_1, their expected uncertainties are about twice as large (see Figure 2.2), in good agreement with the MCMC results.

When the number of epochs N is increased, we find that the fitting uncertainty for the best-fit parameter decreases as $N^{-1/2}$, as expected. The fitting uncertainty increases approximately linearly with the assumed flux uncertainty ($\Sigma = 0.2$ mag, presumably due to variability, rather than due to measurement uncertainties). For example, with $\Sigma = 0.1$ mag and 25 observing epochs (each with four WISE fluxes), the fitting uncertainty for the best-fit parameters D and T_1 decreases by about a factor of ten, to 0.8% and 0.5%, respectively. Of course, here we are using the same model for fitting as was used to generate the synthetic flux – in reality, such small uncertainties are *essentially impossible* due to numerous systematic shortcomings of idealized thermal models (and uncertain $\epsilon(\lambda)$ – see below).

These tests assumed that the correct values of $\epsilon(\lambda)$ are known a priori. When we set ϵ_{W1W2} to an incorrect value of 0.8, instead of the true input value of 0.7, we found that the

best-fit value of D was biased by as much as 24%. Therefore, even though formal fitting precision can be high, the accuracy of fitted parameters can be significantly worse.

We investigated next how the precision of the best-fit D and T_1 decreases when emissivity $\epsilon(\lambda)$ is also fit. Following discussion in Section 2.2.1.5, we fit ϵ_{W1W2}, the value of $\epsilon(\lambda)$ in the WISE bands 1 and 2, and assume that ϵ_{W3W4}, the value of $\epsilon(\lambda)$ in the WISE bands 3 and 4, is known. In other words, we assume a flat prior in the range 0 to 1 for ϵ_{W1W2}, and a Dirac δ function for ϵ_{W3W4}, centered on its true value. The posterior pdf is shown in Figure 2.3. The differences compared to the posterior pdf in Figure 2.2 are substantial; for example, the equivalent 1-σ uncertainties for D and T_1 increased by about a factor of three: from 9.1% and 5.4% to 31% and 13%, respectively. In addition to the covariance between D and T_1, both parameters show covariances with ϵ_{W1W2}. A larger ϵ_{W1W2} can be compensated by a lower T_1 or a larger D. For example, a 0.1 uncertainty in ϵ_{W1W2} can induce a $\sim$25% uncertainty in D! Given that typically $\epsilon(\lambda)$ is not known to better than 0.05, it follows that uncertainty of best-fit D is unlikely below $\sim$10% (in agreement with recent analysis by Masiero et al. (2018a); Wright et al. (2018)).

Compared to D and T_1, the posterior constraint on ϵ_{W1W2} is much weaker. The standard deviation for the marginal distribution of ϵ_{W1W2} shown in Figure 2.3 is 0.23, but note that this marginal pdf is far from Gaussian. The fitting constraints on all three free parameters, and ϵ_{W1W2} in particular, improve as the number of data points increases and the scatter Σ decreases. Figure 2.4 shows the posterior pdf when synthetic data include 25 epochs (each with four WISE fluxes) instead of one in Figure 2.3, and with flux uncertainty of $\Sigma = 0.1$ mag, instead of 0.2 mag. As evident, the posterior constraints for all three fitted parameters are significantly improved. The standard deviations for the marginal distributions are 2.1% for D, 1.0% for T_1 and 0.011 for ϵ_{W1W2} (recall that in the case of fitting only D and T_1, their uncertainties were equal to 0.8% and 0.5%, respectively, using the same number of observations and the same value of Σ). Therefore, adding ϵ_{W1W2} as a free parameter results in about twice as large statistical uncertainties for the best-fit size D.

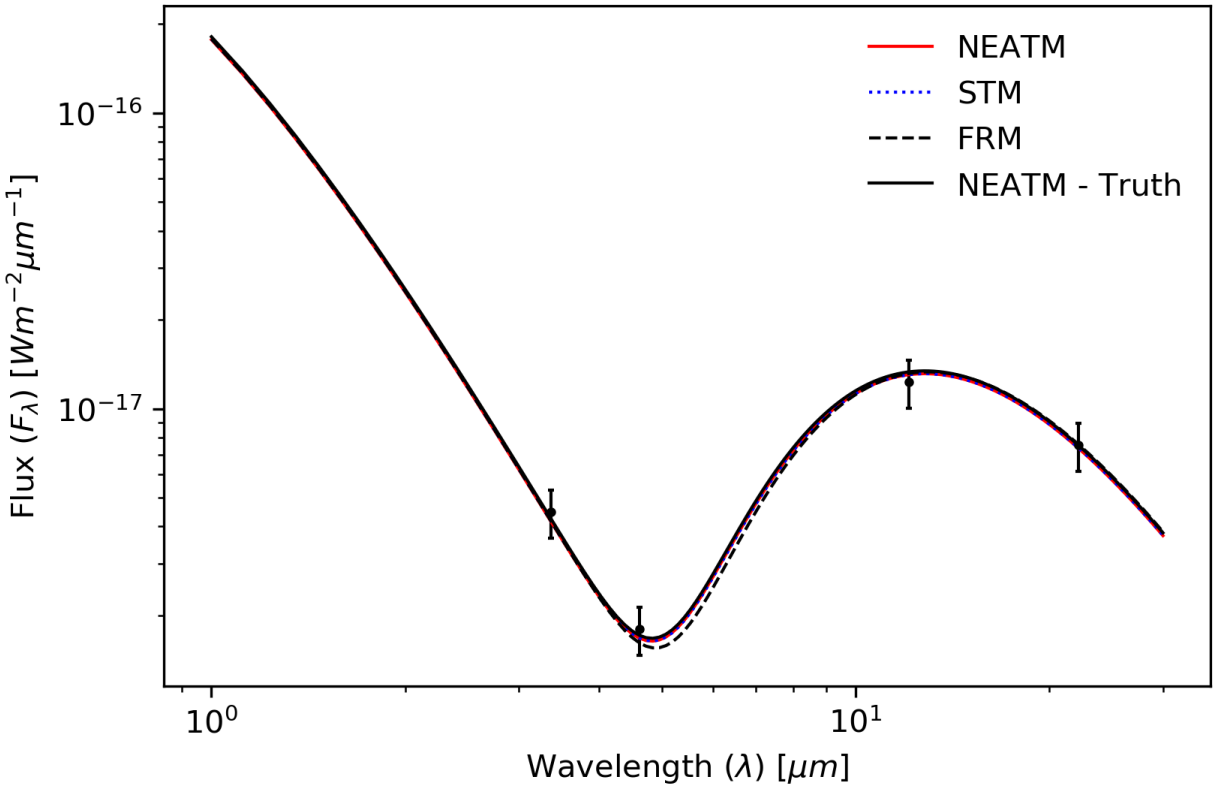

Figure 2.1 The solid line shows the synthetic spectral energy distribution (SED) generated using the NEATM model for an asteroid with $D = 1$ km, $T_1 = 422$ K, $\epsilon(\lambda) = 0.7$, heliocentric distance of 3 au, and an observer at 2 au from the asteroid. The four symbols with error bars correspond to the four in-band fluxes for the WISE observatory at the same distance from the asteroid. The error bars reflect the assumed variability amplitude but the implied scatter is not shown in this plot. The three other lines show the best fits obtained with ATM (see text). The two solid lines are indistinguishable. This figure was generated using example_synthetic_changingNumObservations.ipynb.

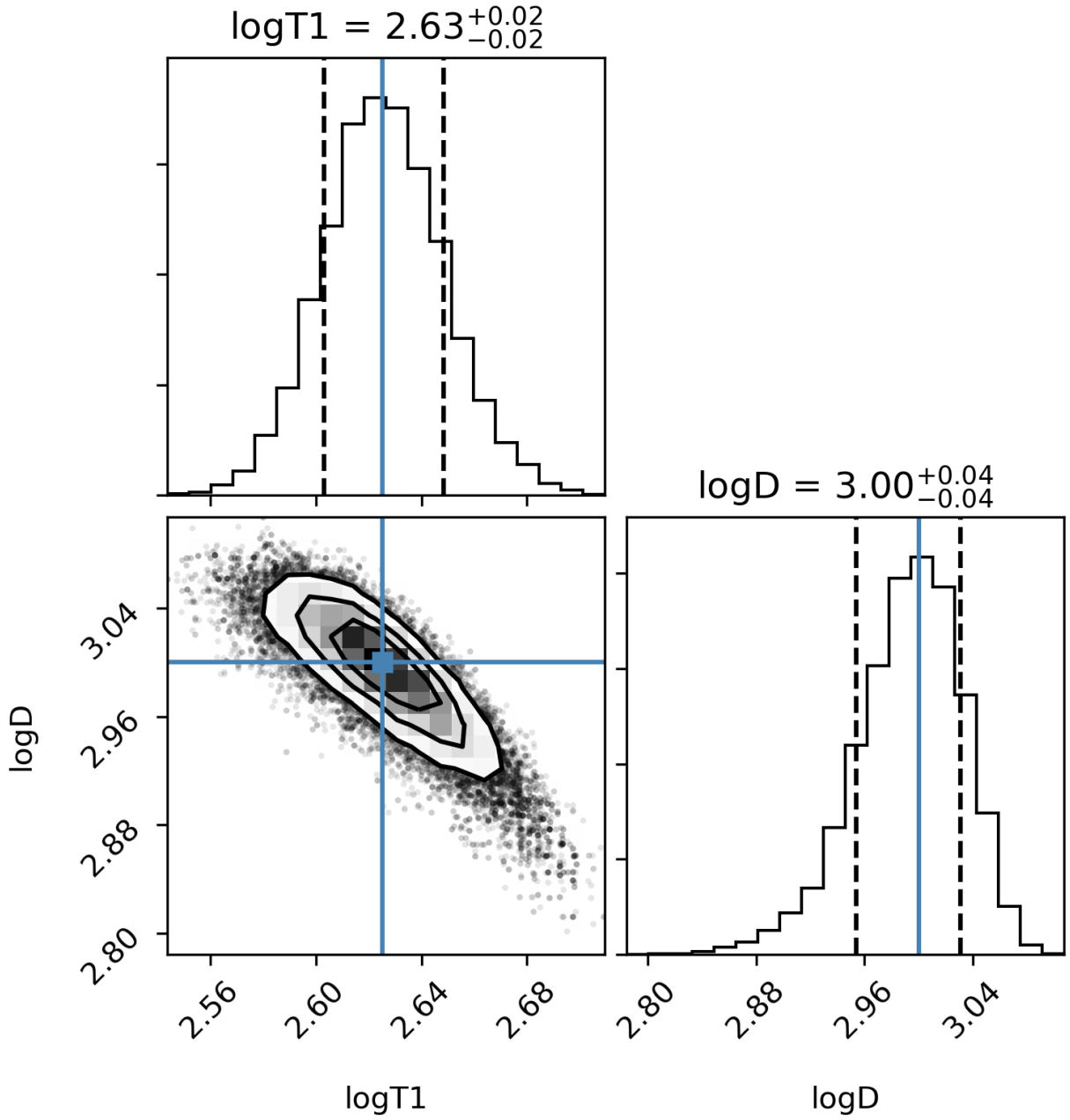

Figure 2.2 The posterior probability density function (pdf) for the case of fitting asteroid diameter D (in m) and temperature parameter T_1 (in Kelvin; see Equation 2.5). The equivalent 1-σ uncertainties for D and T_1 measured as standard deviation of the marginal probability distributions (shown as histograms; the vertical solid lines are true input values and dashed lines mark 1-σ uncertainties), are 9.1% and 5.4%, respectively. Note the strong covariance between D and T_1: a smaller D is compensated by higher T_1. The best-fit SED is shown in Figure 2.1. This figure was generated using example_synthetic_changingNumObservations.ipynb.

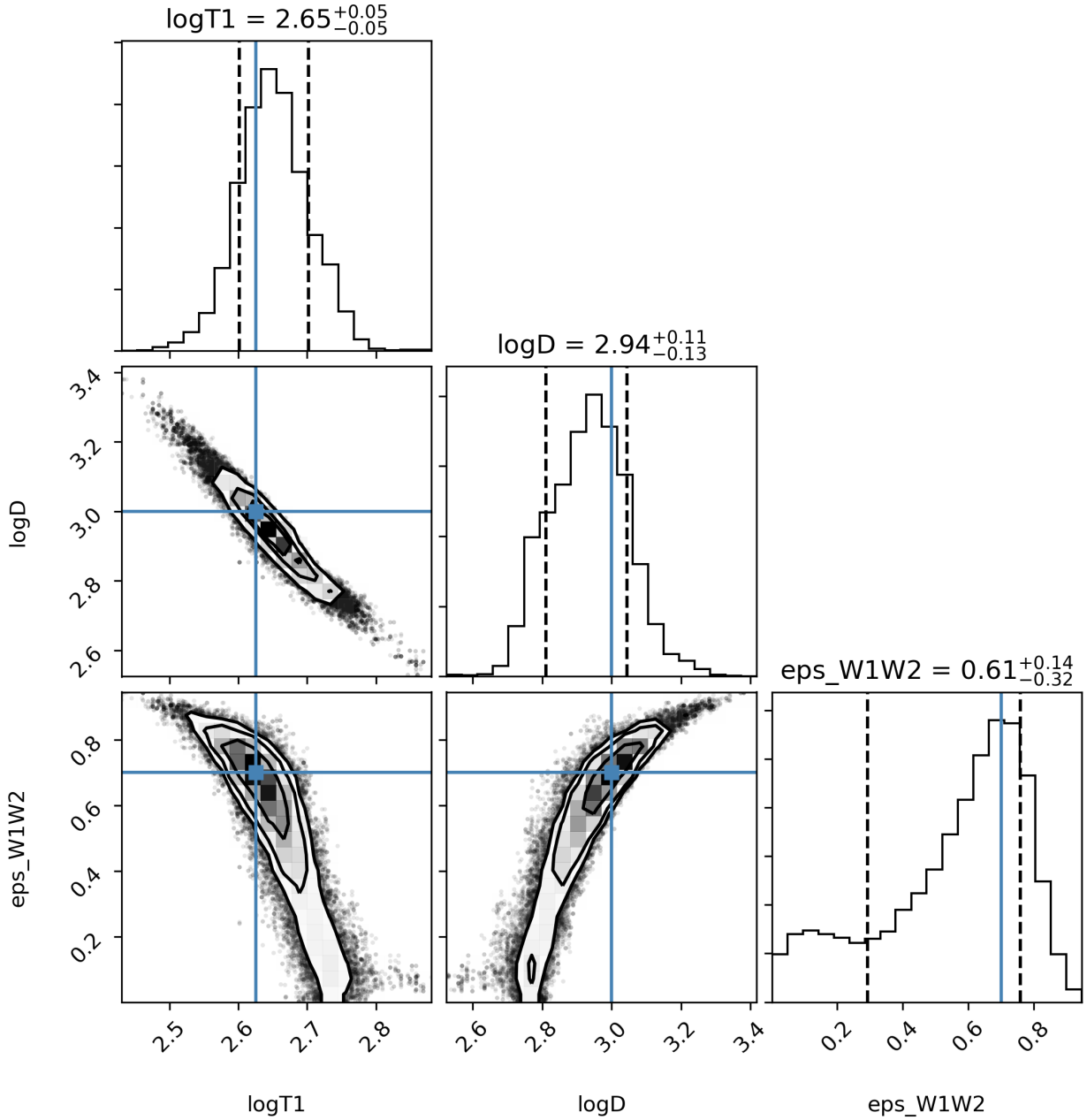

Figure 2.3 Analogous to Figure 2.2, but here for the case of fitting three free parameters: asteroid diameter D, temperature parameter T_1, and the value of $\epsilon(\lambda)$ in the WISE bands 1 and 2, ϵ_{W1W2}. Although all other simulation and fitting parameters are the same as is in Figure 2.2, the differences between the posterior pdfs are substantial. For example, the equivalent 1-σ uncertainties for D and T_1 increased from 9.1% and 5.0% to 31% and 13%, respectively. In addition to the covariance between D and T_1, both parameters show covariances with ϵ_{W1W2}. Compared to D and T_1, posterior constraints on ϵ_{W1W2} are much weaker (the 1-σ uncertainty for ϵ_{W1W2} is 34%!). This figure was generated using example_synthetic_changingNumObservations.ipynb.

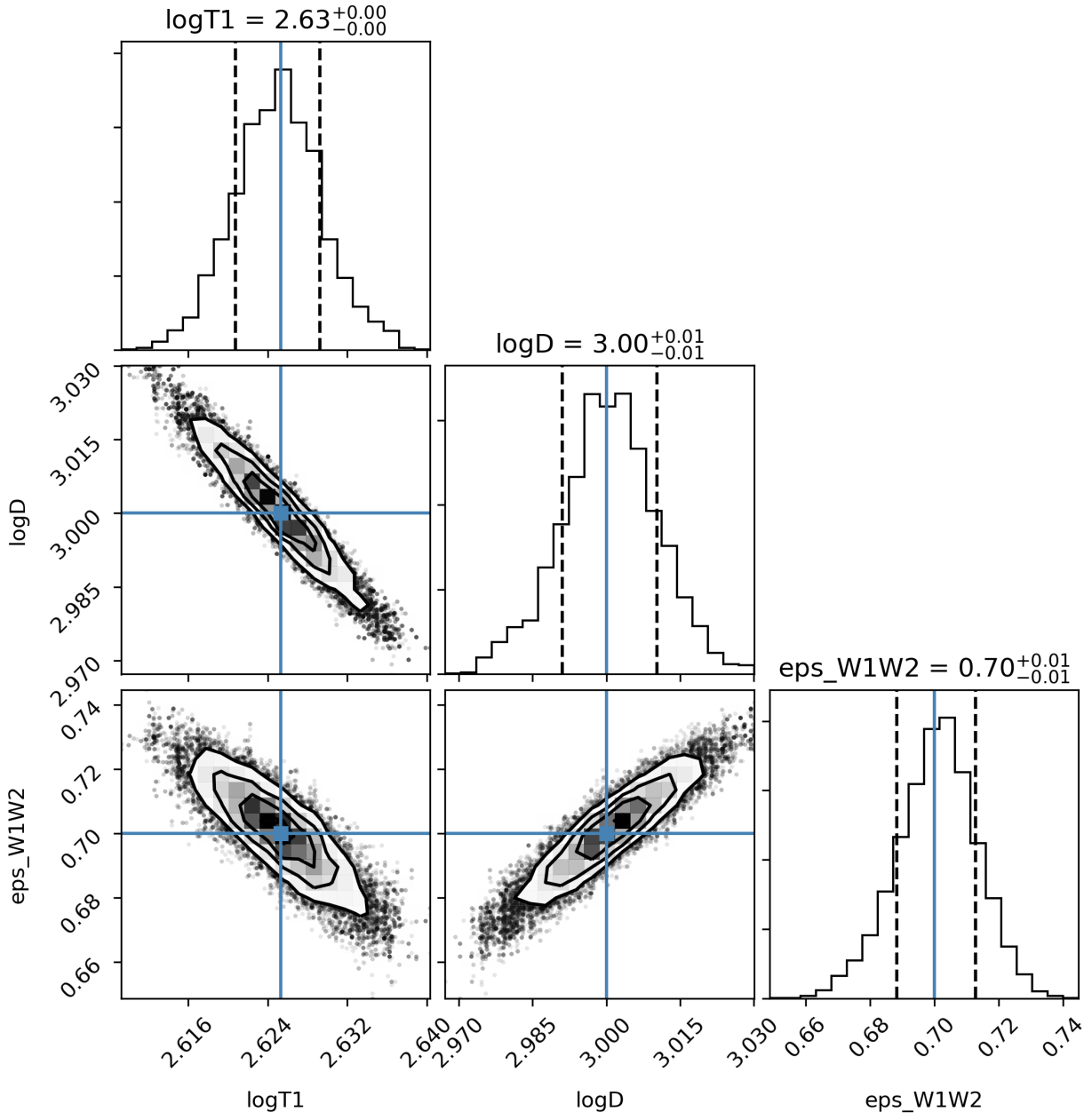

Figure 2.4 Analogous to Figure 2.3, but here with better observational constraints: 25 epochs (each with four WISE fluxes) instead of one, and with flux uncertainty Σ of 0.1 mag, instead of 0.2 mag. Note the significant improvement in best-fit constraints, as well as persistent covariance between the fitted parameters. This figure was generated using example_synthetic_changingNumObservations.ipynb.

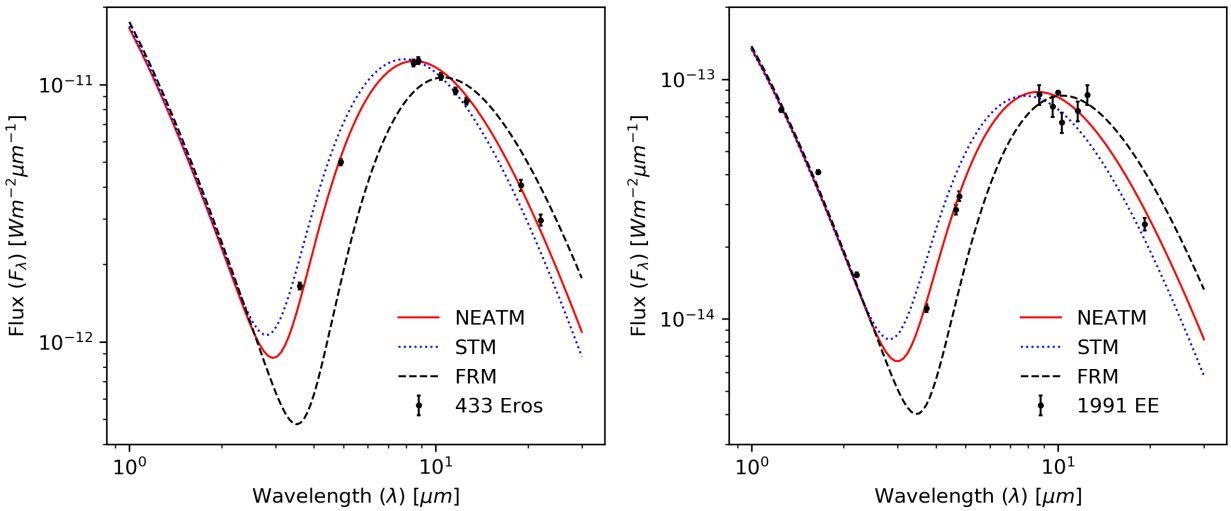

Figure 2.5 In the left panel, physical model validation using data for 433 Eros (symbols) from Lebofsky & Rieke (1979) and models evaluated using best-fit parameters from Harris (1998). This panel recreates Figure 1a from Harris (1998). In the right panel, we plot physical model validation using data for asteroid 1991 EE from Harris et al. (1998). This panel recreates Figure 3 from Harris et al. (1998) and was generated using example_1991EE&Eros.ipynb.

2.2.3.2 Physical Validation

The analysis described in the preceding section validated internal code consistency using synthetic data. While ATM can reproduce its own inputs, it still needs to be validated in an absolute sense. We now use several real asteroids, with sizes known independently from infrared flux measurements, to physically validate the code and implemented models.

First, we consider the observations of asteroid 433 Eros by Lebofsky & Rieke (1979). We use best-fit model parameters from Harris (1998) and demonstrate in Figure 2.5 that ATM produces model fluxes in agreement with measurements. An additional example, with data and analysis for near-Earth asteroid 1991 EE from Harris et al. (1998), also includes and validates the reflected light component, as shown in Figure 2.5. Finally, ATM is also validated against NEOWISE results, as described in detail in Section 2.3.

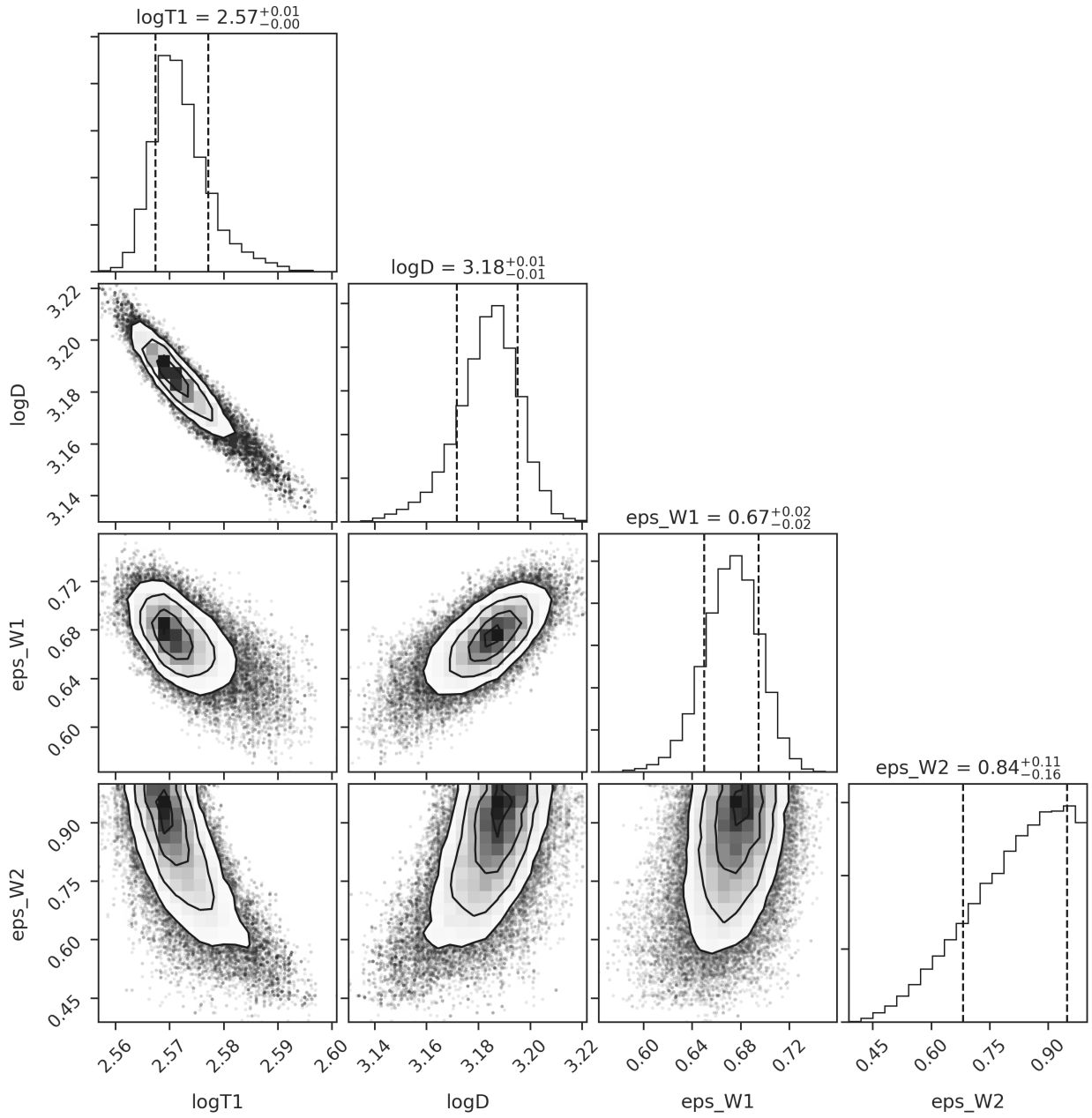

Figure 2.6 ATM "corner" plot for asteroid (54789) assuming model 5. Here both ϵ_{W1} and ϵ_{W2} are free parameters; compare to Figure 7 where an additional constraint $\epsilon_{W1} = \epsilon_{W2}$ is imposed. This figure was generated using single_object_54789.ipynb.

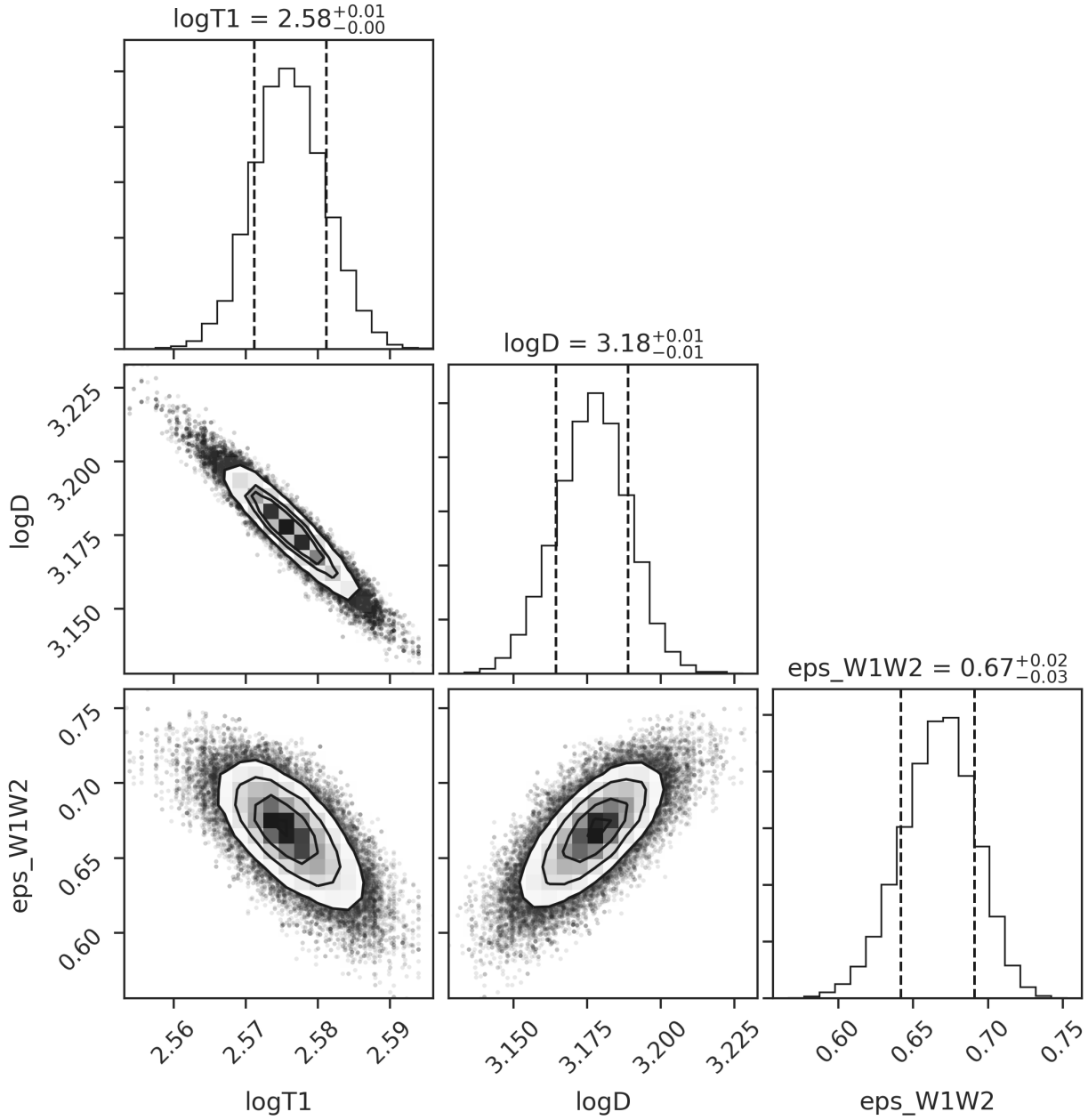

Figure 2.7 ATM "corner" plot for asteroid (54789) assuming model 4. This model imposes a constraint $\epsilon_{W1} = \epsilon_{W2}$. Compare to Figure 6 where both ϵ_{W1} and ϵ_{W2} are free parameters. This figure was generated using single_object_54789.ipynb.

2.2.4 An Illustration of ATM Fitting Capabilities: Different Emissivity Priors

The choice of Bayesian priors when fitting asteroid spectral energy distributions, such as priors for emissivity, can have a significant impact on the best-fit parameters. We illustrate ATM's capabilities for fitting asteroid spectral energy distributions, with emphasis on the treatment of Bayesian priors, using three well-observed asteroids: (25916), (54789), and (90367).

We study how the choices of priors affect the posterior pdf for fitted parameters and we study the resulting biases in point estimates derived from these pdfs (that is, "best-fit parameters"). We consider five different models, where in addition to fitting for asteroid diameter, D, and temperature parameter, T_1, we treat emissivity $\epsilon(\lambda)$ as follows:

- **Model 1**: $\epsilon(\lambda) = \epsilon_0 = \text{const.}$ and we fit for ϵ_0 using a flat prior $0 < \epsilon_0 < 1$.

- **Model 2**: We set $\epsilon_{W3} = \epsilon_{W4} = 0.9$ (that is, the prior is a Dirac δ function) and fit for unknown $\epsilon_{W1} = \epsilon_{W2} = \epsilon_{W1W2}$, where subscripts indicate WISE bands. Here, and in models below, we model $\epsilon(\lambda)$ as a step function, with transition wavelengths[7] between bands at $\lambda_{W1W2} = 3.9\ \mu m$, $\lambda_{W2W3} = 6.5\ \mu m$, and $\lambda_{W3W4} = 18.5\ \mu m$. Note that in SED plots discussed below emissivity ϵ_{W1} is extrapolated to wavelengths shorter than λ_{W1W2}, implying a preset value for the visual albedo, p_V. The same is true for emissivity in the most redward band, ϵ_{W4}; it is extrapolated to wavelengths beyond the W4 bandpass for plotting purposes.

- **Model 3**: We set $\epsilon_{W3} = 0.70$ and $\epsilon_{W3} = 0.86$, motivated by ensemble analysis presented later in Section 2.3, and fit for ϵ_{W1W2}.

- **Model 4**: We set $\epsilon_{W3} = 0.80$ and $\epsilon_{W3} = 0.98$ and fit for ϵ_{W1W2}. The $\epsilon_{W4}/\epsilon_{W3}$ ratio is the same as in the previous model.

---

[7] See http://www.astro.ucla.edu/ wright/WISE/passbands.html.

- **Model 5**: We set $\epsilon_{W3} = 0.80$ and $\epsilon_{W3} = 0.98$ and fit for ϵ_{W1} and ϵ_{W2}. Compared to the previous model, here we do not enforce $\epsilon_{W1} = \epsilon_{W2}$.

Figure 2.6 shows the posterior pdf for four fitted parameters in the case of model 5 and asteroid (54789). We have also considered a case where emissivity values in all four bands are free fitting parameters, for a total of six fitted parameters, but concluded that it is sufficient to start the discussion here with model 5. As the panels in the bottom row of Figure 2.6 show, the pdf for ϵ_{W2} is quite wide and not too dissimilar to its prior, while D, T_1 and ϵ_{W1} have pdfs that are much narrower than their priors, and with a well defined peak. In other words, ϵ_{W2} is much less constrained by the data than the other three parameters. This conclusion is also valid for the other two asteroids considered here.

Following the behavior of emissivity inferred from laboratory spectra for different materials, as illustrated in Figure 1 from Myhrvold (2018a), we enforce an additional fitting constraint: $\epsilon_{W1} = \epsilon_{W2} = \epsilon_{W1W2}$ and fit for three free parameters (models 2, 3, and 4). Figure 2.7 shows the posterior pdf for fitted parameters in the model 4 case. The fitted parameter ϵ_{W1W2} now has a pdf that is much narrower than its priors and with a well defined peak. The same behavior is observed for the other two asteroids. As a result of this analysis, we conclude that for the robust fitting of these and other less observed asteroids, we need to only fit for ϵ_{W1W2}, and not for ϵ_{W1} and ϵ_{W2} separately.

There is a remaining question of what to choose for the values of ϵ_{W3} and ϵ_{W4} that are not fitting parameters. As Table 2.1 shows, models 2, 3, and 4 result in values of best-fit D varying by about 10-15%, depending on the chosen values of ϵ_{W3} and ϵ_{W4}. In addition, adopting $\epsilon(\lambda) = \epsilon_0$ and fitting for ϵ_0 can result in a change of D as large as 16% (compare model 1 and model 2 for asteroid 54789). This approach also leads to a taxonomy-dependent D bias (systematic uncertainty) because the actual bias depends on detailed deviations of $\epsilon(\lambda)$ from the assumed constant value of ϵ_0. We will return to the discussion of optimal choice of ϵ_{W3} and ϵ_{W4} in Section 2.3.

Figure 2.8 compares spectral energy distributions for models 1–5 in the case of asteroid

(90367). As evident, all models are equally successful in explaining observed fluxes in bands W1 and W2. Models 3, 4 and 5 are more successful than models 1 and 2 in explaining observed fluxes in bands W3 and W4. This improvement is due to the choice $\epsilon_{W4} = 1.22\epsilon_{W3}$, instead of $\epsilon_{W4} = \epsilon_{W3}$ for models 1 and 2, and is discussed in detail in Section 2.3. This model degeneracy implies a bias in best-fit asteroid size in the range 10–20%, as discussed above.

Note that the prediction for reflected flux at wavelengths below $\sim$2 μm varies by about 30% due to variation of best-fit D and best-fit ϵ_{W1}, which implies an extrapolated value of p_V. Nevertheless, a measurement of flux at optical wavelengths could only break the model degeneracy if there were a strong prior reason to believe that emissivity values at optical wavelengths were somehow fully determined by the value of emissivity in the WISE W1 band.

2.3 Application of ATM to WISE Data and Comparison with NEOWISE Analysis

In this section, we apply ATM to a "gold" sample of $\sim$7000 best-observed asteroids by WISE. In addition to the best-fit diameter D and characteristic temperature T_1 for each object, we also obtain best-fit emissivity across bands W1 and W2, ϵ_{W1W2}. We compare ATM best-fit parameters to their values as published by the NEOWISE team, discuss physical implications of the best-fit parameters, and derive an approximate method for estimating D from WISE W3 measurements that doesn't require model fitting and is applicable to the majority of asteroids with WISE measurements.

2.3.1 Selection of High-quality WISE Data and Reliable ATM Fits

To compare ATM best-fit parameters with their values as published by the NEOWISE team, we select a relatively small subsample ($\sim$5% of the full sample) with the highest quality and quantity of WISE data. Observations of asteroids by WISE were obtained using the same criteria described by NEOWISE papers. The list of observations were obtained from the

Table 2.1 The Best-fit NEATM Parameters[a] for Asteroids (25916), (54789) and (90367)

Model	$\log(D/\mathrm{m})$	$\Delta D[\%]^b$	$\log(T_1/\mathrm{K})$	${p_{W1}}^c$	${p_{W2}}^c$	ϵ_{W1}	ϵ_{W2}	ϵ_{W3}	ϵ_{W4}
Asteroid (25916)									
1	3.690	0.00	2.587	0.322	$=p_{W1}$	0.876	$=\epsilon_{W1}$	$=\epsilon_{W1}$	$=\epsilon_{W1}$
2	3.682	-2.02	2.589	0.328	$=p_{W1}$	0.874	$=\epsilon_{W1}$	(0.90)	(0.90)
3	3.701	2.45	2.586	0.301	$=p_{W1}$	0.885	$=\epsilon_{W1}$	(0.70)	(0.86)
4	3.684	-1.56	2.589	0.321	$=p_{W1}$	0.877	$=\epsilon_{W1}$	(0.80)	(0.98)
5	3.662	-6.36	2.603	0.243	0.927	0.907	0.644	(0.80)	(0.98)
Asteroid (54789)									
1	3.240	0.00	2.558	0.675	$=p_{W1}$	0.741	$=\epsilon_{W1}$	$=\epsilon_{W1}$	$=\epsilon_{W1}$
2	3.166	-15.7	2.575	0.904	$=p_{W1}$	0.653	$=\epsilon_{W1}$	(0.90)	(0.90)
3	3.197	-9.49	2.571	0.795	$=p_{W1}$	0.695	$=\epsilon_{W1}$	(0.70)	(0.86)
4	3.177	-13.6	2.576	0.867	$=p_{W1}$	0.667	$=\epsilon_{W1}$	(0.80)	(0.98)
5	3.184	-12.1	2.571	0.852	0.431	0.673	0.834	(0.80)	(0.98)
Asteroid (90367)									
1	3.230	0.00	2.580	0.075	$=p_{W1}$	0.971	$=\epsilon_{W1}$	$=\epsilon_{W1}$	$=\epsilon_{W1}$
2	3.254	5.74	2.576	0.072	$=p_{W1}$	0.972	$=\epsilon_{W1}$	(0.90)	(0.90)
3	3.278	11.6	2.572	0.067	$=p_{W1}$	0.974	$=\epsilon_{W1}$	(0.70)	(0.86)
4	3.261	7.49	2.575	0.066	$=p_{W1}$	0.975	$=\epsilon_{W1}$	(0.80)	(0.98)
5	3.261	7.36	2.575	0.065	0.118	0.975	0.955	(0.80)	(0.98)

(a) The values in parenthesis represent priors.

(b) Change in diameter is measured relative to Model 1 diameters.

(c) Geometric albedo is calculated using Equation 2.8 and is stated for convenience.

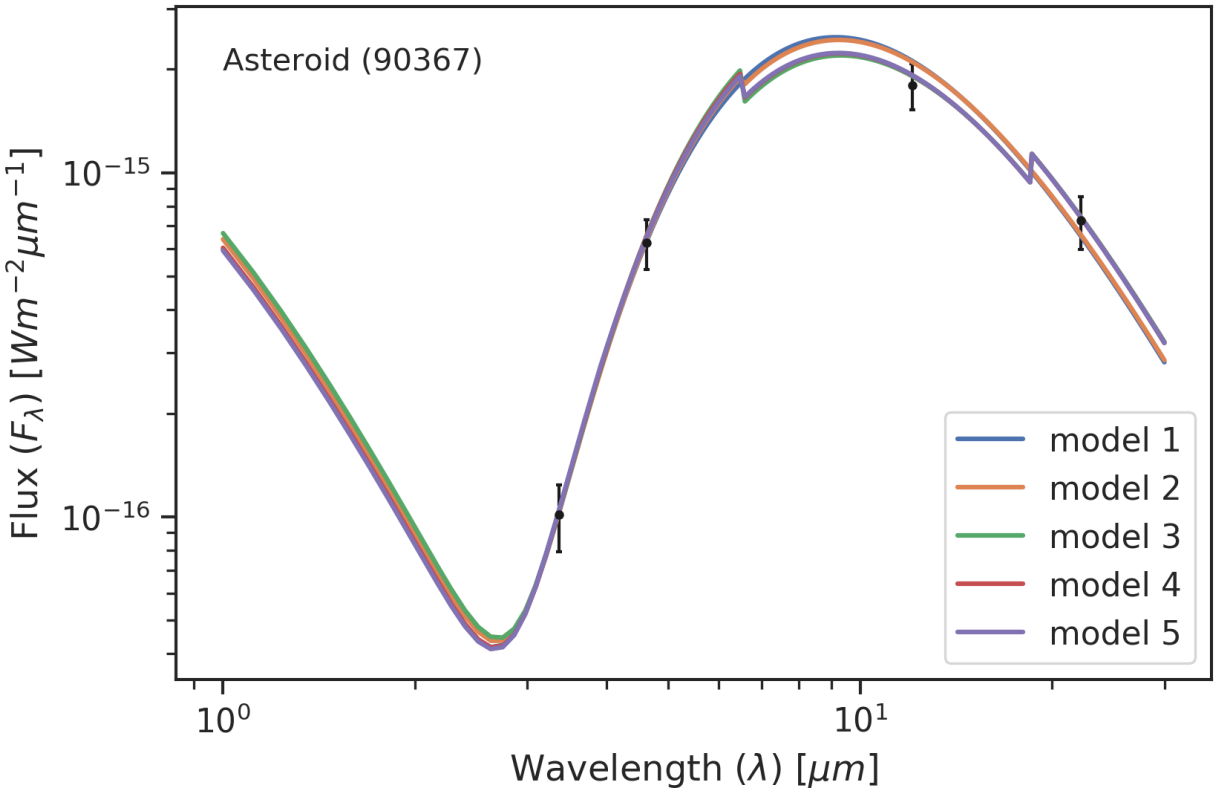

Figure 2.8 The symbols show the median fluxes and their errors for WISE data for asteroid (90367). The errors include the contribution of assumed variability amplitude of 0.15 mag. The solid lines show five NEATM models that differ in chosen priors for emissivity (see text). The best-fit parameters for these models are listed in Table 2.1. The sharp changes in modeled fluxes at $\lambda_{W2W3} = 6.5\ \mu m$ and $\lambda_{W3W4} = 18.5\ \mu m$ is due to a simplistic approximation of emissivity as a function of wavelength as a step function. This figure was generated using single_object_90367.ipynb.

Minor Planet Center database for observatory code C51. The positions and times were then used to search the WISE All-Sky Single Exposure Level 1b Source catalog using a cone search with radius 10 arcsec, and an observation time tolerance of ±2 s. The distance from the asteroid to the Sun, distance from the asteroid to WISE, and the phase angle were obtained with the JPL HORIZONS service[8], using the asteroid designation and observation time. To ensure high quality measurements, we require[9]:

- at least 3 observations in each band with photometric signal-to-noise ratio of at least 4, and

- artifact flag = 0 and quality flag = A, B, C, but not U or X.

The quality flags indicate to what signal-to-noise ratio a source was detected in each band. For example, a quality flag of A indicates SNR $>$ 10, while a quality flag of C indicates 2 $<$ SNR $<$ 3, while U or X indicate an upper limit on magnitude or a non-detection, respectively. This set of criteria reduces ∼10 million observations of ∼150,000 asteroids to ∼350,000 observations of 9672 asteroids.

For some objects, measurements still include outliers despite the above quality cuts. We clip outliers using a simple automated iterative algorithm. For each asteroid, using only the observations that are within 1 mag of the median magnitude a linear time versus magnitude model is fit for each band. Once the models have been fit, all observations that are more than 1 mag away from the best-fit model in each band are flagged as magnitude outliers. The observations may be flagged regardless of if they were or were not used to build the linear model. This outlier clipping reduced the aforementioned 9672 asteroids down to 7363 (14% of observations at this stage were cut as magnitude outliers). Finally, 4 objects were removed as they had missing assumptions on slope parameter, G. The final "gold" sample

[8]See https://ssd.jpl.nasa.gov/horizons/.

[9]For a detailed discussion of WISE data products see http://wise2.ipac.caltech.edu/docs/release/prelim/expsup/sec2_2a.html.

contains ∼308,000 observations (∼77,000 four-band observations) of 7359 objects.[10]

For each object, we fit for diameter D, characteristic temperature T_1 and emissivity across the W1–W2 wavelength range, ϵ_{W1W2}. We produced four sets of fits that differ in priors for emissivity across the W3–W4 wavelength range. Following NEOWISE analysis, we adopted $\epsilon_{W3} = \epsilon_{W4} = 0.9$ for the first set. Informed by discrepancies between the observed and modeled distributions of objects in the W3–W4 vs. W2–W3 color-color diagram (see below for detailed discussion), for the other three sets we adopted $(\epsilon_{W3}, \epsilon_{W4}) = (0.80, 0.98)$, (0.76, 0.93) and (0.70, 0.86).

The top right panel in Figure 2.9 shows the distribution of χ^2 per degree of freedom as a function of the total number of data points in all 4 bands. The computation of χ^2, as well as fitting, assumes an intrinsic scatter due to variability of $\Sigma = 0.15$ mag (see Equation 2.11). The distribution of χ^2, with a mode at 0.88, validates this choice. We note that the mean χ^2 for the entire 7359 sample is 1.77, when cutting outliers with $\chi^2 < 6$ then the mean reduces to 1.11. There is no correlation of the χ^2 distribution with the number of data points; typically, there are ∼50 observations per object.

2.3.2 Comparison to NEOWISE Fits

To compare our best-fit parameters to those obtained by the NEOWISE team, we match objects from our dataset to objects from the 2016 Planetary Data System version of the NEOWISE Diameters and Albedos database[11] (Mainzer et al., 2016), hereafter "NEOWISE values". For this initial comparison, we also require NEOWISE fit code "DVBI" and at least three observations in each of the four WISE bands. A NEOWISE fit code of "DVBI" indicates that diameter (D), visual albedo (V), beaming parameter (B), and infrared albedo (I) were fit for. If a fit code contains a hyphen (-) instead of a letter, the parameter which the hyphen replaced was not fit for and instead had an assumed value. For example, the most numerous fit code with over 87,000 results, "DV- -", fits for diameter and visual albedo

[10] See https://github.com/moeyensj/atm_data/blob/main/paper1/sample.db.

[11] See https://sbn.psi.edu/pds/resource/neowisediam.html.

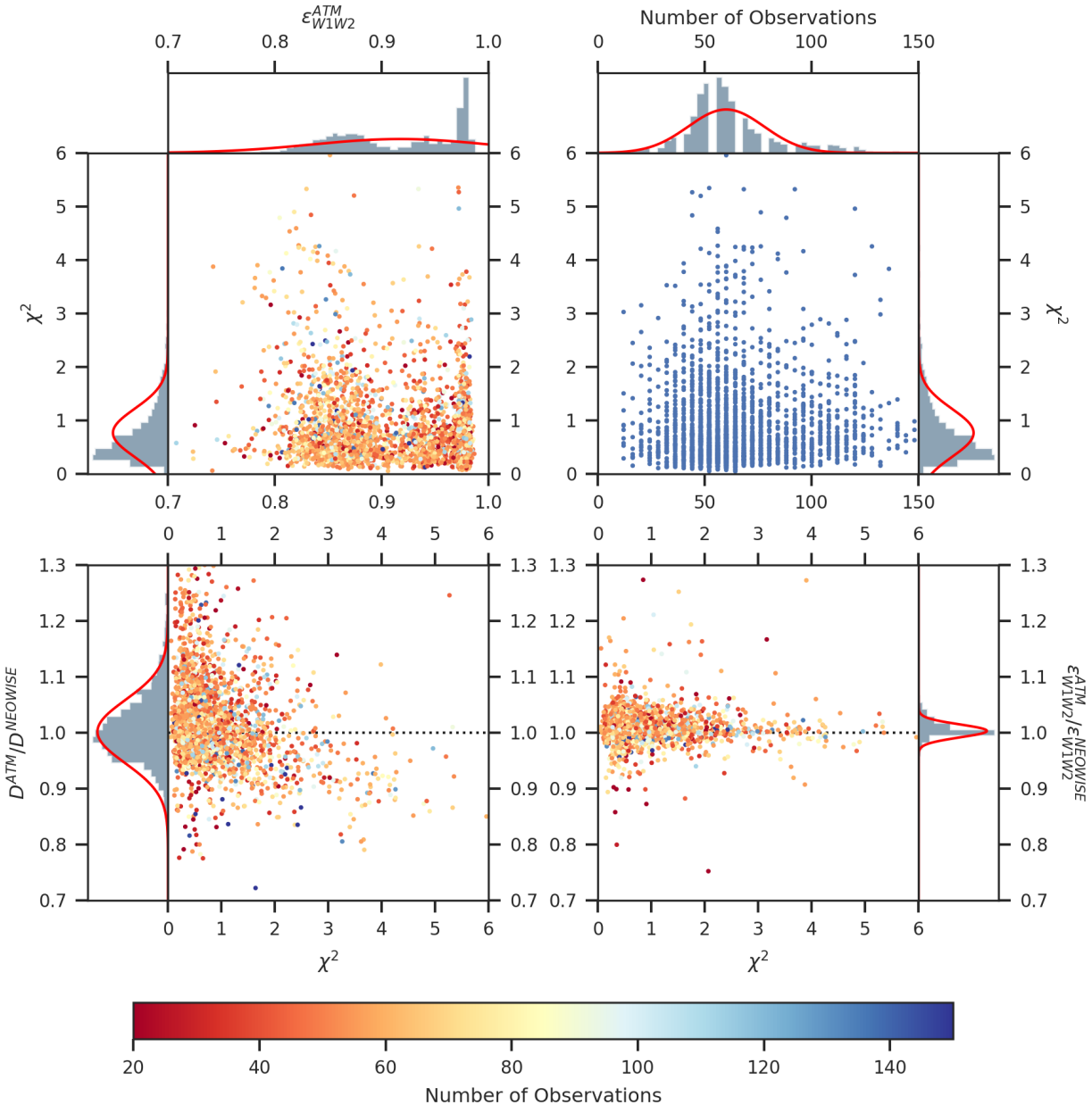

Figure 2.9 The top left panel shows the χ^2 per degree of freedom vs. ϵ_{W1W2} diagram for asteroids that pass initial quality cuts (see Section 2.3.1), have NEOWISE fit code "DVBI" and have at least 3 observations in each of the four WISE bands. The best-fit χ^2 and emissivity ϵ_{W1W2} were obtained using ATM. The symbols are color-coded using the total number of data points, according to the legend below the panel. The top right panel shows the χ^2 vs. the total number of data points diagram. The bottom panels show the ratios of the best-fit diameter (left) and ϵ_{W1W2} (right) obtained by ATM and the 2016 NEOWISE release values vs. χ^2, color-coded the same way as in the top left panel. This figure was generated using analysis.ipynb.

while beaming parameter and infrared albedo had assumed values. We limit our selection of fit codes to "DVBI" as this is the most NEATM-like fitting scenario and is the NEOWISE fit most similar to our modeling assumptions. As "DVBI" fits comprise only about 4600 of the total fits published by NEOWISE, crossmatching with this requirement immediately reduced the number of matches. The additional requirement of having at least three observations in each band further reduces number of matches as the vast majority of NEOWISE results did not use observations in all four bands. We also selected only the NEOWISE results for each asteroid that used the most data points as some asteroids had multiple fitting results under the same fit code. These criteria yielded a final comparison sample of 2807 asteroids. The ratio of best-fit ATM and NEOWISE values vs. χ^2 for diameter D and emissivity ϵ_{W1W2} is shown in the bottom panels in Figure 2.9. As evident, there is a good agreement between the two sets of best-fit parameters.

A more quantitative comparison is shown in the top two panels in Figure 2.10, for the fitting case with $\epsilon_{W3} = \epsilon_{W4} = 0.9$. Both the best-fit diameters and emissivity agree on average to within 0.4%, with a scatter of 5.5% for diameters and 1.1% for emissivity. Plausible reasons for the scatter include different outlier rejection algorithms, different treatments of Kirchhoff's law[12] (for a detailed discussion, see Myhrvold (2018a)) and how ATM accounts for intrinsic variability, although we cannot exclude other causes (e.g., a slight error in the quadrature formula for the W3 band, see Appendix A for details). Whatever the reason, the discrepancies are encouragingly small.

Our ATM results faithfully match the tri-modal distribution of ϵ_{W1W2} emissivity discovered by Masiero et al. (2014). The two bottom panels in Figure 2.10 show the resulting distributions of emissivity and corresponding albedo based on ATM results. The parameters of a best-fit 3-component Gaussian mixture model are listed in Table 2.2. Note that the

[12]We have run ATM with a fiducial dataset both ways: with Kirchhoff's law properly implemented and with the NEOWISE ansatz. The distribution of the ratio of best-fit sizes is centered on 1.003, with a scatter of 1.8%. In other words, although neglecting Kirchhoff's law is physically wrong, the resulting approximation does not contribute significantly to systematic size uncertainty when considering WISE dataset.

Table 2.2. Best-fit Parameters for Emissivity and Albedo Distributions[a]

Quantity	fraction$_1$	μ_1	σ_1	fraction$_2$	μ_2	σ_2	fraction$_3$	μ_3	σ_3
Emissivity ϵ_{W1W2}	0.299	0.977	0.004	0.195	0.945	0.015	0.506	0.862	0.030
Albedo p_{W1W2}	0.296	0.060	0.010	0.203	0.142	0.039	0.501	0.358	0.077

[a]Best-fit 3-component Gaussian mixture shown in two bottom panels in Figure 2.10

fractions of three components listed in Table 2.2 (20%, 30% and 50%) are not corrected for the various sample selection effects.

2.3.3 Understanding ATM Best Fits

Despite good agreement between ATM and NEOWISE fitting results, there are statistical problems with best-fit models when $\epsilon_{W3} = \epsilon_{W4} = 0.9$: the magnitude residuals for all objects are offset from zero by 0.1-0.2 mag in most bands. These offsets indicate that the thermal models used are not fully capable of explaining WISE data. The distributions of observed objects in WISE color-color diagrams offer an efficient way to study possible causes of model deficiencies. We note that the use of a full sample to constrain priors for individual objects is known as Hierarchical Bayesian modeling (e.g., see Chapter 5 in Ivezić et al. (2014)).

For each object, we compute the median observed magnitudes in each WISE band and the median observed colors. The resulting color-color diagrams are shown in Figure 2.11. The position along the model sequences in the W3-W4 vs. W2-W3 color-color diagram (see the bottom right panel in Figure 2.11) is by and large controlled by the heliocentric distance r, while both r and emissivity ϵ_{W1W2} (or, equivalently, albedo) determine the position in the W2-W3 vs. W1-W2 color-color diagram (bottom left panel). Note the three clearly

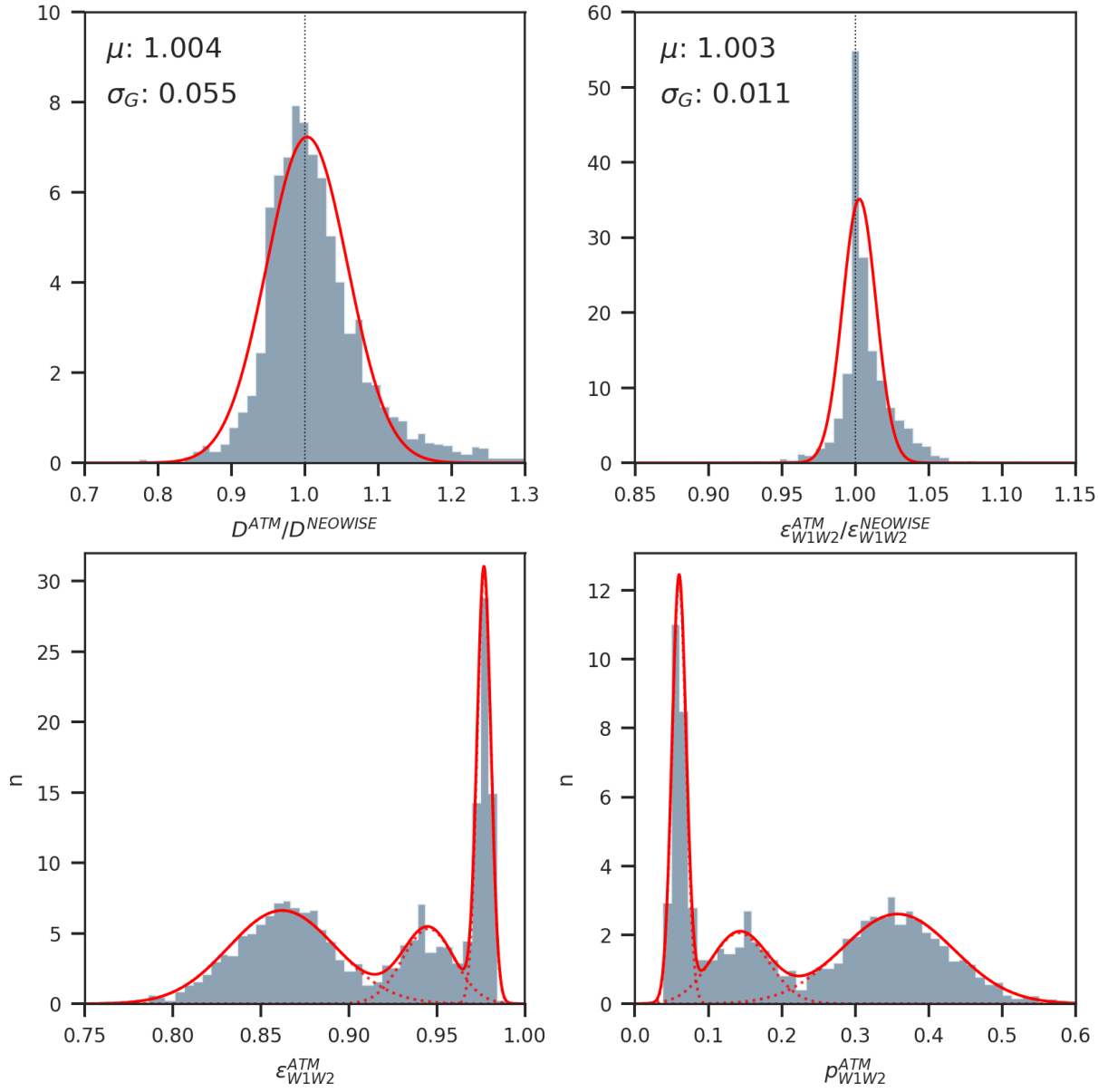

Figure 2.10 The top panels show histograms of the ratios of the best-fit diameter (left) and ϵ_{W1W2} (right) ATM and the 2016 NEOWISE release values, for 2656 asteroids with $\chi^2 < 3$ (ATM) and at least 28 observations. The red lines are single Gaussian fits, with their mean and standard deviation shown in each panel (e.g., ATM matches NEOWISE fits with a bias of 0.4% and a scatter of 5.5%, though note the distribution is leptokurtic). The bottom panels show distributions of emissivity ϵ_{W1W2} (left) and corresponding albedo (right) for the same objects. The solid lines show best-fit 3-component Gaussian mixtures (fit to individual data points rather than to histograms which are shown only for illustration). The best-fit parameters are listed in Table 2.2. This figure was generated using analysis.ipynb.

delineated observed sequences in this diagram. Their finite width is influenced by both the distribution of ϵ_{W1W2} values, and the distribution of observational phase angles (increase in ϵ_{W1W2} and/or phase angle moves model sequences from the bottom left to the top right). Note also that the W2–W3 and W3–W4 colors become redder as r increases, while the W1–W2 color becomes redder for decreasing r (because of the increasing relative contribution of the blue reflected component).

The placement of model tracks in the W3–W4 vs. W2–W3 color-color diagram is controlled by emissivity across the W3 and W4 wavelength range, parameterized as ϵ_{W3} and ϵ_{W4}, while the position along the tracks is controlled by the asteroid's temperature (itself controlled by T1 and the heliocentric distance). The three model tracks shown by dashed lines correspond to $\epsilon_{W3} = \epsilon_{W4} = 0.90$, and are strongly ruled out by the data: they don't overlap the majority of data points.

A small fraction of data points (about 3%) with very blue W3–W4 colors could be, at least in principle, either measurement outliers, or interesting objects with unusual ϵ_{W3} and ϵ_{W4} values. However, all of them are extremely bright in W3 (W3 < 1.8) and essentially represent the top 3% brightest sources. Hence, it is likely that their peculiar behavior in the W3–W4 vs. W2–W3 color-color diagram is simply due to overestimated W3 flux by about 0.5 mag. This conclusion is consistent with the recently reported flux corrections for saturation in the W3 band by Wright et al. (2018).

The only way to move model tracks to reach the data distribution is to assume $\epsilon_{W4} > \epsilon_{W3}$. After some experimentation, we adopted $\epsilon_{W4} \approx 1.22\epsilon_{W3}$. The required scale of ϵ is only set when W1 data is added because the W1 band includes a significant contribution from the reflected light component (proportional to $(1-\epsilon)D^2$, rather than to ^{2} for thermal emission component, as discussed earlier).

We investigated three pairs of values: $(\epsilon_{W3}, \epsilon_{W4})$ = (0.80, 0.98), (0.76, 0.93) and (0.70, 0.86). All three pairs produce model tracks in Figure 2.11 that are in a much better agreement with data than $\epsilon_{W3} = \epsilon_{W4} = 0.90$. The model tracks could be moved further to completely overlap the data distribution by adopting $\epsilon_{W4} \approx 1.34\epsilon_{W3}$ (e.g., $\epsilon_{W3} = 0.73$, $\epsilon_{W4} = 0.98$). We

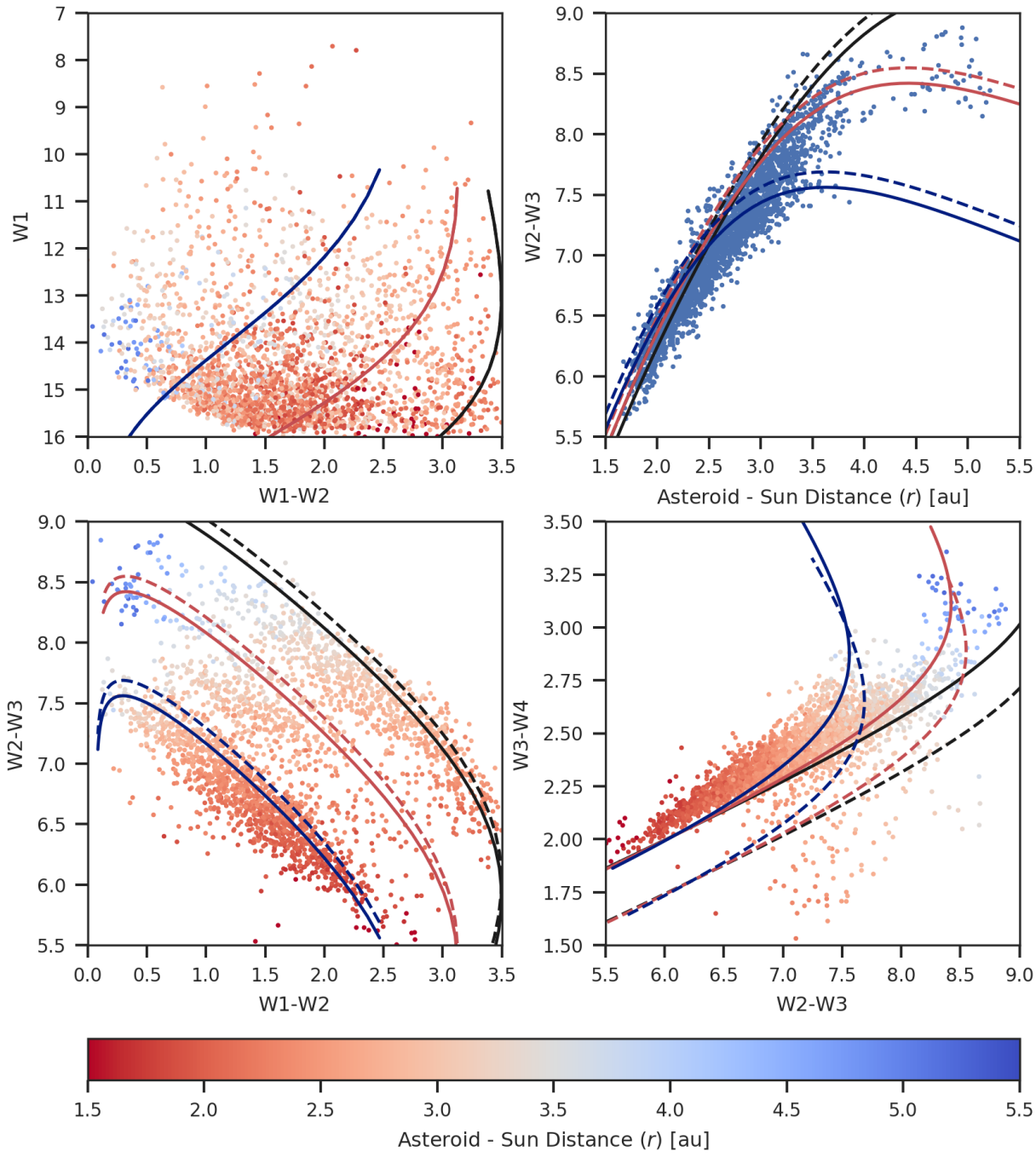
W1
W1-W2
W2-W3
Asteroid - Sun Distance (r) [au]
W2-W3
W1-W2
W3-W4
W2-W3
Asteroid - Sun Distance (r) [au]

Figure 2.11 The symbols correspond to 2656 asteroids with $\chi^2 < 3$ (ATM) and at least 28 WISE observations (the same sample as those plotted in Figure 2.10). Except for the top right panel, symbols are color-coded according to the median heliocentric distance, r. The top right panel shows the variation of W2–W3 color with r. The three solid lines in each panel are model tracks produced with three fixed pairs of characteristic temperature, T_1, and emissivity across the W1 and W2 wavelength range, ϵ_{W1W2} (0.98, 406 K; 0.95, 399 K; 0.86, 391 K; for black, red and blue tracks, respectively). The position along each track is controlled by the heliocentric distance (models are computed for the range 1.5–5.5 AU). Colors do not depend on the object size and distance from the observer; however, the position in the W1 vs. W1–W2 color-magnitude diagram (top left) depends on both of these quantities and model tracks assume a fiducial asteroid diameter of 10 km at a distance of 1 au. Emissivity across the W3 and W4 wavelength range, parameterized as ϵ_{W3} and ϵ_{W4}, controls the placement of model tracks in the W3–W4 vs. W2–W3 color-color diagram. The model tracks shown by solid lines are computed using $\epsilon_{W3} = 0.80$ and $\epsilon_{W4} = 0.98$. Model tracks shown by dashed lines correspond to the usually adopted values $\epsilon_{W3} = \epsilon_{W4} = 0.90$, and are strongly ruled out by the data distribution in the W3–W4 vs. W2–W3 color-color diagram. This figure was generated using analysis.ipynb.

leave detailed investigation of such models to future work.

While in all three cases the agreement between the data and the models in color-color diagrams is essentially unchanged, the values of best-fit diameters change in inverse proportion to the values of ϵ. Relative to the best-fit diameters for $(\epsilon_{W3}, \epsilon_{W4}) = (0.80, 0.98)$ case, which on average agree with the NEOWISE values to better than 1%, the sizes for (0.76, 0.93) case are on average 4% larger, and 10% larger for (0.70, 0.86) case. Note that best-fit diameters scale with ϵ faster ($\propto \epsilon^{-0.7}$) than the naively expected $1/\sqrt{\epsilon}$ due to non-linear fitting effects. We chose $(\epsilon_{W3}, \epsilon_{W4}) = (0.80, 0.98)$ for the rest of analysis here because of agreement with the NEOWISE values but we emphasize that these fiducial diameter values could easily have a systematic error exceeding 10% simply due to incorrect priors for ϵ_{W3} and ϵ_{W4}. It is plausible that these biases could be larger for the $\epsilon_{W4} \approx 1.34\epsilon_{W3}$ case.

This simple analysis illustrates how data can constrain model parameters, and also justifies our updated choice of priors for ϵ_{W3} and ϵ_{W4}. We now proceed with the analysis of best-fit model parameter distributions, their relationship to data properties, and derive an approximate size estimator that can be applied to the majority of objects in WISE sample.

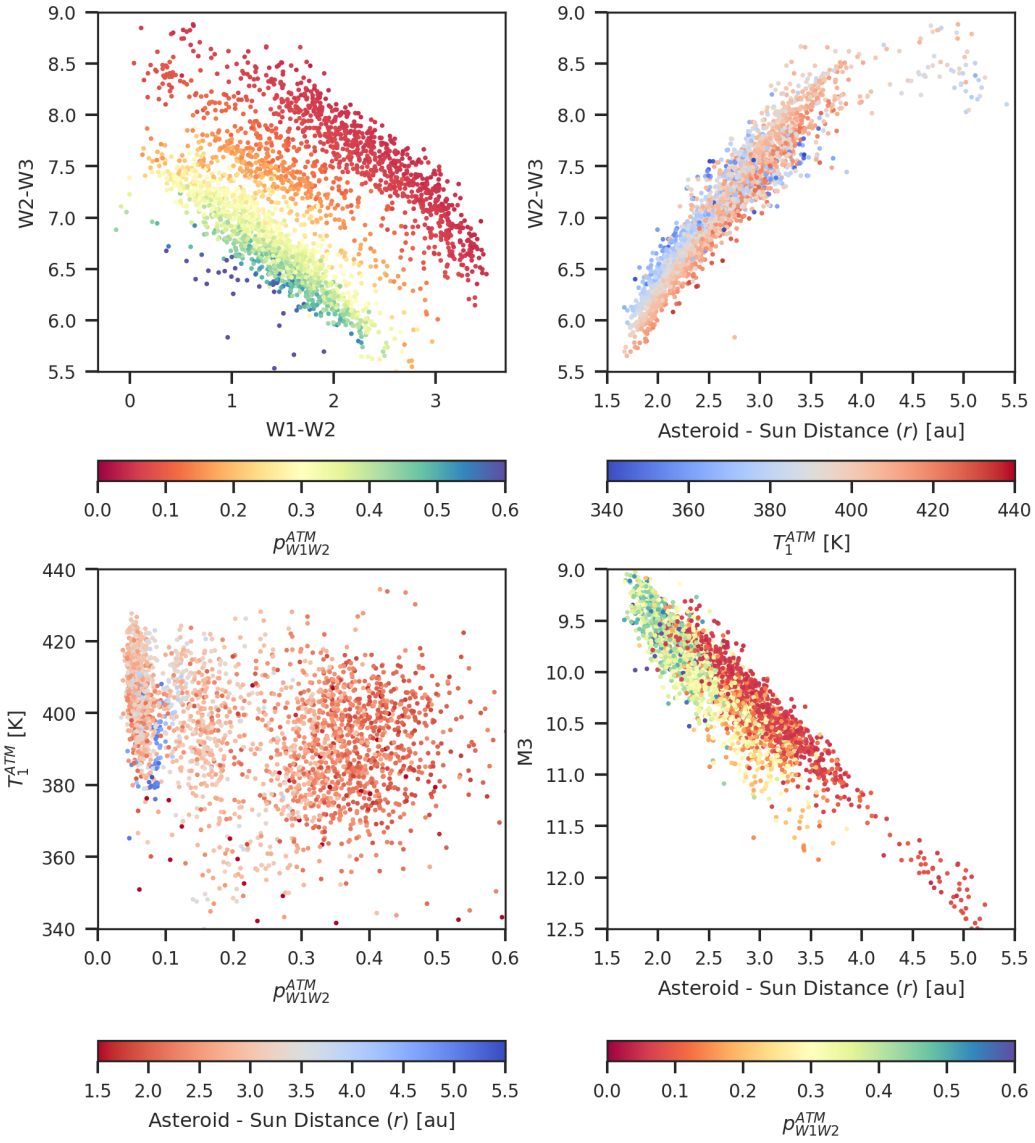

Figure 2.12 The symbols correspond to 2707 asteroids with $\chi^2 < 3$ (ATM) and at least 4 observations in each band. The top left panel shows the W2–W3 vs. W1–W2 color-color diagram, with symbols color-coded by the best-fit values of ϵ_{W1W2}. The three observed sequences clearly correspond to different taxonomic groups. The top right panel shows that the W2–W3 color is by and large controlled by the heliocentric distance, r, while characteristic temperature, T_1, has only a secondary influence due to its small dynamic range. The bottom left panel shows a weak correlation between albedo (or equivalently ϵ_{W1W2}) and T_1 (the albedo distribution is tri-modal, see the bottom right panel in Figure 2.10). The bottom right panel shows a correlation of "absolute" magnitude (corrected for the asteroid size and topocentric distance) with r. The preponderance of low-albedo objects at large r reflects the structure of the asteroid belt. This figure was generated using analysis_W3.ipynb.

2.3.4 An Approximate Single-band W3-based Size Estimator

The availability of WISE measurements varies greatly with bandpass and only a small minority of asteroids have data in all four bands. The majority of asteroids have data only in W3 and W4 bands, or in the W3 band alone. In fact, 45% of published NEOWISE fits (70,000+ out of ∼160,000) used only W3 band data, while only 9% of published fits didn't use any W3 band data. The NEOWISE team developed an elaborate set of different modeling approaches depending on which bands have data (for a concise summary, see Section 3 in Myhrvold (2018b)). Here we propose a much simpler two-step calibration and modeling scheme:

1. In the first step, size, temperature parameter and infrared emissivity ϵ_{W1W2} are fit for several thousand asteroids with WISE data in all four bands. The subset of these objects with direct size measurements is used to calibrate and validate NEATM model parameters and priors (not done here but implicitly used since the NEOWISE team validated their best-fit sizes using about 100 such objects, which we reproduce here with a subpercent bias).

2. In the second step, only W3 measurements are used to estimate the object's size, and the method is calibrated and validated using the four-band sample from the first step.

Hence, for studies requiring large samples of objects with size estimates (e.g., tens of thousands), the second step can provide a uniform dataset with well understood random and systematic size uncertainties from the 4-band sample, which is in turn calibrated using objects with direct size measurements.

The top left panel in Figure 2.12 shows the distribution of objects in the W2-W3 vs. W1-W2 color-color diagram, with symbols coded by the near-IR albedo derived from the best-fit emissivity ϵ_{W1W2}. We use the model with $\epsilon_{W3} = \epsilon_{W4} = 0.9$ because of direct comparison with NEOWISE results further below. As evident, the three sequences displayed by data distribution are related to albedo, and to its tri-modal distribution shown in Figure

2.10. As was demonstrated in Figure 2.11, the position along each sequence is by and large controlled by the asteroid temperature, which itself is controlled by the best-fit T_1 and heliocentric distance. These relations are illustrated in the top right panel in Figure 2.12. The distribution of best-fit T_1 is rather narrow and only weakly correlated with near-IR albedo, as shown in the bottom left panel in Figure 2.12.

As expected from Equation 2.5, T_1 does not strongly depend on heliocentric distance. Given this independence and a rather narrow distribution of T_1, it is possible to derive an approximate but relatively precise predictor for the asteroid flux, given the observing geometry (heliocentric distance, r, and topocentric distance, Δ). We first define a "pseudo-absolute" magnitude that accounts for the dependence of observed flux on asteroid size ($\propto D^2$) and its distance from the observer ($\propto \Delta^{-2}$),

$$M = W + 5\log\left(\frac{D}{\mathrm{km}}\right) - 5\log\left(\frac{\Delta}{\mathrm{au}}\right). \tag{2.13}$$

The variation of this "pseudo-absolute" magnitude constructed using WISE W3 band data with heliocentric distance is shown in the bottom right panel in Figure 2.12. At a given heliocentric distance, the distribution of M3 is rather narrow because both T_1 and ϵ_{W3} distributions are narrow.

A linear fit to M3 as a function of heliocentric distance,

$$M3 = 0.863\left(\frac{r}{\mathrm{au}}\right) + 7.859 \tag{2.14}$$

is adequate and matches observed values with a scatter of 0.21 mag, as shown in the top left panel in Figure 13. Using this fit and the definition of M3 given by Equation 2.13, we derive an approximate estimator for asteroid size,

$$5\log\left(\frac{D_{approx}}{\mathrm{km}}\right) = 7.859 + 0.863\left(\frac{r}{\mathrm{au}}\right) - W3 + 5\log\left(\frac{\Delta}{\mathrm{au}}\right). \tag{2.15}$$

As shown in the top right panel in Figure 2.13, this single-band estimator matches best-fit diameters based on data in all four bands with a scatter of only about 10%. This scatter is primarily due to scatter in unknown T_1 and ϵ_{W3} around their typical values implicitly

assumed in Equation 2.15. For the same reason, systematic errors are correlated with albedo, at about the 5% level as shown in the bottom left panel in Figure 2.13. We also find a bias of about 10% for the largest objects (bottom right panel).

The usefulness of the approximate size estimator given by Equation 2.15 is primarily that it can be applied to a much larger sample of objects than the 4-band sample discussed above. We computed the median W3 magnitude for 128,660 unique objects, that also have NEOWISE size estimates, using about 2.3 million W3 measurements. The formal size uncertainties based on scatter in observed W3 magnitudes and an intrinsic variability of 15% are in the range 2–6% for objects with W3 < 8 and about 10% at the sample faint limit (W3 ~ 10).

The two size estimators are compared in Figure 2.14, as a function of the NEOWISE fit code. They agree on average with a scatter of about 10%, and without appreciable biases for the high data quality "DVBI" subsample (as expected, as this subsample is closely related to the training sample). However, the two by far largest subsamples, "DVB-" and "DV- -" show biases exceeding 10%, and these biases have different behavior for different fit codes. Given that we apply a single estimator, it is likely, although not certain, that biases are introduced by the NEOWISE size estimator. As a possible clue, we find that the bias increases with the formal uncertainty in median W3 magnitude.

We note that because of a strong correlation between the median W3 flux and asteroid size (because of the finite dynamic range of observed distances), the biases with respect to the median W3 flux propagate to biases with respect to the object size and thus may introduce biases when estimating size distributions. Furthermore, WISE observed asteroids at a relatively small range of phase angles, and as stated in Section 2.2.1, the beaming parameter η can vary with phase angle. An increase in range of phase angles observed would create larger variations in scatter on characteristic parameter T_1 as defined by Equation 2.5 which to would contribute to biases on diameter.

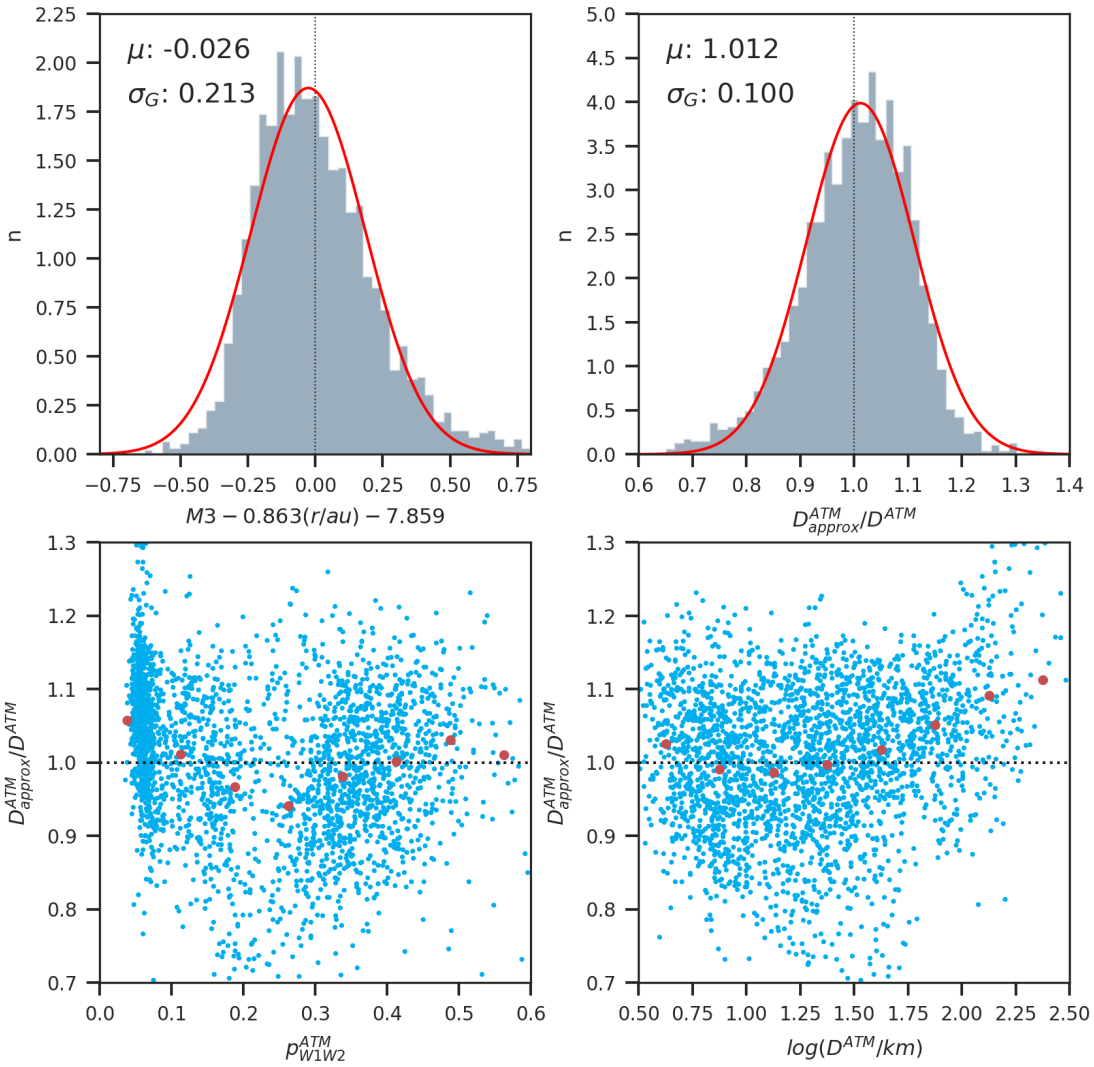

Figure 2.13 The top left panel shows a histogram of residuals after a straight line is fit to the "absolute" magnitude vs. the heliocentric distance relationship shown in the bottom right panel in Figure 2.12. The histogram in the top right panel shows the distribution of the ratio of an approximate asteroid diameter estimate based on W3 measurements alone and the best-fit ATM values based on all four WISE bands. The red lines are single Gaussian fits, with their mean and standard deviation shown in each panel. Note that single-band estimates match 4-band estimates with a scatter of 10% and a bias of 1% (this scatter is not dominated by the formal random uncertainties of single-band diameter estimates). The bottom two panels illustrate systematic uncertainty in this approximate estimate as a function of albedo and diameter. The large symbols are the median values for bins along the horizontal axis. This figure was generated using analysis_W3.ipynb.

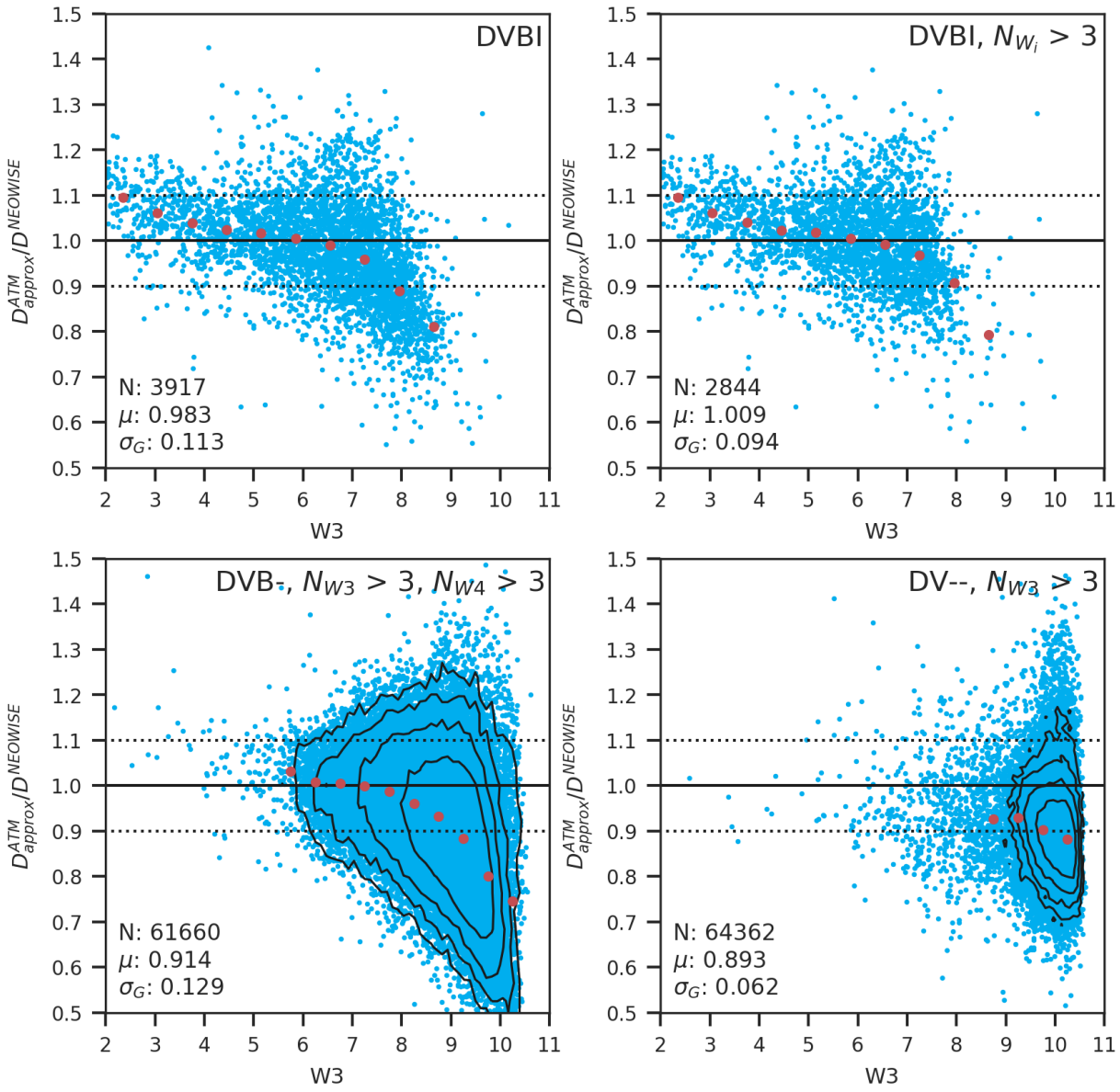

Figure 2.14 The panels show the ratio of W3-based size estimate and the 2016 NEOWISE estimate for subsamples of objects selected by their NEOWISE fitting code and the number of observations (symbols in top two panels and contours in bottom two panels due to much larger subsamples), as a function of the median observed W3 magnitude. The large symbols show the median values of size ratio in bins of W3. The three numbers in each panel list the subsample size, the median value of size ratio, and its robust (interquartile based) standard deviation. Top panels show "DVBI" sample, without (left) and with (right) a limit on the minimum number of observations in all four WISE bands. Bottom panels show "DVB-" subsample (left) and "DV- -" subsample (right), with a limit on the minimum number of observations as shown next to the label. This figure was generated using analysis_W3.ipynb.

2.4 Comparison of WISE-based Model Parameters and SDSS Data

In this section, we match the "gold" sample of 7359 best-observed asteroids from the WISE dataset (see Section 2.3.1 for selection criteria) to asteroids with optical observations listed in the Sloan Digital Sky Survey Moving Object Catalog (hereafter SDSSMOC; (Ivezić et al., 2001; Jurić et al., 2002; Parker et al., 2008)). Following Mainzer et al. (2012), we aim to study correlations between optical and infrared properties, such as colors and albedo. The presence of such correlations provides support that infrared emission models and best-fitting parameters are robust because the two datasets are essentially independent. We quantitatively explore the color vs. albedo correlation and develop an approximate method to estimate asteroid size from optical data alone. In addition, in Appendix B we discuss the selection of candidate M type asteroids using WISE-based best-fit parameters. Their optical color distribution is different than that for the whole sample and this fact further demonstrates that infrared best-fit model parameters are robust – if instead IR outliers were random measurement or modeling failures, their optical color distribution would not differ from that for the whole sample.

The 4th SDSSMOC release[13] lists astrometric and photometric data for 471,569 moving objects observed by SDSS prior to March 2007. Of those, 220,101 objects are linked to 104,449 unique objects with orbits. A match based on object designation to 7359 objects with WISE-based fits yields 1574 objects.

In the remainder of the analysis here, we use optical absolute magnitude, H, based on SDSS measurements (field 47 in SDSSMOC) because it represents an observationally uniform dataset, and because the values obtained from Minor Planet Center were found to have errors (both biases and random scatter) at the level of a few tenths of magnitude (for detailed discussion, see Section 2.3 in Parker et al. (2008)). However, we note that the net offset between the SDSS and MPC values of H, reported by Parker et al. (2008), disappears when using the June 2018 version of the MPC catalog MPCORB.DAT, and the scatter is reduced

[13]The 4th Release of SDSSMOC is available from http://faculty.washington.edu/ivezic/sdssmoc/sdssmoc.html.

to 0.22 mag.

2.4.1 Estimates of Optical Albedo from WISE-based Best-fit Sizes

Using SDSS-based absolute magnitude H and WISE-based best-fit size D, we estimate optical albedo using Equation 2.7. This estimate implies that p_V is a free fitting parameter whose prior is decoupled from emissivity and albedo values at IR wavelengths probed by WISE. If some prior data implied a strong emissivity/albedo relationship across the entire probed wavelength range (e.g., when assuming a constant unknown value of emissivity ϵ), then a joint fit would be more appropriate and D would be constrained by both optical and infrared data. Given that our knowledge of the emissivity vs. wavelength curve for individual objects is usually poor, and that observed emissivity values span a much smaller dynamic range than albedo values, it is better to first estimate size at wavelengths where thermal emission dominates total observed flux, and then use that best-fit size to estimate albedo at wavelengths were reflected light dominates total observed flux.

2.4.2 Optical Albedo-color Correlation

Following Figure 5 from Mainzer et al. (2012), the top two panels in Figure 2.15 show the SDSS a color vs. $i-z$ color-color diagram, with symbols color coded using the V band albedo from Equation 2.17 (left) and IR albedo derived from best-fit ϵ_{W1W2} (right). The a color is the first principal axis for the asteroid distribution in the SDSS $r-i$ vs. $g-r$ color-color diagram, defined as (Ivezić et al., 2001)

$$a = 0.89(g-r) + 0.45(r-i) - 0.57, \tag{2.16}$$

and enables easy separation of C type asteroids ($a < 0$) from S type asteroids ($a > 0$).

As discussed by Mainzer et al. (2012), the addition of IR albedo improves the definition of taxonomic regions, and presumably of subsets of asteroids with different optical albedo distributions, in the optical $r-i$ vs. $g-r$ color-color diagram. Motivated by the morphology of diagrams shown in the top two panels, in addition to the $a = 0$ separator, we add the line

Table 2.3. The Median and Robust Standard Deviation for Visual Albedo, p_V.

Color-selected sample	Median	St.dev.[a]
$a < 0$	0.052	0.016
$a > 0$ & $i - z > a - 0.05$	0.076	0.060
$a > 0$ & $i - z < a - 0.05$	0.203	0.060

[a]The robust standard deviation is estimated using interquartile range.

$i - z = a - 0.05$ and separate the sample intro three subsamples. The median optical albedo for these subsamples is listed in the top left panel in Figure 2.15. The medians and robust standard deviations are listed in Table 2.3.

2.4.3 An Approximate SDSS-based Size Estimator

Asteroid size can be estimated from optical data alone by transforming Equation 2.7 into

$$D = 1329\text{km}\frac{10^{-0.2H}}{\sqrt{p_V}} \tag{2.17}$$

and adopting a median albedo. Due to the large dynamic range of visual albedo, the scatter of such size estimates around the true values is large, about 50–60%, and non-Gaussian. However, the separation of asteroids using optical colors into three subclasses, each with a much narrower albedo distribution than for the whole sample, greatly improves such estimates. We use the following simple algorithm to assign p_V:

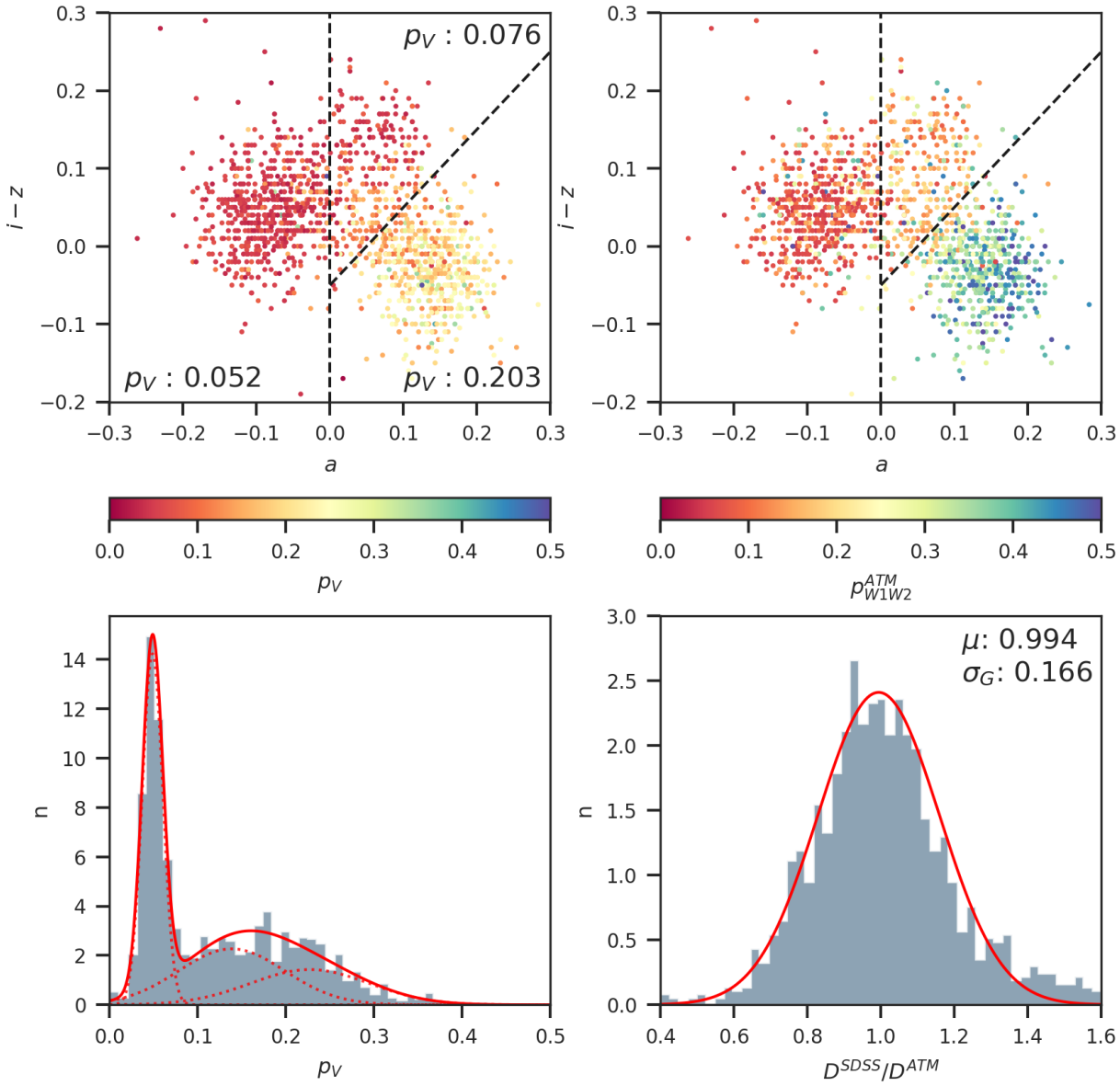
p_V : 0.076
p_V : 0.052
p_V : 0.203
$i-z$
a
p_V
p^{ATM}_{W1W2}
n
μ: 0.994
σ_G: 0.166
D^{SDSS}/D^{ATM}

Figure 2.15 The top panels show the a vs. $i-z$ SDSS color- color diagram, where a color is defined as $a = 0.89(g-r) + 0.45(r-i) - 0.57$ (Ivezić et al., 2001), for 1574 asteroids from SDSS MOC4 catalog that also have WISE-based diameters and IR albedo estimated with ATM. The symbols are color-coded by WISE-based IR albedo (right) and by V-band albedo (left) obtained using WISE-based diameter and SDSS-based absolute magnitude. The vertical dashed line shows the separation between C and S taxonomic classes from Ivezić et al. (2001), while the solid line is a separator between low-albedo and high-albedo objects derived here. The histogram in the bottom left panel shows the distribution of V-band albedo and the dashed lines show the best-fit 3-component Gaussian mixture (fit to individual data points). The darkest component is centered on $p_V = 0.052$. The histogram in the bottom right panel shows the distribution of the ratio of an approximate asteroid diameter estimate based on SDSS measurements alone and the best-fit ATM values based on all four WISE bands. The SDSS estimate is based on three color-assigned values of p_V, listed in the top left panel. The mean and standard deviation listed in the bottom right panel show that SDSS-based diameters match WISE-based best-fit ATM values with a bias of 0.6% and a scatter of 17%. This figure was generated using analysis_SDSS.ipynb.

$$
\begin{aligned}
p_V &= 0.052 \quad \text{if} \quad a < 0 \\
p_V &= 0.076 \quad \text{if} \quad a > 0 \quad \text{and} \quad i-z > a - 0.05 \\
p_V &= 0.203 \quad \text{if} \quad a > 0 \quad \text{and} \quad i-z < a - 0.05
\end{aligned}
\tag{2.18}
$$

As shown in the bottom right panel in Figure 2.15, such SDSS-based size estimates match WISE-based estimates with a scatter of only 17%, and a nearly Gaussian error distribution. For reference, single-band W3-based estimates have an intrinsic precision of about 10%, so the size estimates based on WISE W3 data are about twice as precise as the SDSS-based estimates (and have similar systematic errors because the latter are calibrated using the former). We note that the above scatter of 17% must be at least partially due to the variability-induced scatter in single-epoch SDSS estimates of H.

2.5 Discussions and Conclusions

Motivated by a desire to enable transparency and reproducibility of results, and to foster collaborative software development, we release ATM, a general tool for interpreting infrared

flux measurements of asteroids. With adequate infrared data, this tool can be used to estimate asteroid sizes, constrain asteroid emissivities at infrared wavelengths, and when optical data are available, also estimate visual albedos. The package also includes data files and example Jupyter Notebooks that can help significantly reduce the time to reproduce published results. All the analysis presented here, including all the figures, tables, and catalogs, can be easily reproduced with these Notebooks.

We emphasize that our analysis presented here, and corresponding catalogs with best-fit sizes and other parameters, are far from definitive and can be improved in various ways. For example, an analysis of emissivity in W1 and W2 bands using Hierarchical Bayesian modeling, similar to our analysis of emissivity in W3 and W4 bands (which could also be improved by optimizing the adopted $\epsilon_{W4}/\epsilon_{W3}$ ratio), would likely further decrease systematic uncertainties. As another example, modern machine learning methods could be used to improve our simplistic algorithm for assigning visual albedo using optical colors (Equation 2.18). We leave these improvements for future studies by us and the community.

Nevertheless, our results presented here already yield a number of useful conclusions. We show that ATM can match the best-fit size estimates for best-observed objects published in 2016 by the NEOWISE team with a sub-percent bias and a scatter of only 6%. Plausible reasons for this scatter include different outlier rejection algorithms, different treatments of Kirchhoff's law, and ATM accounting for intrinsic variability, although we cannot exclude other causes. Whatever the reason, the discrepancies are encouragingly small.

Our analysis of various sources of random and systematic size uncertainties show that for the majority of over 100,000 objects with WISE-based size estimates random uncertainties (precision) are about 10% (using W3-based estimates calibrated using high-quality 4-band subsample, see Section 2.3.4), and systematic uncertainties within the adopted model framework, such as NEATM, are in the range 10–20%. We estimate that the accuracy of WISE-based asteroid size estimates is in the range 15–20% for most objects, except for unknown errors due to an inadequate modeling framework (such as spherical asteroid approximation). Of course, there is no implied guarantee of Gaussianity and these statements

need to be interpreted with care. This result is consistent with the statement that accuracy is about 15% by Mainzer et al. (2016), but somewhat larger than the claim of "errors better than 10%" in Masiero et al. (2011).

The treatment of priors for emissivity $\epsilon(\lambda)$ has a direct and fundamental impact on biases in best-fit size estimates. We note that given this role of $\epsilon(\lambda)$, which is presumably shared by all members of an asteroid family, resulting systematic errors will be shared by all family members and thus cannot be seen when analyzing the per-family scatter in WISE-based optical albedo (Masiero et al., 2018a). Studies of such scatter are insensitive to systematic errors due to incorrect $\epsilon(\lambda)$ and thus cannot be used to constrain the absolute uncertainty, or accuracy, of flux-based size estimates. Only direct size measurements can enable a full understanding of the accuracy of size estimates based on thermal models.

Our ATM results faithfully recover the tri-modal distribution of ϵ_{W1W2} emissivity related to taxonomic classes discovered by Masiero et al. (2014). Correlations of SDSS colors and WISE-based best-fit model parameters indicate the robustness of the latter, and also give support to the claim that candidate metallic asteroids can be selected using best-fit temperature parameter and IR albedo (Harris & Drube, 2014). However, it should be noted that the condition on priors $\epsilon_{W1} = \epsilon_{W2} = \epsilon_{W1W2}$, introduced because WISE data do not strongly constrain ϵ_{W2}, may with k different from unity, as we did for ϵ_{W3} and ϵ_{W4}. We noticed some not be optimal. For example, we did not investigate ansatz $\epsilon_{W2} = k\epsilon_{W1}$, evidence for $k < 1$ when both ϵ_{W1} and ϵ_{W2} are free fitting parameters (see Section 2.4 and Table 2.1). Investigation of the optimal value of k using the full high-quality sample and Hierarchical Bayes methodology (as we did for the $\epsilon_{W4}/\epsilon_{W3}$ ratio, see Section 2.3) may shed new light on the behavior of infrared emissivity/albedo (especially if attempted for taxonomic subsamples defined by optical colors).

We utilized a correlation between SDSS optical colors and optical albedo derived using WISE-based size estimates and developed a method to estimate asteroid sizes with optical data alone, with an uncertainty of about 17% relative to WISE-based size estimates. When systematic errors are included, this small difference in accuracy between IR-based and optical

size estimates is further diminished. This remarkable result bodes well for future optical asteroid surveys, such as the Legacy Survey of Space and Time (LSST) (Ivezić et al., 2019), which might deliver such size estimates for over 5 million asteroids (Jones et al. 2018, and references therein).

Nevertheless, we point out that adequate infrared data are crucial for breaking the degeneracy between emissivity and asteroid size. An infrared survey with appropriately placed (at least) three bandpasses, and sensitivity to match the LSST sample, could provide a major breakthrough in our knowledge of the emissivity distribution for asteroid population. We also note that laboratory measurements of emissivity can greatly contribute to this endeavor by providing more robust priors for $\epsilon(\lambda)$. Last but not least, direct asteroid size measurements are of paramount importance for validating thermal asteroid models and quantitatively estimating their intrinsic biases, and they should be greatly encouraged and supported.

Chapter 3

TRACKLET-LESS DISCOVERY

3.1 Introduction

The number of Solar System minor planet discoveries is growing rapidly, thanks to the continuation of present-day surveys such as Pan-STARRS (Denneau et al., 2013) and the Catalina Sky Survey (Larson et al., 2003), as well as upcoming surveys such as the Vera C. Rubin Observatory's Legacy Survey of Space and Time (LSST) (Ivezić et al., 2019) and NEOCam, recently renamed to NEO Surveyor (Sonnett et al., 2020; Mainzer et al., 2021). In about a decade, the number of known objects will grow from the currently known one million to about six million minor planets. Such an increase in discoveries will enable a higher-resolution look into the dynamical evolution of our Solar System. However, identifying minor planets in survey images and linking their detections into orbits continues to be a challenging problem. First, linking asteroid detections across multiple nights is difficult due to the sheer number of possible linkages, and it is made even more challenging by the presence of false positives (Kubica et al., 2007; Denneau et al., 2013; Vereš & Chesley, 2017a,b; Jones et al., 2018). Second, the motion of the observer makes the linking problem nonlinear because minor planets will exhibit higher-order motion on the topocentric sky over the course of weeks (Holman et al., 2018). Finally, once potential linkages have been established, they need to be confirmed as possible orbits using computationally expensive orbit determination software. For example, the Vera C. Rubin Observatory estimates it will discover nearly six million Main Belt asteroids that will be observed hundreds of times over the course of its ten-year survey. Naively attempting to link hundreds of millions of asteroid detections over a ten-year period is not computationally feasible (Jones et al., 2018).

To make the linking problem more computationally tractable, surveys that aim to discover minor planets focus on constructing "tracklets": two-dimensional sky-plane motion vectors consisting of two or more detections spaced typically 20–90 minutes apart that constrain the direction and rate of motion of potential moving objects. Tracklets are constructed to reduce the number of possible linkages that could be formed by providing information on plausible direction and sky-plane angular velocity (Kubica et al., 2007). They are then linked into

internight linkages known as "tracks": sky-plane paths of motion containing several tracklets spanning up to ∼15 nights, typically modeled with low-order polynomials.[1] In the case of the LSST, for a moving object to be discoverable, it must be observed at least twice a night on at least three unique nights within the 15 day window to go through the tracklet-to-track creation process (Jones et al., 2018). In part due to the relative motion of the observer and the rate of motion of moving objects, both tracklets and tracks can exhibit high residuals relative to the fitted low-order polynomial, requiring relaxed fitting tolerances that can in turn lead to the creation of many spurious candidate linkages. Orbit determination (OD) algorithms are therefore required to run on each candidate track so spurious linkages can be identified and be removed.

The Zwicky Transient Facility (ZTF), an optical time-domain survey scanning the entire northern hemisphere of sky at a rate of more than 3700 $\mathrm{deg}^2\,\mathrm{hr}^{-1}$, can be seen as a precursor to the LSST (Graham et al., 2019; Bellm et al., 2019; Masci et al., 2019). ZTF uses the ZTF Moving Object Discovery Engine (ZMODE; Masci et al. 2019) algorithm. Instead of linking tracklets directly into tracks, ZMODE first attempts to build a "stringlet." A stringlet forms an intermediate step between tracklets and tracks that allows for the linking of pairs of detections across nights before tracks are built. This approach was designed to accommodate ZTF's cadence during its main survey, where the cadence is frequently too sparse to form short intranight tracklets.

Recent work by Holman et al. 2018 has shown promising results by shifting the reference frame for linking detections to the heliocenter. By assuming a heliocentric distance and its rate of change, cleverly fitting tracklets for the remaining unknown parameters in inertial space, then propagating the resulting "arrows" to a common epoch, arrows corresponding to the same minor planets will form clusters. These clusters can then be extracted and subsequently validated by orbit fitting. As a testament to the effectiveness of HelioLinC, some 200,000 new minor planet orbits were recovered from the Minor Planet Center's Isolated

[1] More modern variants such as HelioLinC (Holman et al., 2018) use transformations to the heliocentric coordinate system to linearize the linking problem.

Tracklet File (Holman et al., 2018).

Common in all of these approaches is the requirement to build tracklets, which in turn requires a telescope to perform multiple revisits to the same field in a night, then more revisits a few nights later, and so on. For a survey that cannot cover the entire visible sky twice per night, this leads to up to a factor-of-two reduction of the nightly surveyed area. For a survey such as the LSST, which aims to balance four different science drivers, requiring such a cadence decreases the overall ease with which the other science drivers can be accommodated. It is therefore prudent to investigate whether linking algorithms that are cadence-independent can be constructed and whether such algorithms can perform as well as – or better than – the current methods. An algorithm that does not demand a high revisit cadence could increase the efficiency of future surveys, as well as help multi-science missions such as the LSST.

In this paper, we present one such cadence- and observer-independent linking algorithm: "Tracklet-less Heliocentric Orbit Recovery" (THOR). Rather than shifting the origin to the heliocenter like in Holman et al. 2018, we choose to shift the linking frame of reference to a series of dynamically selected heliocentric "test orbits." The main insight is that transforming detections into the frame of a test orbit linearizes the motion of all objects in a relatively thick bundle of orbits near the test orbit (in phase space), which can then be picked out with line-detection algorithms such as the Hough transform. This provides a path to scanning an otherwise voluminous 6D phase space with a finite number of test orbits and at feasible computational cost.

We describe the THOR algorithm in Section 3.2. We test its performance on simulated and real survey observations in Sections 3.3 and 3.4, respectively. Finally, we discuss planned extensions and future work in Section 3.5.

3.2 Algorithm

Algorithms that rely on tracklets to discover moving objects generally search for the subset of orbits that are consistent with the observed tracklets and their constituent observations.

THOR takes a different approach – it aims to identify subsets of observations that are consistent with being close to one or more "test orbits" drawn from the 6D orbital phase space. For each such test orbit, THOR searches for detections that lie near or adjacent to it as it is propagated backward or forward in time using clustering and line-finding algorithms (see Section 3.2.1.4).

This approach is computationally feasible because a single test orbit can be used to recover all objects that exist in a finite (and relatively large) "bundle" around it in phase space. This makes it possible to select and explore a sizable volume of phase space with a finite number of test orbits. It also makes it possible to focus on regions of phase space where we expect to find objects or that we know are populated, which we take advantage of to computationally efficiently search for specific asteroid populations in ZTF (Section 3.4).

There are several advantages to this approach to minor planets searches. First, the propagation of the test orbit is only limited by the robustness of the used orbit propagator. For example, propagating the test orbit using an N-body integrator allows the full gravitational perturbation by the Sun, planets, and massive minor planets to be properly taken into account. This is particularly useful for NEOs or any minor planets that might have nontrivial orbital encounters. Second, a fully defined orbit can be propagated to epochs beyond the scope of any one survey, permitting the use of THOR on archival and combined data sets.[2] Third, by searching for observations near an assumed test orbit, the cadence of observations becomes irrelevant. The orbit can simply be propagated to any epoch where an observation is possible regardless of the temporal spacing between observations themselves. This removes the requirement for tracklets to be observed, and enables the linking of multiple single-night detections instead of the linking of multiple pairs of nightly detections.

The THOR approach does come with challenges. The computational cost, compared to HelioLinC-type algorithms, is higher due to the need to perform reference frame transformations to a (potentially large) number of test orbits rather than a single transformation to

[2]This also means that THOR is observer-independent, from the standpoint of linking detections from multiple observers.

the heliocenter. However, as we show in Section 3.4, this cost can be readily met in practice. A more challenging issue is that of the selection of test orbits to optimally cover the phase space of interest (the phase space where small bodies are likely to be found). While we do show, later in this section, that a relatively ad hoc approach to orbit selection (essentially, importance sampling from the known population) is a very effective way to scan for all objects in a known population, it is also clear this approach will leave the algorithm blind to unknown populations or objects on unusual orbits.[3] We leave the problem of developing an algorithm to select the optimal number of orbits to *exhaustively* search the entire phase space for future work.

3.2.1 THOR

We next describe the five main components of the THOR algorithm that enable the discovery of moving objects using a test orbit. We explain how test orbits are selected from 6D orbital phase space at the end of this section.

3.2.1.1 Select a Test Orbit

A test orbit with heliocentric position vector, $\vec{r}$, and heliocentric velocity vector, $\vec{v}$, at some epoch, t, is placed in a survey that contains the detections of moving objects. In the top panel of Figure 3.1, we show an example test orbit placed in the simulated survey described in Section 3.3.

3.2.1.2 Propagate Test Orbit and Gather Detections

The test orbit is propagated to all possible epochs in the survey. For each unique epoch, all detections within some area, A, measured on the sky plane are gathered. We term the area, A, a "cell" of observations where the test orbit could be representative of the underlying orbital distribution.

[3]Though, occasionally, such objects are still serendipitously found, as witnessed by the recoveries of NEOs in our test data set.

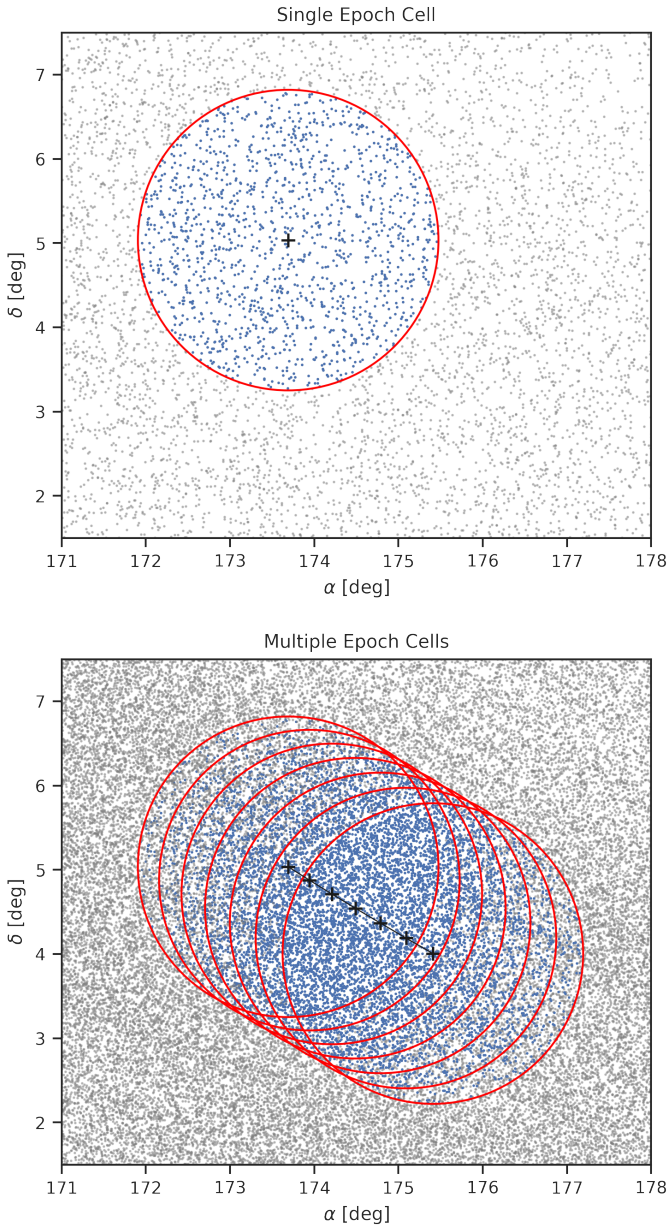

Figure 3.1 In the top panel, simulated detections of real orbits on the first night of a simulated survey are plotted in gray. The survey consists of 16 ten deg^2 fields visited once every other night over a 14 night window. The location of the test orbit on the first night is shown as a black plus sign. The red circle outlines the cell of gathered detections, which are plotted in blue. In the bottom panel, the test orbit is propagated to all possible times in the survey (the remaining six possible exposures) with a cell of observations gathered at each predicted location and epoch. The simulated detections of the subsequent six visits are plotted in gray, in addition to those from the first night. The gathered detections are plotted in blue, as in the top panel. The black line tracks the sky-plane motion of the test orbit, with its location on each line plotted as black plus signs. This figure was generated using plots_simulations.ipynb.

Figure 3.1 shows the topocentric sky-plane motion of a test orbit when propagated through a simulated survey. The black line tracks the sky-plane trajectory of the test orbit. The red circles outline cells of observations appearing at compatible times in the survey.

3.2.1.3 Heliocentric Transformation and Projection into the Test Orbit's Frame of Reference

Once all cells of detections have been gathered, they can be transformed into the heliocentric frame, rotated, and then projected onto a gnomonic tangent plane centered about the motion of the test orbit. We refer the reader to the Appendix E for details on this procedure.

Figure 3.2 shows the transformed, rotated, and then projected detections in the corotating frame of the test orbit. Both θ_X and θ_Y are measured relative to the location of the test orbit at each unique epoch. Here, θ_X lies along the plane of the test orbit, whereas θ_Y is perpendicular to the plane of the test orbit. Once the detections have been transformed to the frame of the orbit and projected onto the tangent plane, each cell of transformed observations can be stacked. Any object that has an orbit similar to the test orbit will appear as a cluster: a circular group in the θ_X–θ_Y plane (or equivalently, a line in θ_X–$\theta_Y - t$ space).In the limit where the test orbit is identical to the orbit of an object in the frame centered on the test orbit's motion, that object will appear as a cluster with a radius proportional to the mean astrometric uncertainty of its constituent detections. Any objects that are on orbits near the test orbit will appear as lines or curves in the corotating frame (there are a few such examples visible in Figure 3.2).

3.2.1.4 Hough Transform

The generalized Hough transform is a template-matching algorithm that can extract arbitrary features from multidimensional space (Ballard, 1981). Lines and curves can be extracted from the test orbit's corotating frame of reference using the equivalent of a 3D generalized Hough transform in θ_X, θ_Y and time, t, space. To do so, we make a 2D velocity

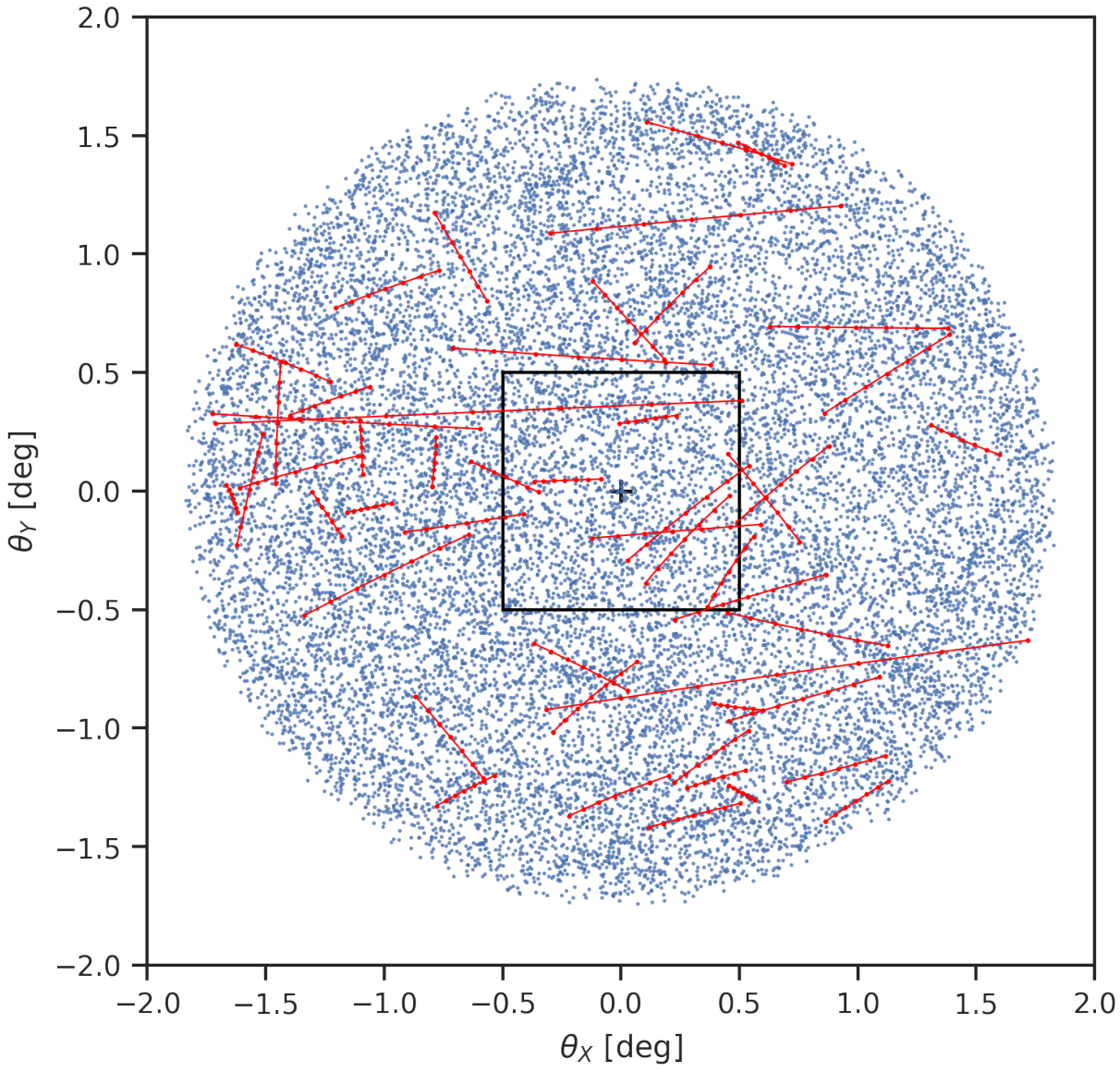

Figure 3.2 Plotted in blue are the cells of detections from Figure 3.1 after they have been transformed to the heliocentric frame and then rotated and projected into the frame of the test orbit's motion. Once transformed and projected, the cells can be stacked. Note that the corotating reference frame is fundamentally three-dimensional. The time dimension (z-axis) has been collapsed onto the gnomonic plane (θ_X–θ_Y plane). In its own frame of reference, the test orbit (black plus sign) lies at $(0, 0)$. Objects with orbits similar to the test orbit will appear as lines or curves, some of which are visible in this image. We plot the detections of 50 such objects in red, with their underlying true linkages plotted as red lines. There are a total of 1397 objects with at least five detections in this image; these are considered findable. The black square outlines the subset of the data used to describe how clustering is performed in Figure 3.3. This figure was generated using plots_simulations.ipynb.

grid in $d\theta_X/dt$ and $d\theta_Y/dt$. By shifting the detections using a grid of assumed linear velocities relative to the test orbit's motion, we effectively perform a 3D Hough transform as a 2D clustering problem.

In Figure 3.3, the subset of data outlined by the black square in Figure 3.2 is plotted. The left panel shows the corotating frame of the test orbit, with the relative motion of nine other objects drawn as lines. The remaining detections with their underlying true linkages are shown in blue. Note the cluster shown in red at $(0, 0)$. Surrounding the cluster is a circle indicating the chosen clustering radius. In the right panel, the same detections after they have been shifted assuming a linear velocity in both X and Y are plotted. Note how one of the nine other objects now forms a cluster; that cluster is also surrounded by a circle indicating the clustering radius (this cluster was found using the actual clustering code). Because the selected test orbit was the orbit of a real object, in the nonzero velocity frame, its detections now also form a line, as shown by the red line toward the middle of the right panel.

With a fine velocity grid, it is likely that similar test velocities will result in the building of clusters with identical subsets of detections. This is also a function of the maximum allowable clustering radius. After all velocities have been tested, the clusters are scanned and any duplicates that have the same constituent detections as another cluster are removed.

We consider any cluster with detections belonging to a single object a "pure" cluster. Some clusters may be contaminated with observations from artifacts or some other, unassociated object(s). We call these "partial" clusters. For example, a cluster with a total of ten detections, where eight of those detections belong to one object and two detections are either false positives or belong to another object or objects, would have a contamination percentage of 20%. Partial clusters could still permit for the discovery of objects, as outlier rejection can be used to filter out the impostor detections. Any cluster that is neither pure nor partial is considered a "mixed" cluster. In the next section, we describe how the process of orbit determination is used to test the validity of the clusters, and how outlier rejection is used to clean partial clusters.

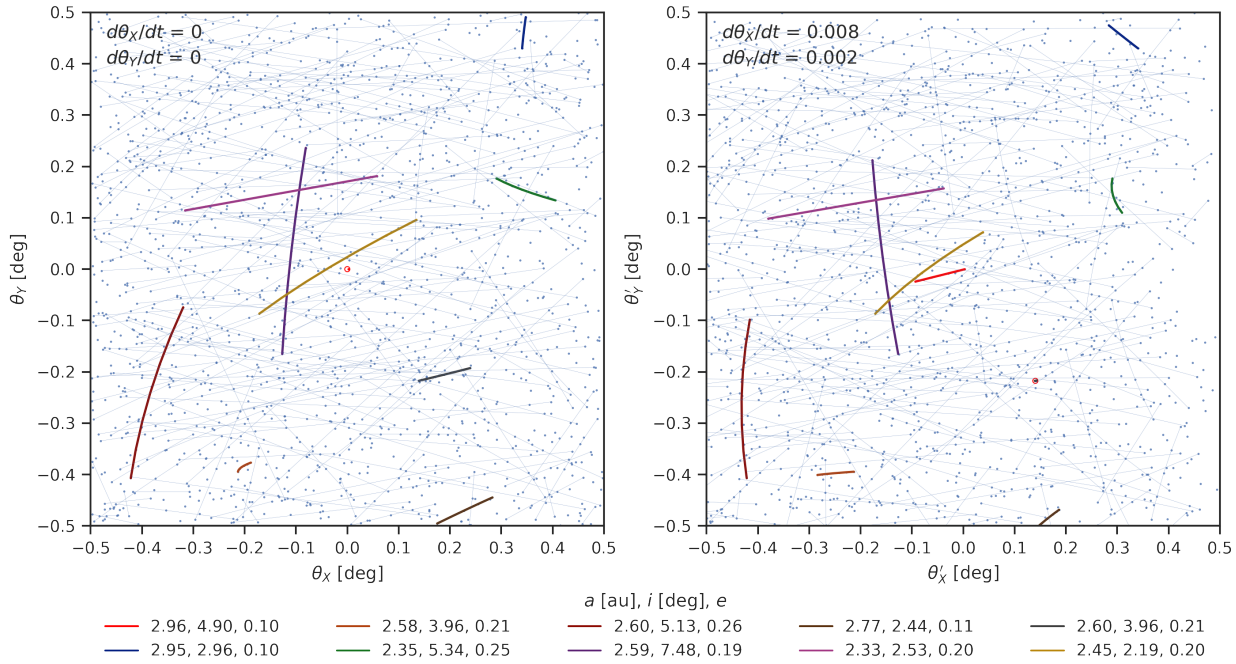

Figure 3.3 By assuming some linear velocity in the corotating frame of the test orbit, objects with similar orbits can be recovered. In the left panel, we plot the same data as outlined by the black square in Figure 3.2. The test orbit is the red cluster of points centered on $(0, 0)$. The red circle surrounding the cluster is the effective cluster radius used (0.005 deg). In blue are the detections of other objects with their true underlying linkages also shown. We plot nine of these other true linkages as colored lines to highlight how we can find other objects. In the right panel, we shift the detections by assuming some velocity in both θ_X and θ_Y relative to the test orbit. An object that originally appeared as a line in the test orbit's projected space now appears as a cluster. Also note that, because our test orbit was a real object, in the nonzero velocity frame it appears as a line, whereas in the zero velocity frame it appears as a cluster. This figure was generated using plots_simulations.ipynb.

3.2.1.5 Orbit Determination

Orbit determination in THOR is a three-step process with the aim of testing the validity of the clusters produced by the Hough transform described in Section 3.2.1.4.

Initial orbit determination (IOD) serves as a first-level filter against erroneous linkages, intended to efficiently reduce the number of clusters to be considered for full differential orbit determination (OD). We implemented the method of Gauss following the formalism introduced by Milani et al. (2008). It seeks to fit a Keplerian orbit to three observations on the sky by solving for the orbit at the time of the second observation. This involves solving an eighth-order polynomial, which can result in up to three solutions per linkage. A challenge for this approach is choosing which three observations to use: as our implementation is not computationally expensive, we elected to fit every combination of three observations for each cluster (i.e., a cluster of five constituent observations would mean testing ten combinations) until a combination is selected such that the χ^2_ν with respect to the observations[4] meets a user-defined threshold. If all combinations are exhausted before the threshold is met, and no outliers can be removed, then the candidate orbit is discarded.

To clean and filter "partial" linkages or clusters, we added outlier rejection to the IOD component of THOR's orbit determination pipeline. The user sets a contamination threshold that defines the maximum percent of observations in a linkage that can be flagged as outliers. If all combinations of three observations have been tested via the method of Gauss, the linkage has not yet met the acceptance threshold, and the linkage has sufficient observations to allow for outlier testing, then the observation with the highest χ^2 residual is removed. If the orbit's overall χ^2_ν meets the user-defined threshold, the orbit is accepted. If the orbit does not meet the threshold, the next-highest residual observation is also removed, and so on, until the maximum number of observations have been removed as outliers or the threshold has been

[4]We note that one must be careful when interpreting the distribution of χ^2 values when only three observations are used to fit the preliminary orbit, but residuals are calculated against all N observations in the cluster. The resulting distribution will not follow the χ^2 distribution with $2N - 6$ degrees of freedom, but generally be wider.

met.

The dynamic range of χ^2_ν values obtained for clusters may span many orders of magnitude. This is also true even for pure clusters, where the approximate nature of the method of Gauss might yield residuals as high as several arcseconds. To further improve these solutions, a refit could be attempted using a more robust technique such as the method of Herget (1965). However, we found it sufficiently computationally efficient to simply pass all orbits satisfying a very relaxed IOD cut to the differential corrector. The corrector then acts as a much stronger and more accurate filter.

After IOD, the preliminary orbits are differentially corrected (OD) using the method of weighted linear least squares in combination with numerical differencing to determine the Jacobian (the matrix of partials of the observations with respect to changes in the orbit state). We implemented the differential correction algorithm described in Vallado (2013), with a few notable differences: we added outlier rejection in a similar fashion to the IOD code, a χ^2_ν acceptance threshold, and adaptive finite differences.

The goal of differential correction is to improve a preliminary orbit solution by minimizing the residuals across all its constituent observations. The improved orbit is propagated to all observation times, and residuals with respect to the observations are calculated. If such a differentially corrected orbit meets a user-defined χ^2_ν threshold, it is accepted. If the differentially corrected orbit does not meet the user-defined threshold, iteration continues until the maximum number of iterations is reached. If, at this point, the orbit does not meet the user-defined χ^2_ν threshold, the observation with the highest χ^2 value is removed and iterative improvement is continued using the new set of observations. The iterative outlier removal continues until the maximum allowable number of observations have been removed as determined by the user-settable contamination threshold. This process is an effective filter against erroneous ("mixed") orbits, and is an effective method to clean contaminated ("partial") orbits.

Orbits that pass the differential correction step will likely contain only a subset of the observations that belong to those orbits. Some orbits also will share observations with other

orbits (e.g., orbits that represent different sub-arcs of one longer arc). The last step in THOR's orbit determination process aims to merge orbits that represent the same object, and extend orbits that could have more observations associated with it. This is done through an iterative combination of OD and observation attribution. We perform attribution by propagating the differentially corrected orbits to all epochs of the observations contained in the gnomonic tangent plane of the test orbit. Any observations that lie within a user-defined angular distance of the predicted location are identified as possible attributions. If, at this point, any observations are identified as belonging to multiple orbits, these orbits are merged to form one larger orbit (a "child" orbit) and refit. If such a refitted child orbit passes the OD filter (i.e., meets the χ^2_ν threshold), its parent orbits are removed from the active pool of possible orbits. Any orbits that pass OD but could not be improved further are unlikely to have more observations attributed to them and are output as high-quality candidate orbits. This iterative combination of attribution and refitting is done until each orbit has only unique constituent observations and the orbits are not further improved by OD.

In summary, the orbit determination stage filters and transforms the candidate clusters produced by the Hough transform step into a catalog of high-quality, differentially corrected orbits. These successfully linked orbits and their observations are removed from the active pool of observations, and steps 3.2.1.1 - 3.2.1.5 are repeated with new test orbits to attempt to link the remaining observations.

3.2.2 Phase Space Gridding

A key element of THOR is the selection of test orbits and how to best populate the relevant phase space to maximize completeness for different populations of minor planets. The bulk of the Solar System small bodies that LSST will discover are Main Belt asteroids (Jones et al., 2018). We initially focus on populating the phase space to discover populations in the Main Belt and the slower-moving outer Solar System populations. Future work is planned to target the faster-moving NEO population, and to further optimize test orbit selection.

Guided by tests on real ZTF data, the current implementation of THOR uses nine average

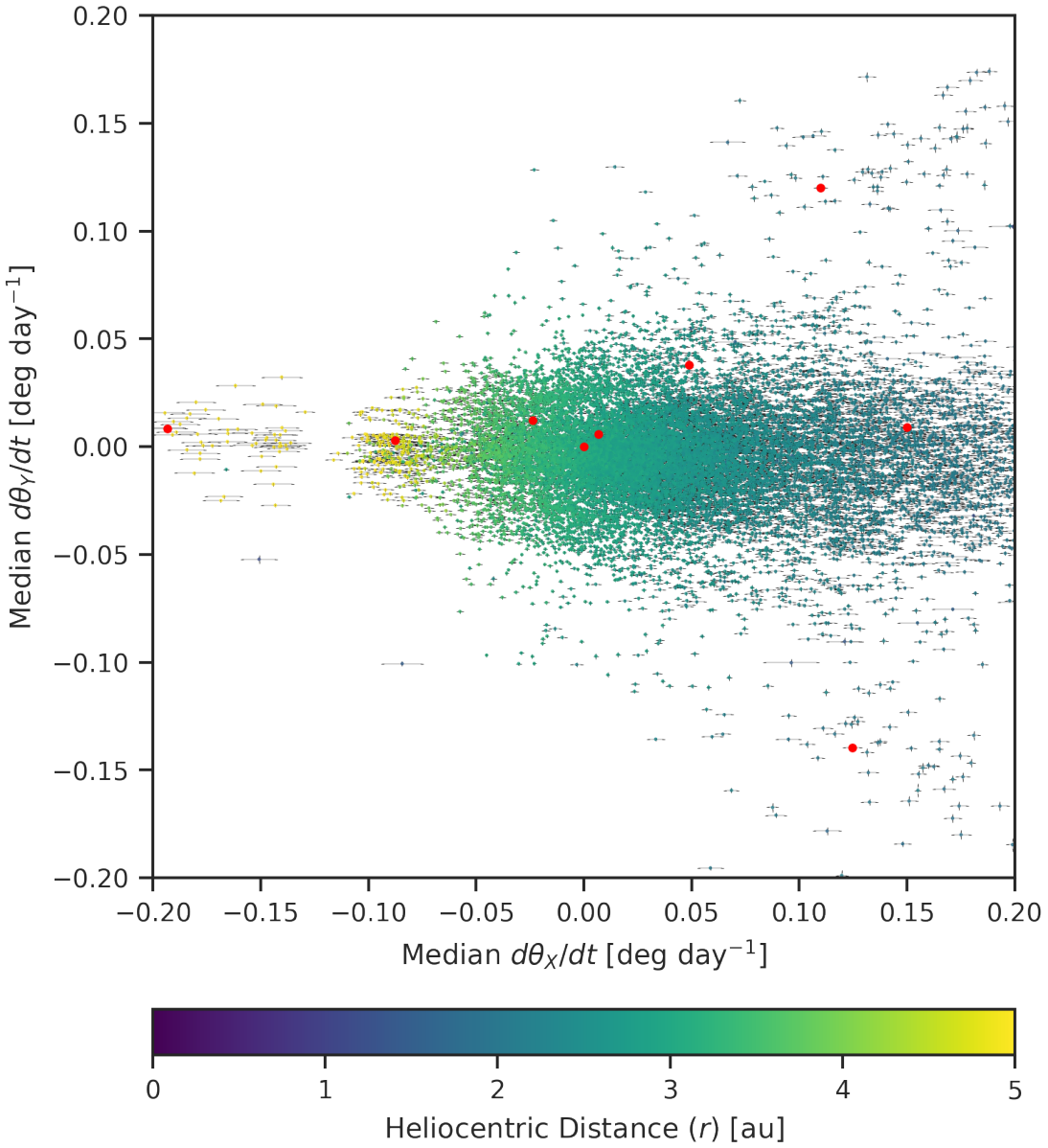

Figure 3.4 Using one of the test orbits from the simulated survey but now setting the cell size to 500 deg^2, we plot the median velocities in θ_X and θ_Y relative to the test orbit for each findable object in the corotating frame of the test orbit. The error bars are the respective standard deviations of $d\theta_X/dt$ and $d\theta_Y/dt$, which serve as a proxy for the curvature in the projected space. The points are colored by the median heliocentric distance. The red points show the locations of the test orbits selected for the survey using the orbit selection criteria described in Section 3.2.2. As before, the test orbit on which this image is centered has a velocity of $(0, 0)$ deg day^{-1}. The default velocity grid is set to 300 equally spaced bins between -0.1 and 0.1 deg day^{-1} in both coordinates. The range of the velocity grid allows different heliocentric distances to be probed, which in effect allows one test orbit to recover other objects as long as their motion relative to the test orbit appears linear and the rate of motion lies within the searched grid. This figure was generated using plots_simulations.ipynb.

orbits from the known population of minor planets per patch of sky. The patches of sky are currently a simple rectangular sky-plane subdivision 15 by 15 deg in (α, δ) with no overlap. This notional subdivision could easily be replaced by something more robust such as HEALPix (Gorski et al., 2005). The cell size was set to 1000 deg^2 to maximize the number of detections gathered by each test orbit at each epoch. This size permits a test orbit to be selected along any edge of a patch and still gather all detections in that patch of sky during the first epoch of observation. The patches in which an average orbit is calculated do not have any effect on the detections that propagating a test orbit can gather; they are simply used to find a set of test orbits.

The test orbits are selected in bins of semi-major axis (a), with bin edges roughly corresponding to the location of the most prominent mean-motion resonances (the Kirkwood gaps).To target the Main Belt, we set the inner most edge at $a = 1.7$ au with the outermost edge at $a = 5.0$ au. The remaining bin edges were set to $a = 2.06$, 2.5, 2.82, 2.95, and 3.2 au, which correspond roughly to the Saturn ν_6 secular resonance, and the 3:1, 5:2, 7:3, and 2:1 Jovian mean-motion resonances, respectively (Minton & Malhotra, 2009). To target the outer Solar System, we added a single bin with edges $a = 5.0$ au and $a = 50.0$ au.

For each bin, an object with average eccentricity (e) and inclination (i) is selected from the Minor Planet Center's (MPC) catalog of known objects. No test orbits with eccentricity greater than 0.5 were selected. This relatively simple orbit selection is not sufficient to target the complete Main Belt population. To target the missed populations from earlier trials, we subdivided the first semi-major axis bin further – in particular, the bin located at the approximate location of the Hungaria family of asteroids ($e < 0.18, 1.78 < a < 2.00$ au, $16 < i < 34$ deg). The Hungarias are notable for their wider spread of orbits in orbital phase space, especially in inclination (Warner et al., 2009). Three test orbits were selected in different bins of eccentricity (e) with bin edges (including the exterior edges) $e = 0.0$, 0.1, 0.2, and 0.4.

In Figure 3.4, we plot the median linear velocities in θ_X and θ_Y of all the findable objects relative to one of the test orbits selected for the simulated survey. The simulated survey

is described in the following section. The plotted error bars are the standard deviations of the velocities, and are a proxy for the curvature in the 3D space of the gnomonic plane and time. The test orbits that were selected for use on the simulated survey are plotted in red. We used many different iterations of Figure 3.4 to inform decisions regarding the test orbit selection code. The test orbit selection criteria and code were designed in such a way as to maximize coverage in this space.

Note that, by choosing different upper and lower velocity bounds, different heliocentric distances are probed by each test orbit. For example, the velocity grid can be set to sweep out a wider heliocentric distance at the cost of greater computation time. Equivalently, an additional test orbit could be used with a smaller grid to the same effect and at a possibly lower computational cost. These are trade-offs that may depend on the survey and its data properties – in particular, a survey's observational depth influences the number of observations, and its hardware and software characteristics influence the occurrence of false positives. In general, the propagation and gathering of detections takes on the order of seconds on a single thread, whereas the current Hough transform implementation takes on the order of minutes to hours with moderate parameters on a few dozen threads. The Hough transform component of THOR is where future optimization will be focused, in addition to improving test orbit selection.

3.3 Validation with Simulated Data

3.3.1 Simulation Setup

To test the efficacy of our algorithm, we created a simulated mini-survey composed of 16 square fields 10 deg^2 in size. Each field is visited once every other night over a 14 day period, yielding seven unique visits per field over the course of the simulated survey. We downloaded the Minor Planet Center's catalog[5] of known orbits and used open-source orbit propagation and determination software, `oorb`, to generate ephemerides for each known object at the

[5]

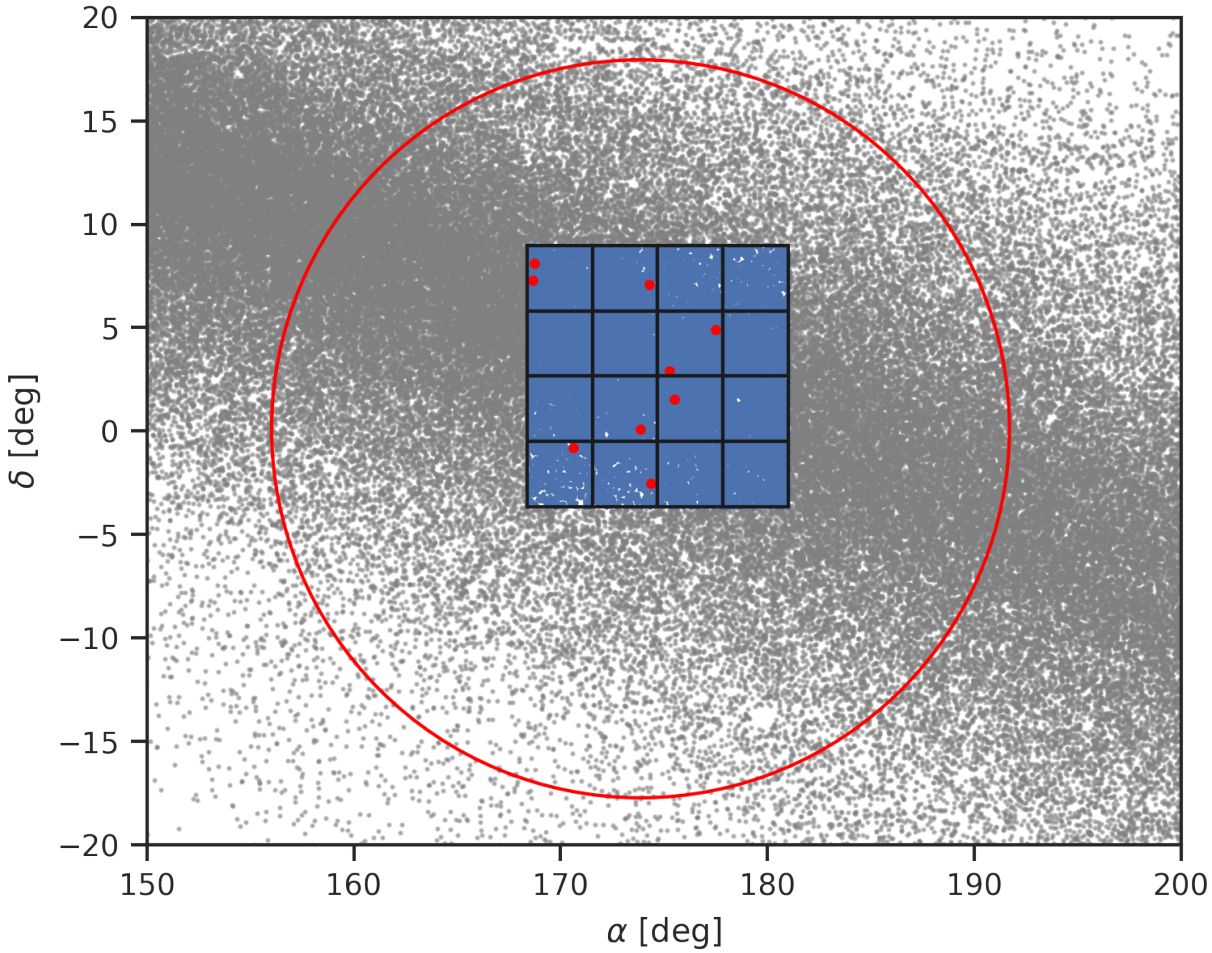

Figure 3.5 Plotted in blue are the simulated detections of real orbits generated using the MPC's orbit catalog for the first night of the simulation. The ephemerides for the underlying orbit catalog from which the blue detections are drawn are shown in gray. The red points indicate the locations of the selected test orbits on the first night. The red circle traces the 1000 deg^2 cell size for one of the orbits. The 16 square fields (black squares) each ten deg^2 in size are visited once every other night over a period of 14 nights. The footprint was roughly centered on the ecliptic to maximize minor planet density. No false-positive detections are plotted, even though they were included in the simulated survey. This figure was generated using plots_simulations.ipynb.

different survey times (Granvik et al., 2009). In practice, THOR will run on the detections resulting from difference imaging; however, simulated ephemerides are sufficient to test the performance of linking detections.

The survey footprint was roughly centered on the ecliptic to maximize the density of minor planet detections (see Figure 3.5). We added 100 ± 10 randomly distributed false-positive detections per square degree to simulate the effect of image differencing artifacts. The final survey consisted of $139,120$ (55.4%) simulated observations belonging to real orbits and $112,078$ (44.6%) false-positive observations. The simulated observations were randomized with 100mas scatter to reflect the expected astrometric errors for moving objects as observed by current-generation surveys.

We consider any minor planet with at least five detections throughout the survey to be findable, yielding a total of $18,332$ such objects. Any object that is recovered in a pure orbit is considered to be found. For comparison, the Vera C. Rubin Observatory's current implementation of MOPS would require three tracklets (six detections in three intranight pairs) for an object to be findable – as a tracklet-based algorithm, it would recover none of the objects in the simulated survey.

The contamination threshold for partial clusters was set to 20%, with a minimum cluster size of five detections (see Section 3.2.1.4). The default grid in both $d\theta_X/dt$ and $d\theta_Y/dt$ for the Hough transform was set to 300 equally binned velocities between -0.1 and 0.1 deg day^{-1} relative to the motion of the test orbit, yielding a total of 90,000 velocities tested per test orbit. The minimum cluster size was set to five detections (equal to the findability criterion), with a cluster radius of 0.005 deg. To extract clusters for every velocity tested, we use scikit-learns implementation of the Density-Based Spatial Clustering of Applications with Noise (`DBSCAN`) algorithm (Ester et al., 1996; Pedregosa et al., 2011). `DBSCAN` is a nonparametric clustering algorithm that finds clusters in higher-dimensional space of approximately equivalent density given a maximum distance parameter and marks unclustered data as outliers. We use open-source orbit determination software, `oorb`, to propagate the test orbits to all possible times in the survey (Granvik et al., 2009).

For orbit determination, the χ^2_ν and contamination thresholds were set to (10^5, 20%), (10, 20%), and (10, 0%) for IOD, OD, and the iterative combination of differential correction and attribution, respectively. The maximum attribution distance was set to one arcsecond. The minimum number of observations and minimum arc length for an orbit to be accepted were set to five observations and one day, respectively.

3.3.2 THOR Performance

Using nine test orbits, selected as described in Section 3.2.2, THOR recovered 16,729 candidate orbits representing $16,728$ (91.3%) of the findable objects in the simulated survey. Only one erroneous orbit was produced.

The top panel of Figure 3.6 shows the recovered orbit completeness in bins of semi-major axis (a) and inclination (i). The red contours indicate the number of objects that should be findable, while the red points show the selected test orbits. The vertical dashed lines trace the semi-major axis bin edges. The percentages state the overall completeness in each bin of semi-major axis (note that the first bin has three test orbits). In the bottom panel, we repeat the same plot configuration as in the top panel, but we plot eccentricity (e) instead of inclination. Instead of percentage completeness in the five semi-major axis bins, we state the numbers of unique objects found to give the percentages context. As shown in both panels, the orbits selected in the bins beyond 2.5 au recover the vast majority (96.1%) of the findable objects within that region. Performance in the semi-major axis bin between 2.06 au and 2.5 au (85.2%) suggests that either more test orbits are needed or that the test orbit we selected is not adequately representative of the underlying orbital distribution. The same is true for the three orbits selected in the semi-major axis bin between 1.7 au and 2.06 au (61.5%). Unsurprisingly, the two Hungaria orbits did not completely recover their constituent population, because of the larger spread of Hungaria orbits in orbital element phase space. The high-eccentricity orbit selected did, however, enable other high-eccentricity objects to be recovered in addition to many objects in bins with larger semi-major axis. The test orbits were run in series, starting in order of increasing semi-major axis. The The second orbit that

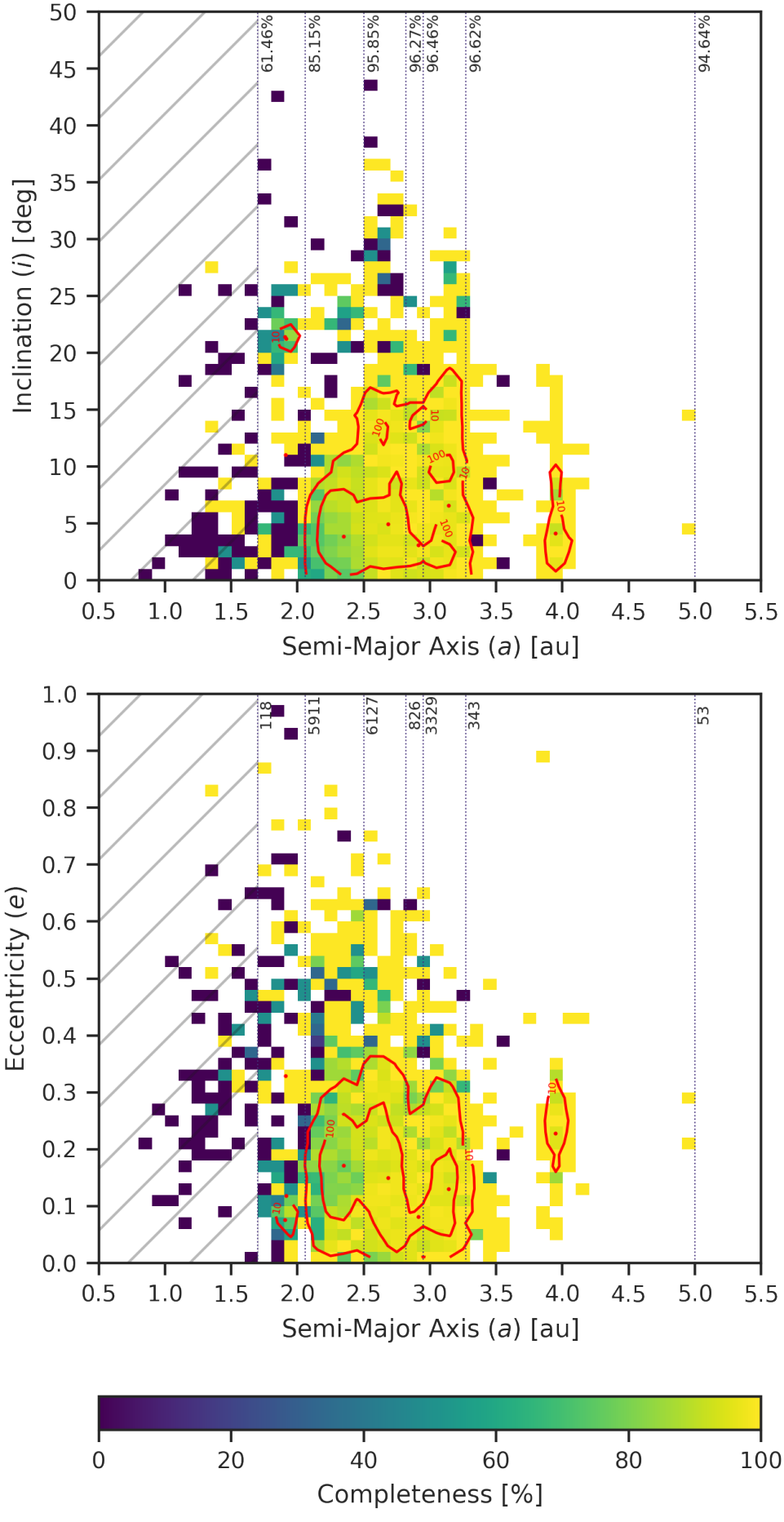
61.46%
85.15%
95.85%
96.27%
96.46%
96.62%
94.64%
Inclination (i) [deg]
Semi-Major Axis (a) [au]
118
5911
6127
826
3329
343
53
Eccentricity (e)
Semi-Major Axis (a) [au]
Completeness [%]

Figure 3.6 In the top panel, the percentage of objects found in pure orbits is visualized in bins of semi-major axis (a) and inclination (i). Number density contours are drawn as red lines to show the number of objects findable (five or more simulated detections throughout the survey). The nine test orbits used are plotted as red points. The vertical dashed lines indicate the five semi-major axis bin edges, with the overall percentage completeness per bin written at the top. The hatched area indicates the region of semi-major axis space where no test orbits were chosen. In the bottom panel, we plot percent completeness in bins of semi-major axis (a) and eccentricity (e); contours and test orbits are plotted in the same style as in the top panel. Instead of percent completeness in each of the five bins, we now explicitly state the number of objects found. This figure was generated using plots_simulations.ipynb.

ran was the one selected in the inner Main Belt, and over 10,000 of the findable objects in the survey were successfully found by using this test orbit. Overall completeness for the Main Belt and beyond ($a > 1.7$ au) is 91.6%. While we did not focus test orbit selection to tackle the NEO population and populations interior to the Main Belt ($a < 1.7$ au), serendipitous completeness for these objects is 18.8%.

3.3.3 THOR Performance

The complete simulations run, starting from a catalog of simulated observations to producing a catalog of $16,729$ orbits using nine test orbits took seven hr on a personal workstation using 60 threads on 30 cores. The workstation has the following specifications: a 32-core/64-thread 2990WX AMD Threadripper, 128 GB DDR4 RAM, and 2 TB of SSD drive space. The run required a moderate amount of memory, especially during the attribution of orbits, approximately 1.5 GB per thread. Given the heavy reliance of the current implementation on Pandas data frames, it is likely the memory usage could be further optimized. Theoretically, the algorithm should require little more than 2–3x the space occupied by the observations.

3.3.4 Completeness and Purity Considerations

In Figure 3.7, we plot completeness (in blue) and purity (in green) as functions of the different pipeline components. Plotted in red is the number of linkages at each stage. From clustering

to the final stage, we see only a 0.2% drop in completeness, while the purity increases from 48.5% to 99.9%. Moreover, the number of clusters, initial orbits, and recovered orbits, quickly approaches the true number of objects recovered at each subsequent stage of the pipeline. The nearly constant completeness across the different pipeline components indicates that further completeness is likely to be found by improving our test orbit selection algorithm.

3.4 Validation with Zwicky Transient Facility Data

The Zwicky Transient Facility (ZTF) is a robotic time-domain survey of the northern sky capable of scanning more than 3700 deg^2 hr^{-1} with its 47 deg^2 field of view, 600 megapixel camera, and readout time of just 8 s. ZTF can observe to a median 5σ r-band depth of 20.6 mag (Bellm et al., 2019; Graham et al., 2019; Masci et al., 2019). During a full night of observing, ZTF may issue as many as 1.2 million alerts ($\sim$1/10 of what is expected for LSST) to science users through its alert broker system (Patterson et al., 2019). Contained within the alerts are the detections of moving objects. In particular, if a previously discovered Solar System object is predicted to be at the location of a difference image point source, the source can be attributed to the known object (Masci et al., 2019). In its first three months of operations, ZTF submitted $\sim$600,000 observations of minor planets and has discovered about 320 new objects (Bellm et al., 2019). As only its public survey cadence is designed to emulate the baseline cadence of LSST and produce tracklets, the full ZTF data set provides a good framework to test THOR on data that otherwise would not be readily useful for minor planet discovery.

3.4.1 ZTF Alert Data Set

We downloaded 15 nights worth of ZTF alerts dating from 2018 September 3 to 2018 September 17. We filtered out static sources and selected all alerts with real bogus values above 0.5, emulating future LSST requirements. This yielded a final count of 255, 358 (30.9%) known object observations that could be matched to objects in the Minor Planet Center's orbit catalog. The remaining 572, 188 (69.1%) alerts are unassociated and therefore unknown. It

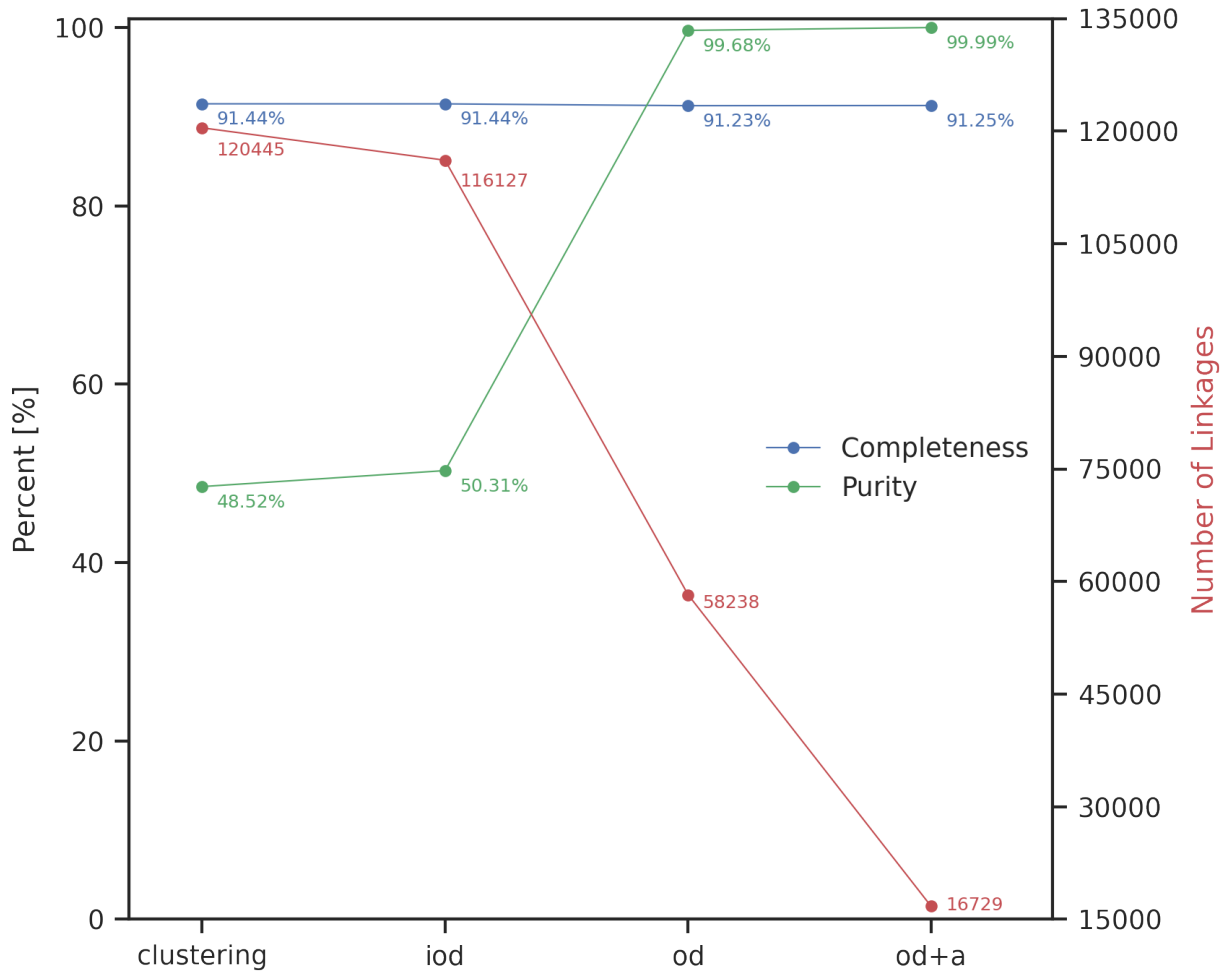

Figure 3.7 Completeness and purity as functions of pipeline components are plotted in blue and green, respectively. The number of linkages is plotted in red. Each subsequent component of THOR enhances linkage purity, while completeness remains largely unchanged. The number of linkages quickly approaches the true number of recovered objects. This figure was generated using plots_simulations.ipynb.

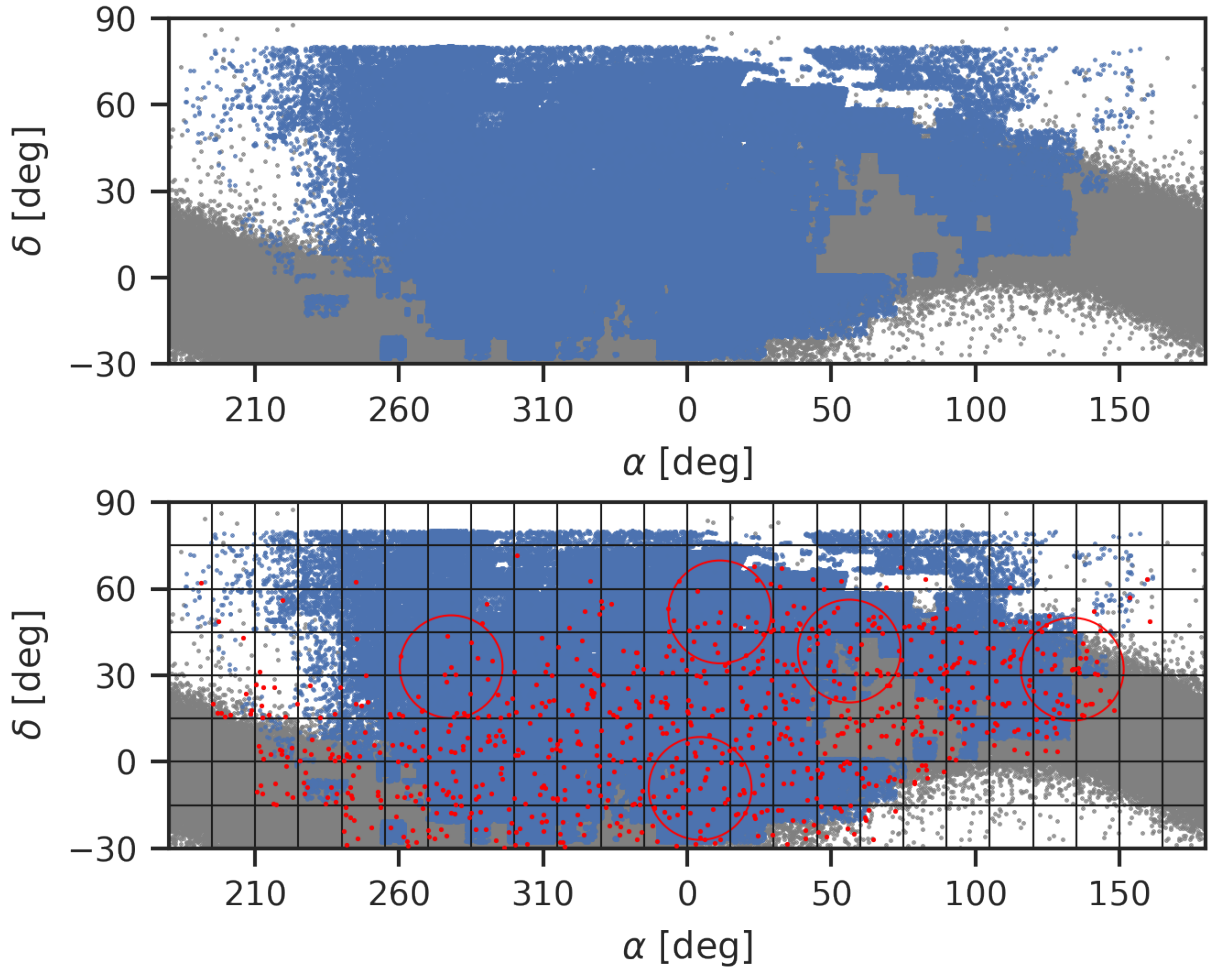

Figure 3.8 In both panels, 15 nights of filtered ZTF alerts are plotted in blue. The catalog of known orbits with ephemerides generated for the first night of alerts by ZTF are plotted in gray. From the gray population, test orbits are selected. In the bottom panel, we plot the same as the top panel but show the chosen sky-plane subdivision for test orbit selection (the 15x15 deg^2 patches). For each patch, test orbits (in red) from the catalog of known objects were selected. If no ZTF observations occurred in any individual patch, no test orbits were selected. If no known object is predicted to be in any patch or bin of semi-major axis, no test orbit was selected. We randomly selected five of the test orbits and outline their 1000 deg^2 cells of observations as red circles. This figure was generated using plots_ztf.ipynb.

is likely that among the unassociated detections there are undiscovered moving objects in addition to astrophysical transients and false positives (e.g., image subtraction artifacts).

During the 15 night window of alerts, ZTF scanned over $20,000$ deg^2 of sky (for comparison, the simulated survey had a footprint of 160 deg^2). We divided the sky plane into 15 by 15 deg patches of sky, and for each patch, selected nine average known objects in different bins of a as described in Section 3.2.2. If no known object exists in a bin of a for a particular patch, no test orbit was used for that bin. A total of 821 test orbits were chosen. In Figure 3.8, we plot the ZTF observations in blue with the underlying catalog of known orbits in gray. The red dots represent the on-sky location of the 821 test orbits chosen and the five red circles indicate the effective cell area on the sky. The black lines outline the 15 by 15 deg patches used to select up to nine test orbits per patch.

The Hough transforms and orbit determination for each test orbit were conducted with the same parameters and configuration as in the simulated survey (300 equally binned velocities between -0.1 to 0.1 deg day^{-1} and a maximum allowable clustering radius of 0.005 deg). For clustering, the contamination threshold was set to 20% and the minimum cluster size was set to five observations. For orbit determination, the χ^2_ν and contamination thresholds were set to (10^5, 20%), (10, 20%), and (10, 0%) for IOD, OD, and the iterative combination of differential correction and attribution, respectively. The maximum attribution distance was set to one arcsecond. The minimum number of observations and minimum arc length for an orbit to be accepted were set to five observations and one day, respectively.

Compared to the simulated survey, nearly 100 times as many test orbits were chosen. We grouped the test orbits into chunks of five patches (yielding a maximum of 45 test orbits per chunk). Each chunk was submitted to an individual node on the University of Washington's HPC cluster, Hyak. After each chunk finished processing, the recovered orbits were concatenated and a simple algorithm was run to select the orbit with most observations and longest arc for each set of orbits that shared observations. After shared observations were assigned to their best-fitting orbit, a final step of differential correction was executed to update any orbits that had any observations observations removed.

Like the simulated survey, we consider any known object with at least five detections throughout the 15 nights of alerts to be findable, resulting in a total of 21,542 such objects. Assuming a classical MOPS approach, we find that only 9,381 (44.8%) of the known objects would be findable. With the ZMODE algorithm, we find 14,291 (68.2%) would be findable[6]. ZMODE is designed to work with ZTF's cadence and attempts to link objects with a minimum of four detections over four consecutive nights, while the classical LSST MOPS algorithm looks to link three pairs of detections over 15 consecutive nights and thus requires a minimum of six observations. We consider any object recovered in a pure orbit with at least five observations to be found.

3.4.2 THOR Performance

THOR recovered 21,723 candidate orbits consisting of nearly 167,000 unique observations from the described two-week slice of the ZTF data set. These include 21,018 pure orbits for 20,940 (97.2%) of the 21,542 minor planets with at least five detections. The remaining 705 orbits contain either only unassociated/unknown observations or a mix of associated and unassociated observations. Of the 21,723 recovered orbits, 1,783 orbits contain only singleton observations (orbits made completely without tracklets). In the top panel of Figure 3.9, we plot completeness in bins of semi-major axis (a) and inclination (i). The red number density contours show the number of objects that should be findable, while the red points indicate the chosen test orbits. The vertical dashed lines trace the bin edges. The outer Solar System bin stretches from 5.0 to 50 au, but is plotted in its entirety. The percentages at the top of each bin list the overall completeness in each bin. In the bottom panel, we repeat the same plot configuration, but instead of inclination, we plot eccentricity (e). We replace the percentage completeness values in each bin with the number of found objects, to give the percentages greater context. As in the simulated survey, the performance on ZTF alerts beyond 2.5 au suggests most of the findable objects were recovered (98.4%). The performance

[6]The numbers stated here assume *ideal* performance by both pipelines – i.e., 100% completeness.

in the bins nearer than 2.5 au is better than in the simulated survey. This is likely due to the designed overlap of selecting nine orbits per 15 × 15 deg^2 patch of sky and setting the cell size to 1000 deg^2 (see Figure 3.8). Note that, in the bottom panel, it is clear that high-eccentricity objects were not favored in recovery. Test orbit selection was limited to orbits with eccentricity less than 0.5, and some of these objects may appear nonlinearly within the corotating frames of the test orbits. Overall completeness for the Main Belt and beyond ($a > 1.7$ au) is 97.4%, including serendipitously linking 28.9% of objects with orbits interior to the Main Belt ($a < 1.7$ au).

Our test orbit selection yielded a total of 821 orbits; however, only 621 of those orbits lead to known objects being found. Two hundred and twenty of the test orbits yielded no known object discoveries, of which 88 had no findable objects in their corotating frame of reference. This is due to several reasons: first, some of the selected test orbits might belong to objects that are too dim to be detectable by ZTF. Second, because the orbits are run in series, moving outward in semi-major axis space, it is likely that some orbits that ran earlier swept out their population of findable objects – resulting in the later test orbits not having any objects to find, especially if the two orbits were similar.

In all, the total number of known objects recovered is 20,940 (97.2%) of the 21,542 with five or more observations. This represents a factor of ∼2 discovery increase over traditional MOPS, and a factor of ∼1.5 increase over ZMODE.

3.4.3 Computational Performance

We grouped each set of test orbits into chunks of five patches. Each group of test orbits was then set to run on a node on UW's HPC cluster, Hyak. A total of 23 nodes, with 28 cores per node, were used. The SLURM jobs were submitted to the backfill ("checkpoint") queue, meaning that jobs are frequently preempted and paused in favor of jobs submitted by the node owners. Regardless of multiple preemptions and job restarts, the full run on two weeks of ZTF took less than 18 hr. Some nodes finished in as little as 10 minutes, and the longest running job took just under 16 hr.

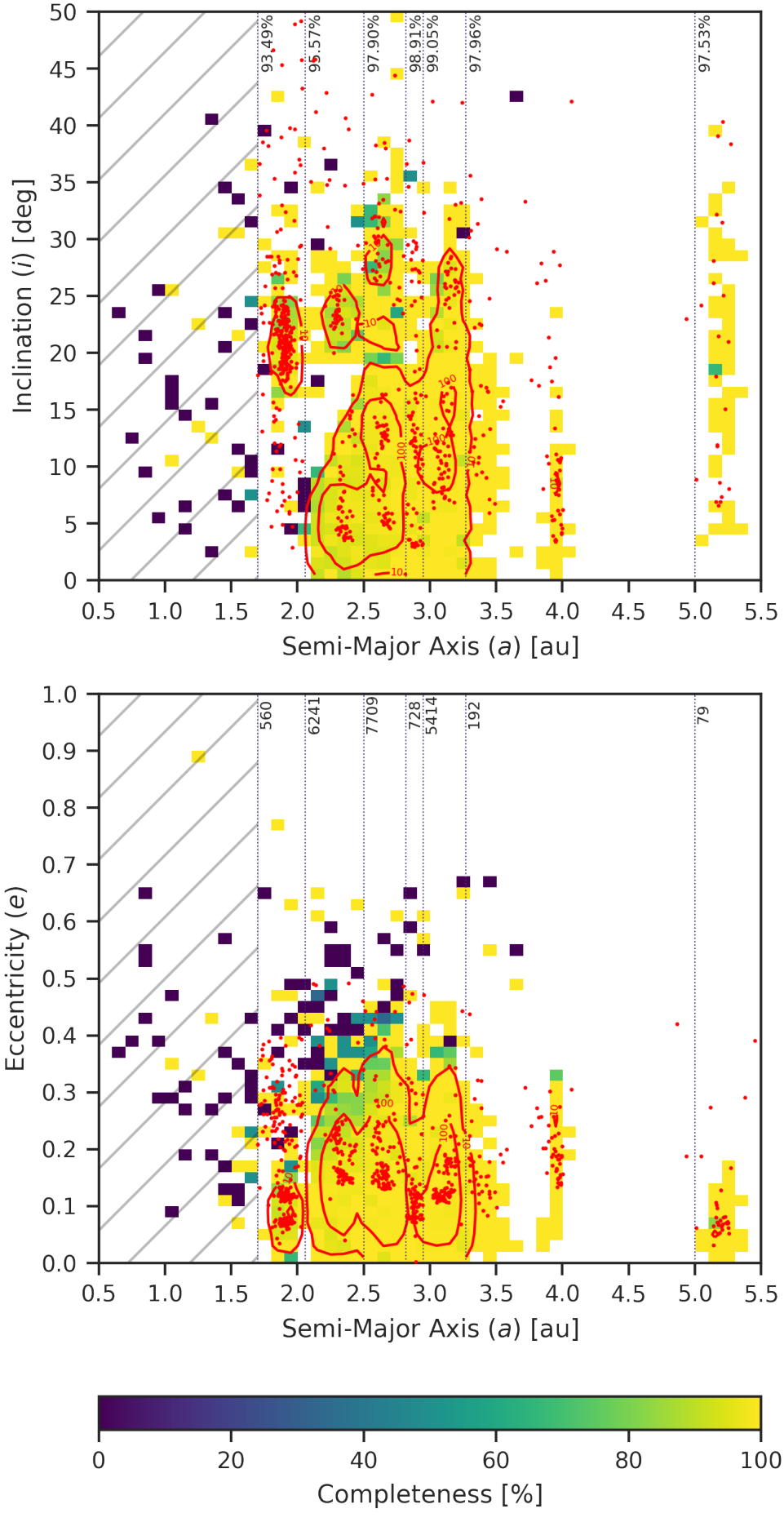
93.49%
95.57%
97.90%
98.91%
99.05%
97.96%
97.53%
Inclination (i) [deg]
Semi-Major Axis (a) [au]
560
6241
7709
728
5414
192
79
Eccentricity (e)
Semi-Major Axis (a) [au]
Completeness [%]

Figure 3.9 In the top panel, the percentage of objects found in pure orbits is visualized in bins of semi-major axis (a) and inclination (i). Number density contours are drawn as red lines to show the number of objects findable (five or more detections throughout the two weeks of ZTF alerts). The hatched area indicates the region of semi-major axis space where no test orbits were chosen. The 821 test orbits used are plotted as red points. The vertical dashed lines indicate the five semi-major axis bin edges, with the overall percentage completeness per bin written at the top. In the bottom panel, percent completeness in bins of semi-major axis (a) and eccentricity (e); contours and test orbits are plotted in the same style as in the top panel. Instead of percent completeness in each of the five bins, we now explicitly state the number of objects found. This figure was generated using plots_ztf.ipynb.

3.4.4 Completeness and Purity Considerations

We next characterize the completeness, purity, and any additional filtering required for submission of candidate new discoveries to the Minor Planet Center. In Section 3.4.4.1, we evaluate which objects we would have discovered and reported in 2018, given the state of MPC's orbit catalog at the time. In Section 3.4.4.2, we use a more recent catalog of known orbits from April 2021 to assess the purity of the sample we would have submitted in 2018, and look for still-unknown objects.

3.4.4.1 2018 Orbit Catalog Analysis

The ZTF alerts that correspond to possible observations of known minor planets are labeled as such as part of the ZTF alert pipeline. These labels or associations are the "ground truth" with which we compare THOR's performance. As stated in the previous section, THOR recovered 21,723 candidate orbits including 21,018 pure orbits for 20,940 (97.2%) of the 21,542 minor planets with at least five detections.

In Figure 3.10, we plot completeness and purity as functions of the different pipeline components. The purity[7] rises from only 3.8% at the Hough transform stage to 96.8% for

[7]Note that the purity as computed here is actually a *lower limit*. The 705 orbits that have been identified as mixed or unknown will contain objects that were undiscovered in 2018, objects that had known orbits not of sufficient quality such that observations could be attributed to them, as well as truly erroneous linkages.

the final recovered orbits, while – quite strikingly – completeness drops by only 0.8%. This suggests that further completeness is likely to be found by improving our test orbit selection algorithm. In absolute terms, the number of linkages drops from $\sim 7 \times 10^6$ clusters to 21,723 recovered orbits. Among the 21,018 pure orbits are 150 orbits that contain only sub-arcs of the longer observations arcs of 72 unique objects in the data set. The continued presence of these duplicates suggests further improvements to the orbit merging and extension algorithm are necessary.

We apply two additional filters to produce a final set of high-confidence candidates that could be submitted to the MPC. First, we remove any observations that occurred within 30 minutes of another observation. This helps in removing observations of stationary sources that may not have been filtered out during preprocessing. Doing so reduces the number of unknown orbits from 705 to 526. Second, we remove any orbits that contain the observations of known objects. In practice, if THOR were deployed as a discovery algorithm, it would not run on observations that have already been associated with known objects. This reduces the number of unknown orbits to 488 high-quality discovery candidates that, had THOR been running in 2018, would have been submitted to the MPC.

3.4.4.2 2021 Orbit Catalog Analysis

At the time of writing, two-and-a-half years of minor planet discovery have passed since ZTF made the observations in 2018. We downloaded the Minor Planet Center's (MPC) catalog of known orbits on 2021 April 20, and ran attribution of the MPC's orbits on the two weeks of ZTF observations, creating a new set of updated associations. We then analyzed how THOR performed, with the benefit of hindsight, to establish the lower limits on the purity of the sample that would have been reported to the MPC in 2018.

Of the 488 orbits that were identified as high-quality, submission-ready discoveries using the 2018 associations, we find that 477 correspond to observations of 476 unique minor planets that have been discovered or had their orbits improved since 2018 (one object was identified twice, as two shorter sub-arcs). The remaining 11 orbits in the high-quality sample

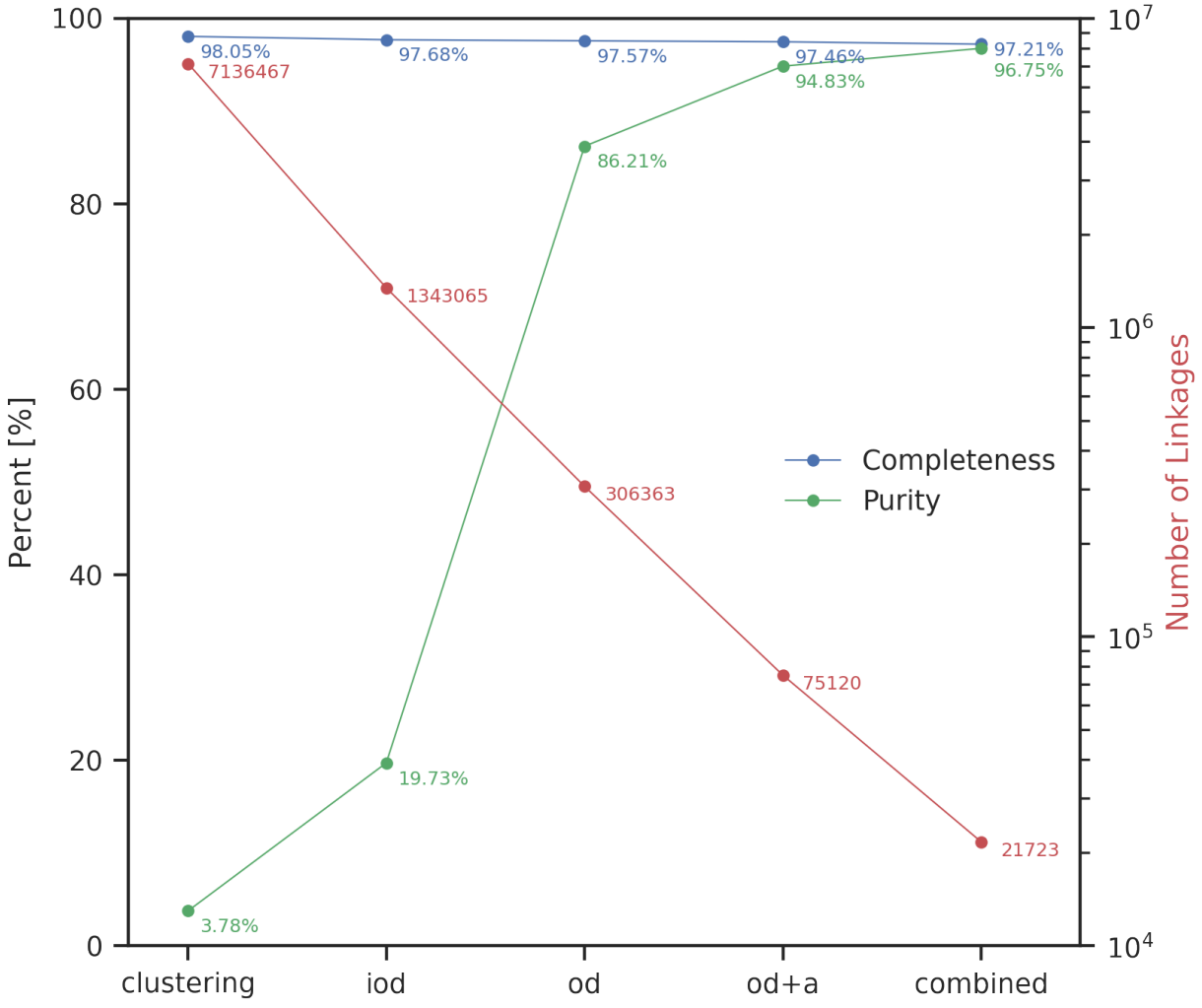

Figure 3.10 Completeness and purity as functions of pipeline components are plotted in blue and green, respectively. The number of linkages is plotted in red. The "combined" step is the simple algorithm used to combine orbits recovered by individual nodes on the HPC cluster. Each subsequent component of THOR enhances linkage purity while completeness remains largely unchanged. The number of linkages quickly approaches the true number of recovered objects. This figure was generated using plots_ztf.ipynb.

are as-yet undiscovered objects or false linkages; based on visual inspection and comparisons of their magnitudes, they do appear to be real objects. The observations and predicted motion of the 11 discovery candidates are plotted in Figure 3.11.

Using the IMCCE SkyBot service (Berthier et al., 2006), we conducted a cone search for all observations belonging to the 11 orbits. The observations of 10 candidates could not be associated to any known objects within 10". Only the candidate with $e > 1$ orbit had all observations within 5" of a known object—the comet C/2018 U1. C/2018 U1 is classified as a hyperbolic comet and was discovered on 2018 October 27. We compared the orbit fit of C/2018 U1 to the orbit as reported in JPL's SBDB.[8] The inclinations agree to within 0.3 deg, and the perihelion distances agree to within 0.2 au, while our fit overestimates eccentricity by 0.1. Given the short arc (6.13 days) and small number of observations (five observations) with which the orbit was fit, we conclude our solution agrees remarkably well with the one present in the SBDB. The ZTF observations were made during a two-week slice that took place six to eight weeks prior to the discovery date. Had THOR been running at the time when data was taken, this comet would have been discovered by ZTF.

Therefore, we conclude that the purity of submissions—an important consideration for the Minor Planet Center – is at least 97.7%. If all 11 discovery candidates are confirmed as real, then the purity of the submitted sample could be as high as 100%. This demonstrates that, at a practical level, THOR is ready to be used for discovery and reporting of new discoveries with a high degree of confidence.

3.5 Summary and Future Work

We present an observer- and cadence-independent asteroid linking algorithm, which we name "Tracket-less Heliocentric Orbit Recovery" (THOR), and show its performance on both simulated and real detections. We have shown that THOR can link Main Belt asteroids and more distant populations at high completeness and at moderate computational cost. On two

[8] https://ssd.jpl.nasa.gov/sbdb.cgi

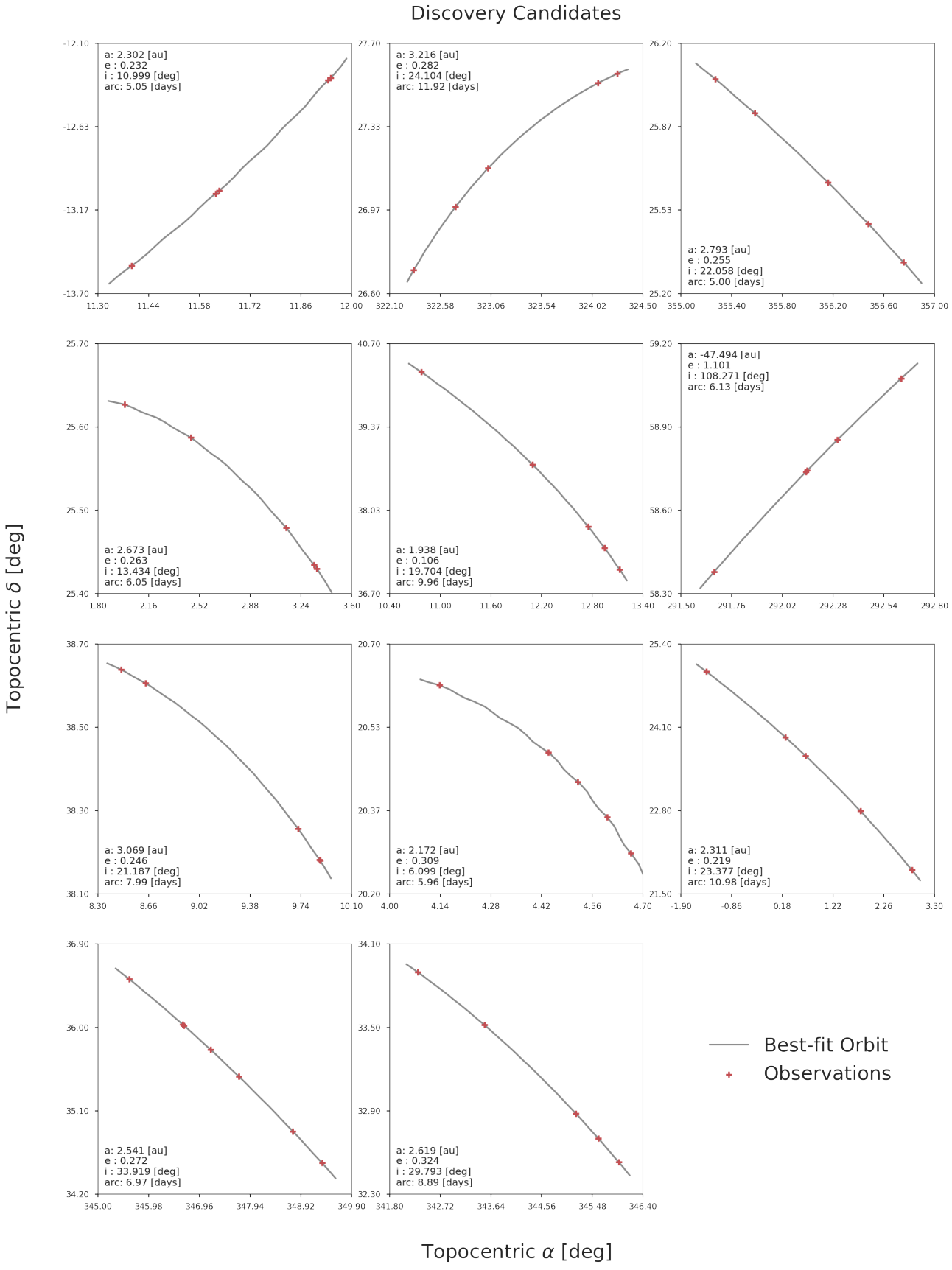
Discovery Candidates
a: 2.302 [au]
e : 0.232
i : 10.999 [deg]
arc: 5.05 [days]
a: 3.216 [au]
e : 0.282
i : 24.104 [deg]
arc: 11.92 [days]
a: 2.793 [au]
e : 0.255
i : 22.058 [deg]
arc: 5.00 [days]
a: 2.673 [au]
e : 0.263
i : 13.434 [deg]
arc: 6.05 [days]
a: 1.938 [au]
e : 0.106
i : 19.704 [deg]
arc: 9.96 [days]
a: -47.494 [au]
e : 1.101
i : 108.271 [deg]
arc: 6.13 [days]
a: 3.069 [au]
e : 0.246
i : 21.187 [deg]
arc: 7.99 [days]
a: 2.172 [au]
e : 0.309
i : 6.099 [deg]
arc: 5.96 [days]
a: 2.311 [au]
e : 0.219
i : 23.377 [deg]
arc: 10.98 [days]
a: 2.541 [au]
e : 0.272
i : 33.919 [deg]
arc: 6.97 [days]
a: 2.619 [au]
e : 0.324
i : 29.793 [deg]
arc: 8.89 [days]
Topocentric δ [deg]
Topocentric α [deg]
Best-fit Orbit
Observations

Figure 3.11 The observations of the 11 discovery candidates identified in Section 3.4.4.2 are plotted in red, with the sky-plane motion of their best-fit orbits plotted as lines. Ten of the objects show MBA-like best-fit orbit solutions. The remaining object has a hyperbolic orbit solution and corresponds to recovery observations of the hyperbolic comet C/2018 U1. The "wiggles" apparent in some of the best-fit orbit lines are due to the motion of the observer (topocentric motion). This figure was generated using plots_ztf.ipynb.

weeks of simulated data, THOR recovered 91.3% of the 18,332 that were ideally findable, whereas a tracklet-based linking algorithm would have recovered none. On two weeks of ZTF data, THOR linked 97.2% of the 21,542 objects with at least five detections (a factor of ∼2 recovery increase over MOPS and a factor of ∼1.5 increase over ZMODE). THOR recovered orbits for 97.4% of objects beyond 1.7 au, with 98.4% of objects recovered beyond 2.5 au. Furthermore, by comparing the 2018 sample to the catalog of orbits as presently known (April 2021), we show that the lower limit on purity of THOR submissions to the MPC would be 97.7% and—assuming all candidates shown in Figure 3.11 are confirmed as real—possibly as high as 100%. This, in combination with its capability to discover objects regardless of cadence or observer, renders it immediately useful for Main Belt and KBO searches on survey data and archival data sets.

While THOR can be applied to running surveys, application to archival data sets – such as ZTF, CSS, or PanSTARRS – is interesting as well. The most straightforward way to do this would be to run a sliding ∼two-week window over the data set, from beginning to the end, running THOR at each window instance. For each window run, orbits could be computed for the discovered objects and projected to the past (and future) for discovery of additional observations (and improvement of orbital solutions). This is the approach we ourselves plan to take with the ZTF archival data. Assuming the 11 objects discovered here are representative of remaining undiscovered objects in ZTF, this search would likely yield on order of 1,000 asteroids. The discovery potential with deeper archival data sets (e.g., DES data or the DECam archive) is likely to be significantly larger.

3.5.1 Future Directions

There are several avenues for improvement, which we discuss here below.

Test Orbit Selection: As discussed previously, THOR can discover objects at very high completeness in the Main Belt and beyond using the current test orbit selection criteria. Those selection criteria are, however, not ideal for the inner Main Belt, NEOs, objects interior to Earth's orbit, or objects on generally unusual orbits (e.g., ISOs). For the linking algorithm to be universally viable, these populations need to be reliably discoverable. The focus of this paper has been to demonstrate that the overall approach works even with a rather simplistic orbit selection algorithm; developing an optimal test orbit allocation algorithm is the focus of follow-up work and is currently in progress.

Hyperparameter Optimization: Throughout the description of the algorithm's performance on both the simulated survey and on two weeks of ZTF alerts, we stated the values of the hyperparameters used. These values were decided on after numerous trial runs, early testing, and optimizations. We plan to quantify in much greater detail how THOR performs as these parameters change, and how these parameter might change with different combinations of data sets and surveys. This includes automated data-driven selection of the clustering radius, the velocity grid, and its granularity, as well as decisions on the width of the time window over which to attempt linking.

Orbit determination: With our goal of extending THOR to the NEO population will likely come the need to make improvements to the orbit determination component of THOR. Improvements such as iteration at the IOD stage, as opposed to delegating the work directly to the differential corrector, could be useful here. We have also started researching the possibility of bootstrapping orbit determination directly from the corotating reference frame of the test orbit. The test orbit effectively serves as the initial orbit guess for the orbits of objects contained in the tangent plane centered on the motion of the test orbit. The relative velocities (determined by the Hough transform), along with the location in the tangent plane, should adequately constrain four of six orbital parameters. The remaining two can be fit

for at lower computation cost than performing initial orbit determination and differential correction on sky-plane coordinates.

Scalability and "Discovery-as-a-Service": We plan to deploy THOR as a scalable cloud-based service to which observations can be submitted for automated identification of Solar System objects, with candidate discoveries returned to the user and/or submitted to the Minor Planet Center. This would effectively provide the Solar System community with a "discovery-as-a-service" platform. The advantage of such a cloud-based platform is that it can elastically scale to meet THOR's computational needs as data sets are submitted. Additionally, a centralized discovery service would enable cross-survey and cross-data set discovery.

Chapter 4

SEARCHING THE NOIRLAB SOURCE CATALOG

4.1 Introduction

Present-day surveys such as the Panoramic Survey Telescope and Rapid Response System (Pan-STARRS) (Chambers et al., 2016), Catalina Sky Survey (CSS) (Larson et al., 2003), the Asteroid Terrestrial-impact Last Alert System (ATLAS) (Tonry et al., 2018), and the Zwicky Transient Facility (ZTF) (Graham et al., 2019) have discovered ∼10,000s of new asteroids over the course of the last few years increasing the number of known minor planets to nearly 1.2 million. In the near future, the Vera C. Rubin Observatory's Legacy Survey of Space and Time (LSST) (Ivezić et al., 2019), expected to begin in July 2024, is predicted to increase the number of known minor planets by about five million over the course of its ten-year survey (Jones et al., 2018) – LSST will discover more minor planets in a decade than have been discovered in the preceding 220 years that started with the discovery of Ceres by Giuseppe Piazzi in 1801. The number of discoveries per population varies greatly, with an order of magnitude increase expected in the number of trans-Neptunian objects and Jupiter Trojans, and a factor of few increase expected in the Main Belt and Near-Earth Object populations (Jones et al., 2018). Such an increase in the number of known minor planets will provide an unparalleled view of the dynamical history of the Solar System.

Asteroids and comets are typically discovered by observing "tracklets": an intranight combination of two or more observations that constrain the direction and rate of motion of potential moving objects. Surveys such as Pan-STARRS, Catalina Sky Survey, and ATLAS, aim to observe triplets or quads: tracklets made of three or four observations, respectively. The Vera C. Rubin Observatory will require at least three pairs (tracklets made of at least two observations) over the course of a linking window, typically 15-30 days. Tracklets are observed to reduce the total number of combinations of observations that would need to be tested for the presence of moving objects. Once observed, tracklets are stored and linked across nights to form candidate tracks: internight combinations of typically two or more tracklets. LSST will need three tracklets, each observed on a unique night within the nominal linking window, to go through the tracklet-to-track creation process (Jones et al., 2018).

Work by Holman et al. (2018), has shown that substantial computational performance and linking efficiency can be gained by shifting the frame of linking to the heliocenter, where minor planets would appear to move along great circles. As part of the Vera C. Rubin Observatory's Solar System Pipelines, LSST will discover minor planets using a purpose-built modification to the HelioLinC algorithm called HelioLinC3D (Eggl et al., 2020).

Requiring tracklets to discover minor planets has two main consequences. First, the requirement to observe and construct tracklets puts a strong constraint on the cadence of a survey – to discover minor planets surveys *must* revisit the same field multiple times in a night. Second, any dataset that was not constructed with a tracklet building cadence is a dataset that is not suited for Solar System small body discovery – the vast majority of astronomical datasets!

"Tracklet-less Heliocentric Orbit Recovery" (THOR), introduced in Moeyens et al. (2021) and described in Chapter 3, is a small body discovery algorithm capable of discovering minor planets without using tracklets. Moeyens et al. (2021) showed that when applied to two weeks of observations from the Zwicky Transient Facility (ZTF) (Bellm et al., 2019; Graham et al., 2019; Masci et al., 2019), THOR was able to recover 97.4% of known objects with $a > 1.7$ au. Due to cadence constraints alone, a tracklet-based algorithm such as HelioLinC could have at best recovered 2x fewer objects, while ZTF's in-house algorithm, ZMODE (Masci et al., 2019), could have at best recovered 1.5x fewer objects. In addition to recovering 97.2% of all minor planets with five or more detections, had THOR been running as ZTF's linking algorithm when the observations were made in 2018, it would have submitted at least 488 high-quality discovery candidates to the Minor Planet Center. Using a catalog of known orbits from 2021 to re-attribute the ZTF observations, the authors showed that of the 488 high quality candidates submission candidates, 477 of the orbits corresponded to 476 minor planets that had been discovered since the observations were made in 2018. Of the remaining 11 orbits, 10 were identified as new discovery candidates with the one remaining orbit representing precovery observations of comet C/2018 U1. The observations found by THOR were taken six to eight weeks prior to the discovery of the $e \sim 1$ comet by

the Mt. Lemmon Survey. Had THOR been deployed as ZTF's linking algorithm, it would have enabled ZTF to claim the discovery of this comet in addition to discovering at least 476 new minor planets.

Despite its success on ZTF data, there remain several unaddressed concerns surrounding THOR as a potential successor of tracklet-based discovery algorithms. First, THOR has yet to be extended to the $a < 1.7$ au population. Second, compared to the LSST, ZTF is on average three-four magnitudes shallower, and hence, has significantly fewer observations and lower source density. In this work, we focus on addressing the latter and leave the work of extending THOR to the $a < 1.7$ au for a future work.

The second data release of the NOIRLab Source Catalog (Nidever et al., 2021) contains some 68 billion measurements of 4.3 billion unique co-added objects. The catalog has extracted observations from over 400,000 publicly available images taken by three observing facilities: CTIO (4m Blanco + DECam), KPNO (4m Mayall + Mosaic3), and Steward Observatory (2.3m Bok + 90Prime). The majority of the exposures (340,000+) were taken using DECam on the 4m Victor M Blanco at the Cerro Tololo Inter-American Observatory (CTIO) during the Dark Energy Survey (Abbott et al., 2018). Images taken by DECam on the 4m Victor M Blanco can reach a limiting magnitude of $r \sim 23.5$ magnitudes, approaching – but still one magnitude shallower than – the expected depth of the LSST. Therefore, the NSC represents an excellent candidate to test THOR in the regime of higher source density. Additionally, the cadence of the images taken during the Dark Energy Survey is such that it does not consistently enable the construction of tracklets making the NSC an excellent candidate to test and validate tracklet-less discovery.

The B612 Foundation is a private, charitable, non-profit organization dedicated to protecting Earth from asteroid impacts. In 2018, the Asteroid Institute, a program of the B612 Foundation, launched the Asteroid Discovery, Analysis, and Mapping (ADAM) platform. ADAM is a cloud-based astrodynamics platform designed to enable computationally intensive research in planetary science and planetary defense. In Section 4.2 we briefly describe the ADAM platform and detail the work done to deploy THOR on ADAM. We name this

discovery software and hardware ecosystem ADAM::THOR, the beginnings of a cloud-based asteroid discovery service. In Section 4.3 we describe two new open-source Python packages called `precovery` and `cutouts`, which were built to further enhance the prototype discovery service. In Section 4.5 we introduce the NOIRLab Source Catalog and describe the results of our initial search using ADAM::THOR. Finally, in Section 4.6 we detail the ongoing work and future plans to prepare ADAM, and ADAM::THOR, for the LSST dataset.

4.2 Asteroid Discovery, Analysis, and Mapping (ADAM)

To enable computationally intensive research into planetary science-related problems, the Asteroid Institute, a program of the B612 Foundation, has launched the Asteroid Discovery, Analysis, and Mapping (ADAM) platform. ADAM has the capability to run Monte Carlo impact probability calculations for potential impactors, it exposes a variety of propagators such as Systems Tool Kit (STK)[1] and Find_Orb[2], and maintains a variety of standardized datasets. Additionally, ADAM has two recently released services with a website frontend: an orbital element converter built on code contained in THOR, and a precovery service that exposes the `precovery` and `cutouts` codes as a service.

4.2.1 ADAM::THOR

Due to ADAM's innate scalability and its ability to deliver significant computational power, we elected to deploy THOR on the ADAM platform. We now describe the work done to enable discovery searches with the THOR code on ADAM. We name this software/hardware ecosystem ADAM::THOR – the beginnings of a cloud-based asteroid discovery service.

In Figure 4.1 we show the cloud-based architecture of the ADAM::THOR service. Discovery searches with THOR are launched using a Cloud Run hosted API. Cloud Run is a Google Cloud service that allows users to host containerized applications (such as via Docker) in the cloud (the "thor-api" component). The containerized application for "thor-api" is built using

[1]See https://www.ansys.com/products/missions/ansys-stk.

[2]See https://www.projectpluto.com/find_orb.htm.

FastAPI[3]. The "thor-api" allows users to upload discovery search inputs, submit discovery searches (i.e., jobs), query the status of THOR jobs, and finally, download the results of the THOR job once it is completed. The inputs to a typical discovery search contain the observations file, a file containing the test orbits to use for the search, and finally a configuration file that sets the parameters with which the THOR algorithm should run. The API uploads the inputs to Google Cloud Storage and splits the input test orbits file into a single "task" per test orbit. Each task is added to a publication and subscription queue (a.k.a, pub/sub). The pub/sub queue is managed by RabbitMQ[4] and is hosted on a dedicated virtual machine (the "rabbit" VM).

Once all the tasks have been added to the queue and the inputs to each task have been appropriately separated into separate sub-directories in the Google Cloud Storage Bucket, then processing can begin. The width of the distributed application is managed by an autoscaler VM. In Figure 4.1 the autoscaler VM is called "thor-production-autoscaler". The autoscaler monitors the pub/sub queue and will create (spawn) worker VMs as a function of the size of the queue. The maximum width of the distributed application is set by a configuration setting and is typically limited by Google Cloud Platform quotas. Each worker VM is spawned using a Docker image that contains the THOR code and the required software depedencies to run the pipeline. The version of the THOR code deployed to the VMs is set by a configuration parameter, and so it is conceivable to run different versions of the THOR code across different deployments. For example, ADAM::THOR currently has two deployments: "production-tasks" which is depicted in Figure 4.1, and "staging-tasks" which is not shown in the figure but is an identical secondary queue which can run simultaneously with "production-tasks". The gray-dashed boxed marks the components of the architecture that are managed by a Python package called `thorctl`[5]. `thorctl` was purpose built to manage THOR discovery searches across a parallel backend and was built

[3]See https://fastapi.tiangolo.com/.

[4]See https://www.rabbitmq.com/.

[5]See https://github.com/B612-Asteroid-Institute/thorctl

in-house to distribute the THOR algorithm over a series of virtual machines hosted by Google Compute Engine. It should be noted that `thorctl` can be generalized to work with other cloud-service providers such as Amazon Web Services or Microsoft Azure.

In its default configuration, each worker VM has 32 virtual CPUs and 128 GB of RAM (the "e2-standard-32" machine family). The maximum number of workers that can be spawned is currently set to 48 (this number is determined by a quota on the number of cores). Hence, during a fully-scaled discovery search, the number of cores and the amount of RAM that is available are 1536 vCPUs, and 6144 GB, respectively. Since each task is defined to be a single test orbit, a total of 48 test orbits can be processed simultaneously, with the individual THOR pipeline components using up to 32 cores on each individual machine. For example, a single THOR task at the clustering stage will distribute that clustering work to each core available on the machine. The type of machine each worker VM should be deployed on is determined and set by configuration files within `thorctl`. These configurations can be changed arbitrarily depending on the use case, cost allowances, and user preference.

As each task is completed, the results of that task, namely the output orbits and their observation members are saved and stored in dedicated Buckets for the task. Once all the jobs have been completed (processing either exited successfully, or processing exited due to crashes of the underlying software), then the results of each task (if available) are aggregated by the API and stored in a separate cloud storage Bucket. If any tasks did not return results whether that be due to no objects being found or due to problems with the software, these tasks are marked as such and can be analyzed later. The aggregated results can then be downloaded by the user and analyzed for completeness and searched for discovery candidates.

We stress that ADAM::THOR is currently still in heavy development and that many of the characteristics and details of the architecture described here may change as we continue to build out the platform. Our goal from the beginning has been to build ADAM::THOR to process the full LSST data volume and to simultaneously offer ADAM::THOR to the community as a discovery service – any changes in the architecture and the underlying software will be completed with this goal first and foremost in mind.

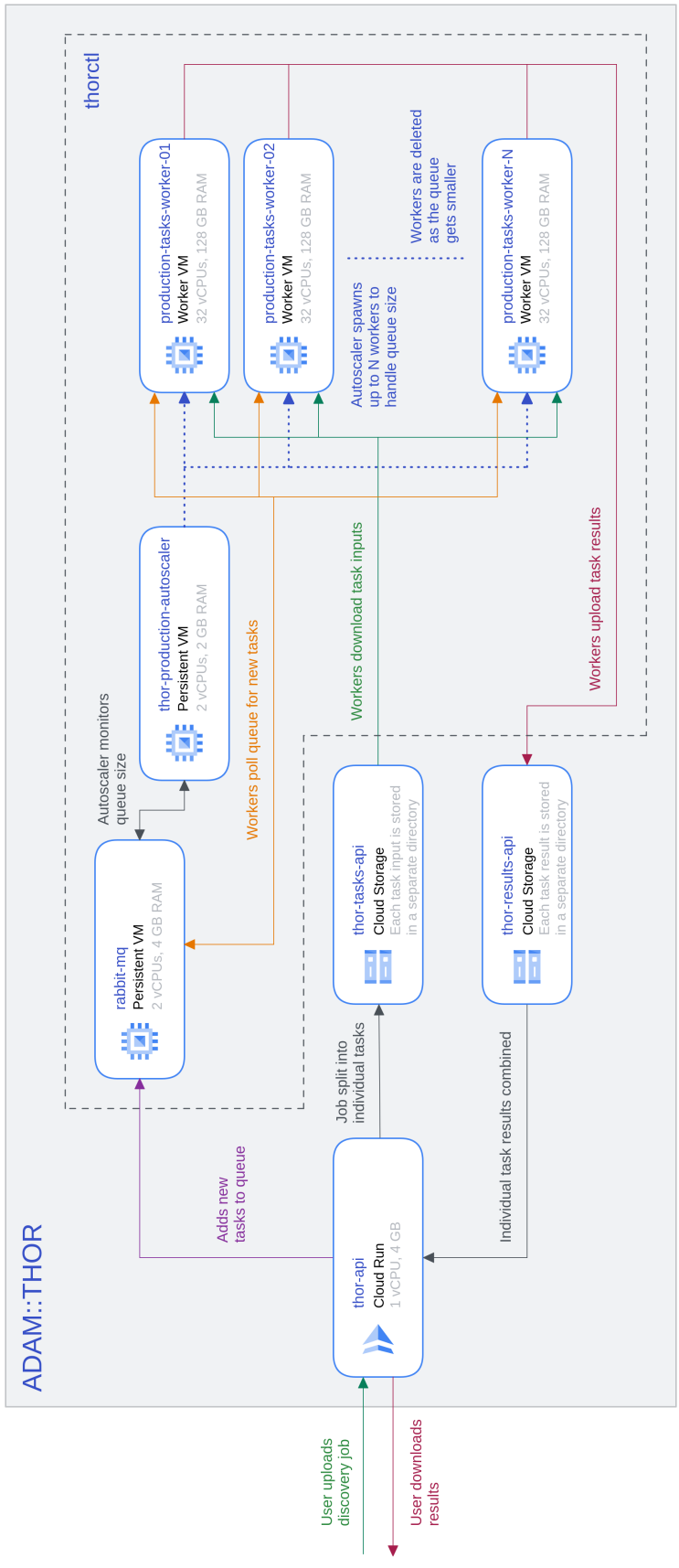
ADAM::THOR
thorctl
User uploads discovery job
User downloads results
thor-api
Cloud Run
1 vCPU, 4 GB
rabbit-mq
Persistent VM
2 vCPUs, 4 GB RAM
thor-production-autoscaler
Persistent VM
2 vCPUs, 2 GB RAM
thor-tasks-api
Cloud Storage
Each task input is stored in a separate directory
thor-results-api
Cloud Storage
Each task result is stored in a separate directory
production-tasks-worker-01
Worker VM
32 vCPUs, 128 GB RAM
production-tasks-worker-02
Worker VM
32 vCPUs, 128 GB RAM
production-tasks-worker-N
Worker VM
32 vCPUs, 128 GB RAM
Adds new tasks to queue
Job split into individual tasks
Individual task results combined
Autoscaler monitors queue size
Workers poll queue for new tasks
Workers download task inputs
Workers upload task results
Autoscaler spawns up to N workers to handle queue size
Workers are deleted as the queue gets smaller

Figure 4.1 ADAM::THOR is built on Google Cloud Platform (GCP) using services such as Cloud Run, Cloud Storage, and Compute Engine Virtual Machines (VMs) to scale to the needs of a THOR discovery search. In brief, the user uploads a THOR discovery job which contains test orbits and observations. The job is divided into individual tasks (each test orbit is a single task) and added to a queue. The queue is monitored by an autoscaler that will launch additional "Worker" VMs as a function of the length (size) of the queue. Each Worker VM pulls a task from the queue, its corresponding input data, and begins processing. The results of each task, namely the output orbits and other data products, are stored in a separate Google Cloud Bucket. Once the job is completed, the results are aggregated and then returned to the user. This figure was generated using Google Cloud Architecture Diagramming Tool.

4.3 THOR Enhancements

In this section, we describe several of the key enhancements that were made to the THOR algorithm to enable preliminary processing of the NOIRLab Source Catalog. That is, we describe the changes made to the core of the discovery algorithm which is deployed at scale as the ADAM::THOR service.

Moeyens et al. (2021) described the efforts undertaken to run THOR on two weeks of observations from the Zwicky Transient Facility (ZTF). In a two-week period from 2018 September THOR found 97.2% of all known objects with five or more observations, including ten discovery candidates and precovery observations of a comet. Motivated by experiments running THOR on the NOIRLab Source Catalog, we added and modified several components of the THOR code to enable searches in these data.

4.3.1 Hotspot_2D: Faster clustering

One of the bottlenecks identified in Moeyens et al. (2021) is the Hough transform component. In THOR, the projected and transformed coordinates are fundamentally three-dimensional spanning θ_X-θ_Y-t space. By leveraging the time dependence of the two positional dimensions, that is, by assuming a set of velocities ($d\theta_X/dt$,$d\theta_Y/dt$) in θ_X, θ_Y, the 3D Hough transform is reduced into a 2D clustering problem.

In practice, the 2D clustering problem is performed for each combination of assumed velocities. In its default configuration, THOR uses 90,000 different velocity combinations ranging between -0.1 and 0.1 deg day^{-1} in 300 bins per coordinate. Once the projected observations have been shifted with the assumed velocities, `DBSCAN` (Ester et al., 1996) is used to extract clusters, and in turn, to find potential moving objects.

Instantiating 90,000 instances of `DBSCAN` is computationally expensive, made even more expensive as the number of observations increases. For example, Moeyens et al. (2021), stated that the clustering component for a single test orbit could take several hours while the remaining pipeline components would take tens of minutes or less.

We developed a different approach to the clustering problem which we call "hotspot_2d". In brief, "hotspot_2d" bins the transformed and shifted observations on a common 2D grid, and looks for occurrences where multiple observations fall in the same bin creating a "hotspot". As this approach relies purely on array operations using `numpy` (Harris et al., 2020), we have found that it is at minimum an order of magnitude faster than performing clustering with `DBSCAN` which bears an upfront initialization cost.

To test the efficacy of the new clustering algorithm we repeated the analysis described in Section 4.4.2 of Moeyens et al. (2021). We reprocessed the ZTF alert data set using the latest version of the THOR code with both `DBSCAN` and "hotspot_2d"[6]. Using the 2021 April catalog of known orbits described in the aforementioned section, THOR using `DBSCAN` recovers 21,459 of the 22,091 objects deemed findable (97.14% completeness). THOR using "hotspot_2d" recovers 21,446 objects of the 22,091 objects deemed findable (97.08% completeness). The "hotspot_2d" variation finds 13 fewer objects but took 2 days less processing time.

4.3.2 Test Orbit Selection: HEALPix

In Moeyens et al. (2021) test orbits were selected by subdividing the sky into 15 x 15 deg patches, and selecting up to nine test orbits from the known population in bins of semi-major

[6]We note that there were several optimizations made to the THOR code since the work described in Moeyens et al. (2021) but none of substantial impact to the overall performance of the code.

axis (for details see Section 2.2 in Moeyens et al. (2021)). Dividing the sky into patches as described introduces issues at the poles where the size of the patches decreases with the polar angle. Instead, we changed the test orbit selection algorithm to use HEALPix to subdivide the sky into pixels of equal area, and then for each pixel test orbits were selected as before. The default configuration is to select test orbits using *nside*=16, or equivalently, to select up to nine test orbits per pixel that are 13.6 $\deg^2$ size.

4.4 Pipeline Enhancements

In this section, we summarize the functionality of two software components that were developed to further improve the ADAM::THOR discovery pipeline.

4.4.1 Precovery

"Precovery" is the process of finding observations of minor planets in past images. An asteroid may not be discovered until sufficient observations have been made within the span of a nominal linking window. However, that same asteroid may have been previously observed but not identified as a new object because there were insufficient observations for a linking algorithm to discover it. Or, perhaps, the specific observing program was not actively conducting Solar System small body discovery searches but still observed the asteroid in question. These observations are still important to identify as they allow the observational arc of the object to be extended without relying on possibly observing the object again in the future. Extending the observation arc decreases the uncertainty of the asteroid's orbit which improves predictions on when it can be observed again. Additionally, identifying observations of asteroids that have been discovered also reduces the number of observations that remain to be processed in subsequent discovery searches.

We note that "attribution": the process of identifying known objects in astronomical images or catalogs and "precovery" are closely related. In this work, we consider "attribution" and "precovery" to be functionally equivalent – the same code used for "precovery" can also be used for "attribution". In the literature, the attribution problem is often described as

the process by which observations are attributed to already known objects (see, for example, Milani et al. (2001)), whereas precovery is the process by which observations of discovered objects are found in past images. In both cases, knowledge of the uncertainty of the known object's orbit and the propagation of that uncertainty is vital to determine if an observation could belong to the orbit. In the absence of uncertainty propagation, precovery searches are typically limited to short time spans around the discovery epoch for uncertain orbits. The time span can be extended for objects with highly certain orbits, such as the population of minor planets that are "numbered". In practice, running a precovery search also involves running orbit determination algorithms to iteratively update the orbit fit as new observations are found and to identify possible outliers.

We developed an open-source Python package called `precovery` (Nelson et al., 2022), specifically to find observations of objects in large astronomical datasets (volumes on the order of TBs of data). A critical design criterion is that these searches need to be fast and parallelizable. To accomplish this, `precovery` restructures input catalogs of observations into two components: an index or table containing information on the exposures contained within a catalog, and a series of binary files that contain the point source measurements. The input exposures are mapped to a common HEALPix grid of user-defined size which allows multiple datasets to be mapped to a dataset-agnostic grid and thereby allows these datasets to be searched simultaneously. The "index" stores information such as the HEALPixels visited, the midpoint time of the exposures, the observatories from which the exposures were made, the dataset from which the exposures were indexed, and the filters. Critically, the index also stores the URI of the binary file, the starting location of where the observations are stored in said binary file (the offset), and finally, the length of the data. The process of restructuring an input catalog into this dual-type format we call "indexing", and bears a one-off upfront cost. In the case of the NOIRLab Source Catalog, indexing the 3.3 billion observations downloaded as described in the next section took 2.5 days on a single core.

Once the data has been "indexed", the precovery code is then able to start searching for potential matches given an input orbit. `precovery` optimizes the search solution by

first dividing the data into windows of time. An input orbit is propagated using N-body propagation to the midpoint of each window. From there the orbit is 2-body propagated to each time an exposure occurs. If the 2-body propagated orbit intersects any of the HEALPixels within that window, these HEALPixels are marked as "FrameCandidates". Once the FrameCandidates have been identified, the orbit is then N-body propagated from the window midpoint to the time of each candidate exposure. Only then are the actual observations of the exposures read into memory. The predicted position of the object within that frame is then compared to all observations within the frame, and if any fall within a user-defined angular tolerance, then these observations are marked as "PrecoveryCandidates" and returned to the user. Instances where observations could have occurred, namely, when a trajectory intersects a HEALPixel but no observations fall within the angular threshold, can also be returned to the user. There is a significant amount of information that can be gleaned from exposures where an observation could have occurred but none were found – cosmic rays, interference with background sources, and source saturation, can all be identified once the individual postage stamps are visualized (see Figure 4.2 in the next section for an example). Orbit propagations and ephemeris generation in `precovery` are performed using `oorb` (Granvik et al., 2009).

In the context of ADAM::THOR, `precovery` is used to find additional observations of initially short discovery arcs within an extended search window that is centered about the linking window. This is accomplished through an iterative process: each orbit is run through the `precovery` code looking for any observations that are within 1" of the predicted position of the object ±60 days from the orbit's initial epoch. The initial epoch typically falls somewhere within the initial search window. Any newly found observations are then refitted with the differential corrector (see Section 2.1.5 in Moeyens et al. (2021) for details). The observations that passed OD are marked as belonging to the candidate. In the next iteration, the search window is increased to ±120 days from the initial epoch, and the orbit is differentially corrected again when new observations are found. This entire process is repeated until the first two or three years of the dataset are searched for each orbit. In the absence of complete

integration of the `precovery` into the ADAM::THOR pipeline, we initially chose to limit the search to the first few years due to computational constraints. In practice, the `precovery` and OD iterations are currently performed on a local system as an afterburner analysis step. In full operations, this iterative process will also be run as part of ADAM::THOR and will likely replace the orbit merging and extension code currently in the THOR pipeline.

The precovery code is capable of running a single search through all seven years of the NOIRLab Source Catalog (NSC) dataset filtered as described in Section 4.5.1 in 35 s. At the time of writing, the Asteroid Institute is working on exposing precovery for large batch jobs (>=1000s of orbits) to the community. Upon completion, this service will be used to find all observations for the 1.2 million known orbits contained within the Minor Planet Center's catalog of known orbits. These observations, if not already maintained by the MPC, will be submitted, and furthermore, will be used to analyze THOR's performance as we work to process the full NSC dataset.

4.4.1.1 Challenges and Future Work

As the precovery code has become a critical component in the ADAM::THOR discovery pipeline, we now describe some of the computational challenges and design considerations of the code that we aim to address in the future.

Error Propagation: Orbital uncertainties, typically represented as covariance matrices, are currently unsupported. To make robust attributions of observations to orbits, the uncertainty of the state on its predicted position at the time of the observations should be taken into account. There are a variety of methods to propagate orbital uncertainties: brute-force Monte Carlo propagations via sampling of the covariance matrix (as are currently supported in ADAM), more robust sigma-point propagation techniques, and in the simplest case, analytic 2-body propagation using the error state transition matrix or the variational equations.

High *nside* Indexing with Downsampling: We experimented with a variety of different assumed values of *nside* for indexing. Our experiments indicated that *nside* of 32 is

a performant enough configuration for the NSC data. However, the size of the individual pixels is 3.3 $\deg^2$. In this instance, if the user requests "FrameCandidates" be returned, i.e., those HEALPix-mapped exposures where an observation could have occurred but did not, the number of candidates returned can be many tens of times the number of pixels where an observation did occur. In other words, the larger the pixels the more likely they will be intersected by the input trajectory even though no observations are nearby the predicted position of the object. A method to optimize this search would be to index input catalogs at a significantly higher *nside*, and have *nside* be a search parameter. A property of nested HEALPixels is that they can be trivially downsampled to a smaller *nside* value. By indexing at a high *nside* value we can allow the user to determine the appropriate *nside* value at which the search should occur. This could even be defined as a function of source density for different indexed datasets.

Support for Varying Observation Times within an Exposure: It is currently assumed that all observations within an exposure are measured at the same time. In the case of surveys such as the LSST, each individual observation within an exposure will have a reported observation time that takes into account the motion of the shutter across the focal plane. If such a dataset were to be ingested by the precovery code with the code in its current state, each observation would be assigned a unique HEALPix frame thereby negating the benefits gained from mapping exposures to a common HEALPix index. As the number of frames approaches the total number of observations, the precovery code's computational performance approaches that of the brute-force method of propagating an orbit to each individual observation time and checking for possible intersections in observation space rather than the more spare exposure space.

Space-based Observatories: The precovery code currently does not have the means to scan observations made by space-based observatories as the underlying propagator does not have the functionality to get observatory positions for such observatories. In the future, we would need to add support for the reading and ingesting of SPICE kernels that describe the trajectories of space-based observatories.

Checkpointed and Multiprocessed Indexing: Indexing observations currently poses an upfront cost of a few days of indexing time. For example, indexing the second data release of the NSC with *nside*=32 takes 2.5 days on a single core. However, any interruptions during indexing means the entire dataset needs to be re-indexed from the start. There is currently no checkpointing. We are planning to restructure the underlying data to support multiprocessed indexing and checkpointing.

4.4.2 Cutouts

In addition to the precovery code described in the previous section, we developed an open-source Python package called `cutouts` (Moeyens & Kiker, 2022) to visualize the outputs of the precovery code. `cutouts` uses the metadata stored by `precovery` to query the appropriate image access service using the Python package `pyvo` which implements the IVOA's Simple Image Access 2.0 protocol (Dowler et al., 2015).

In Figure 4.2, we show an example of a mosaic produced by the cutouts code. The green and red crosshairs are centered on the predicted position of the object in each image. A green crosshair indicates that a point source was found within the angular tolerance during the precovery search. A red crosshair indicates that no such match was found in the image, but the images are still shown. The example mosaic is striking, in that it is clear there were several instances where this particular moving object could have been detected but was not. Note, for example, the last three images in the second to last row – 2013 RR_{163} is moving in front of a background galaxy preventing the point source extractor used during the creation of the NSC dataset from isolating the point source. The cyan arrow in each postage stamp indicates the direction of motion. It should be noted that the length of the arrow is the same in each image and is not proportional to the rate of motion. Note that the point source elongation visible in some of the images is in the direction of motion. The mosaic clearly indicates that 2013 RR_{163} is a real object. In fact, 2013 RR_{163} is one of the objects discovered by ADAM::THOR during the initial search experiment described in the latter sections of this work.

(2013 RR163)

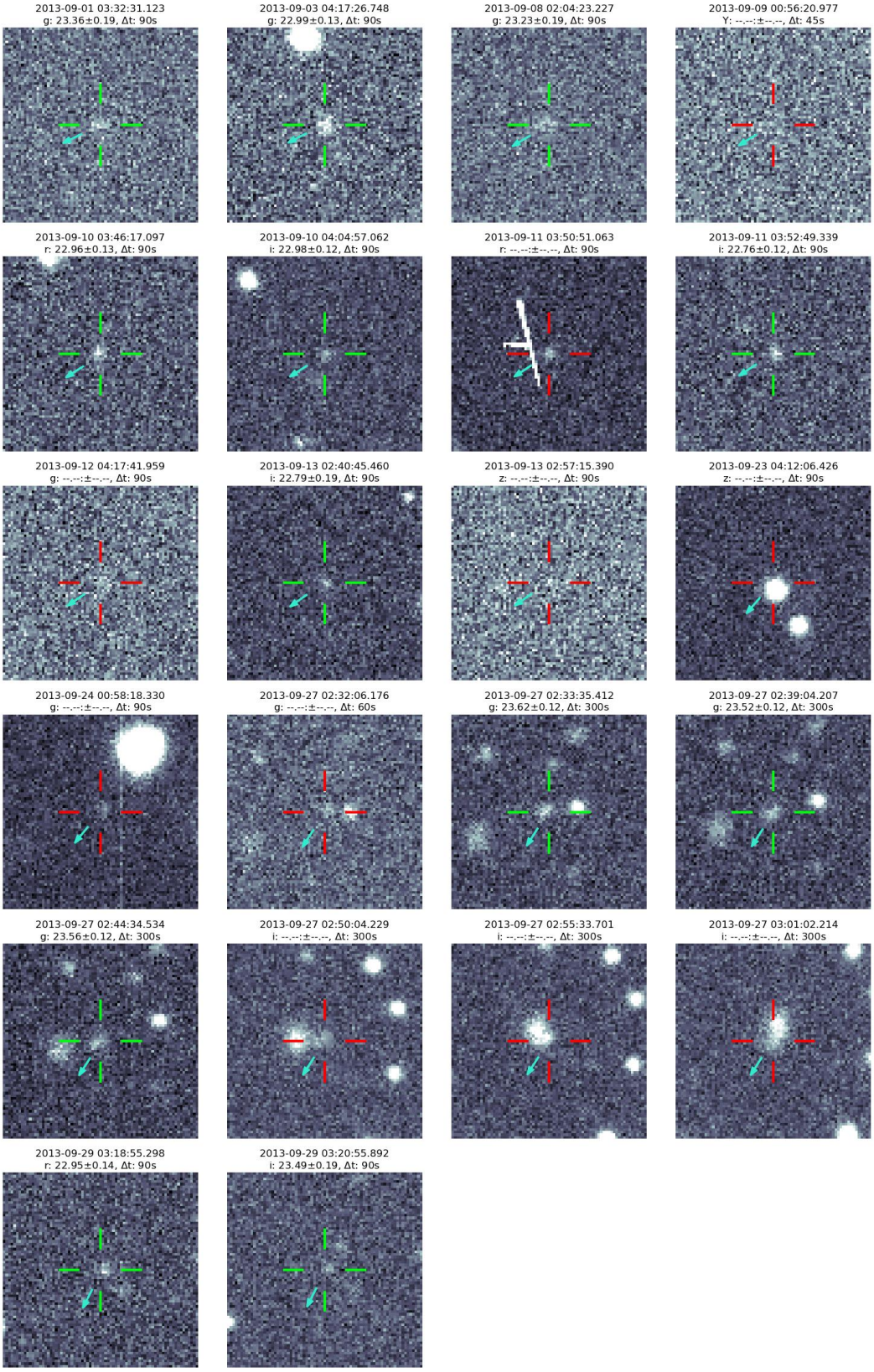
2013-09-01 03:32:31.123
g: 23.36±0.19, Δt: 90s
2013-09-03 04:17:26.748
g: 22.99±0.13, Δt: 90s
2013-09-08 02:04:23.227
g: 23.23±0.19, Δt: 90s
2013-09-09 00:56:20.977
Y: --.--:±--.--, Δt: 45s
2013-09-10 03:46:17.097
r: 22.96±0.13, Δt: 90s
2013-09-10 04:04:57.062
i: 22.98±0.12, Δt: 90s
2013-09-11 03:50:51.063
r: --.--:±--.--, Δt: 90s
2013-09-11 03:52:49.339
i: 22.76±0.12, Δt: 90s
2013-09-12 04:17:41.959
g: --.--:±--.--, Δt: 90s
2013-09-13 02:40:45.460
i: 22.79±0.19, Δt: 90s
2013-09-13 02:57:15.390
z: --.--:±--.--, Δt: 90s
2013-09-23 04:12:06.426
z: --.--:±--.--, Δt: 90s
2013-09-24 00:58:18.330
g: --.--:±--.--, Δt: 90s
2013-09-27 02:32:06.176
g: --.--:±--.--, Δt: 60s
2013-09-27 02:33:35.412
g: 23.62±0.12, Δt: 300s
2013-09-27 02:39:04.207
g: 23.52±0.12, Δt: 300s
2013-09-27 02:44:34.534
g: 23.56±0.12, Δt: 300s
2013-09-27 02:50:04.229
i: --.--:±--.--, Δt: 300s
2013-09-27 02:55:33.701
i: --.--:±--.--, Δt: 300s
2013-09-27 03:01:02.214
i: --.--:±--.--, Δt: 300s
2013-09-29 03:18:55.298
r: 22.95±0.14, Δt: 90s
2013-09-29 03:20:55.892
i: 23.49±0.19, Δt: 90s

Figure 4.2 A mosaic of "postage stamps" or "cutouts" centered on the predicted position of 2013 RR_{163} in each image. The predicted position of the asteroid in each image is indicated by the red or green crosshairs, where the color of the crosshair indicates a point source detection was found within the angular tolerance during the precovery search. A green crosshair indicates that a point source was found, whereas a red crosshair indicates a non-detection. The cyan arrow shows the predicted direction of motion in each image. If the object is real, and its orbit has been correctly calculated, we would expect to see a point source near the center of each image where the seeing and exposure time would have allowed it to be observed. Images like these can be used for quick by-eye verification of potential discovery candidates. This figure was generated using plots_nsc_dr2_initial.ipynb.

4.4.2.1 Challenges and Future Work

As in the case of `precovery` in Section 4.4.1.1, we now summarize two key challenges and upgrades that remain to be made to `cutouts`.

Multi-observatory Support: The cutouts code currently uses a user-defined fixed image width and fixed image height in units of pixels. However, different observatories have different optical properties such as varying pixel scale. Ideally, we would want the user to define an image width and height in arcseconds and support different pixel scales from different observatories.

Fast Image Access: The cutouts code will cache downloaded images to avoid making expensive repeat queries against the underlying APIs. However, download rates are still the predominant limiting factor. It can take ∼5 minutes to construct a single mosaic when downloading the images for the first time, thereby, effectively ruling out the possibility of doing this at scale. To generate mosaics for all candidate orbits would require local access to the images or faster access at the server level.

4.5 Results of Initial NSC Processing

The second data release of the NOIRLab Source Catalog (NSC) (Nidever et al., 2021) contains observations extracted from over 400,000 publicly available images from three of NOIRLab's observing facilities. The bulk of the images come from the Dark Energy Survey (Abbott

et al., 2018) and span from 2012 September to 2019 November. In Figure 4.3 we show the number of visits contained in the second data release of the NSC visualized on a HEALPix grid with *nside*=32. Note that the Dark Energy Survey "tank" is clearly visible.

4.5.1 NSC Data Set for Initial Searches

The NOIRLab Source Catalog contains 68 billion measurements of 4.3 billion objects. The catalog is constructed primarily for use as a catalog of co-added objects – objects extracted from multiple epochs of observation combined into a single source. As part of its processing, the number of detections of each unique object is stored. Since the catalog contains observations of "static" objects such as stars and galaxies, we applied a first pass filter to select the observations of objects that most likely correspond to the observations of moving objects such as asteroids. We queried for only those observations that had been detected four or fewer times in the same location ($ndet \leq 4$). This reduced the total number of observations from 68 billion to 3.38 billion observations of 1.76 billion objects. We downloaded the NSC observations using the `astrodatalab`[7] package developed by NOIRLab. This was accomplished by querying for all observations within slices of 0.05 deg in Right Ascension yielding a total of 7200 such files. We restructured the spatially divided files into temporally divided files of 31-day lengths. The total time to download and restructure the data was approximately seven days.

For our initial experiments, we selected observations dated between 2013 August 30 and 2013 September 30. The linking window was extended to 30 days rather than the typical 15 in large part due to the cadence of the underlying observing programs. Observations of the same area of the sky, which primarily come from the Dark Energy Survey, typically occur every one to two weeks which does not always guarantee that there will be enough observations of minor planets for discovery. The 30-day window contains a total of ∼18 million observations.

[7]See https://github.com/astro-datalab/datalab.

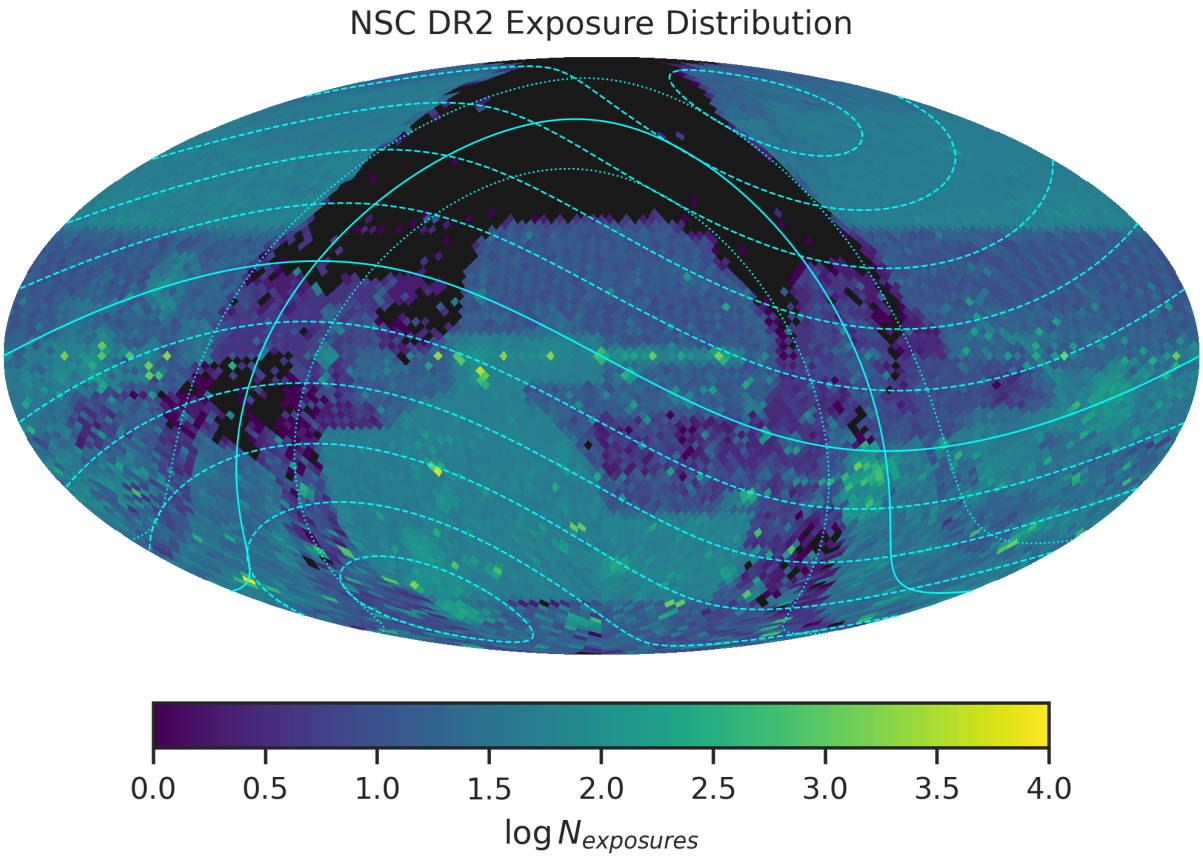

Figure 4.3 The number of exposures in different regions of the sky are visualized. NSC DR2 exposures are mapped to a HEALPix scheme with characteristic parameter *nside*=32. The ecliptic plane is plotted as a solid cyan line with corresponding dashed lines of constant ecliptic latitude at ±15, 30, 45, 60, 75 deg. The galactic plane is also plotted as a solid cyan line, with corresponding dotted lines of galactic latitude at ±15 deg. This figure was generated using plots_nsc_dr2_initial.ipynb.

In Moeyens et al. (2021), the two-week window of observations from ZTF spanned ∼20,000 deg^2 and contained ∼800,000 observations. By comparison, the 30-day window of observations from the NSC spans ∼3700 deg^2 and contains ∼18 million observations. On average, the source density is 200 times greater, with 5000 observations per square degree over the course of the 30-day window in the NSC compared to 40 observations per square degree over the course of the 15-day window of ZTF observations. The difference here is in large part due to ZTF's use of difference imaging to generate alerts and it being ∼3-4 magnitudes shallower. We further filtered the observations by mapping them to a HEALPix grid, only keeping those observations that come from exposures with the following properties:

- Ecliptic latitude ($\beta < 60$): Select exposures near to the ecliptic plane.
- 10-σ Depth >= 20: Select only those exposures that have a 10-sigma depth greater than 20 to preferentially select some of the deeper exposures which are more likely to contain undiscovered minor planets.
- Observatory and Instrument = "c4d": Limit the exposures to those made with DECam on the 4-m Victor M Blanco at CTIO.

Applying each of these criteria as a final filter reduces the total number of observations in the 30-day window to ∼3 million or about 15% of 18 million. In Figure 4.4, we plot the reduced and filtered observations. The full sample of ∼18 million observations are shown as blue points, whereas the observations contained in the reduced dataset are shown on a HEALPix map of the number of exposures per pixel.

Based on the performance of THOR in Moeyens et al. (2021), we elected to use mostly the same pipeline configuration settings. For the Hough transform component, we set THOR to use a grid of velocity combinations ranging between -0.1 and 0.1 in 300 bins deg day^{-1} for both $d\theta_X/dt$ and $d\theta_Y/dt$, yielding a total of 90,000 different velocities tested per test orbit. For the orbit determination components, the χ^2_ν and contamination thresholds were set to (10^5, 20%), (10, 20%), and (10, 0%) for initial orbit determination (IOD), the differential corrector (OD),

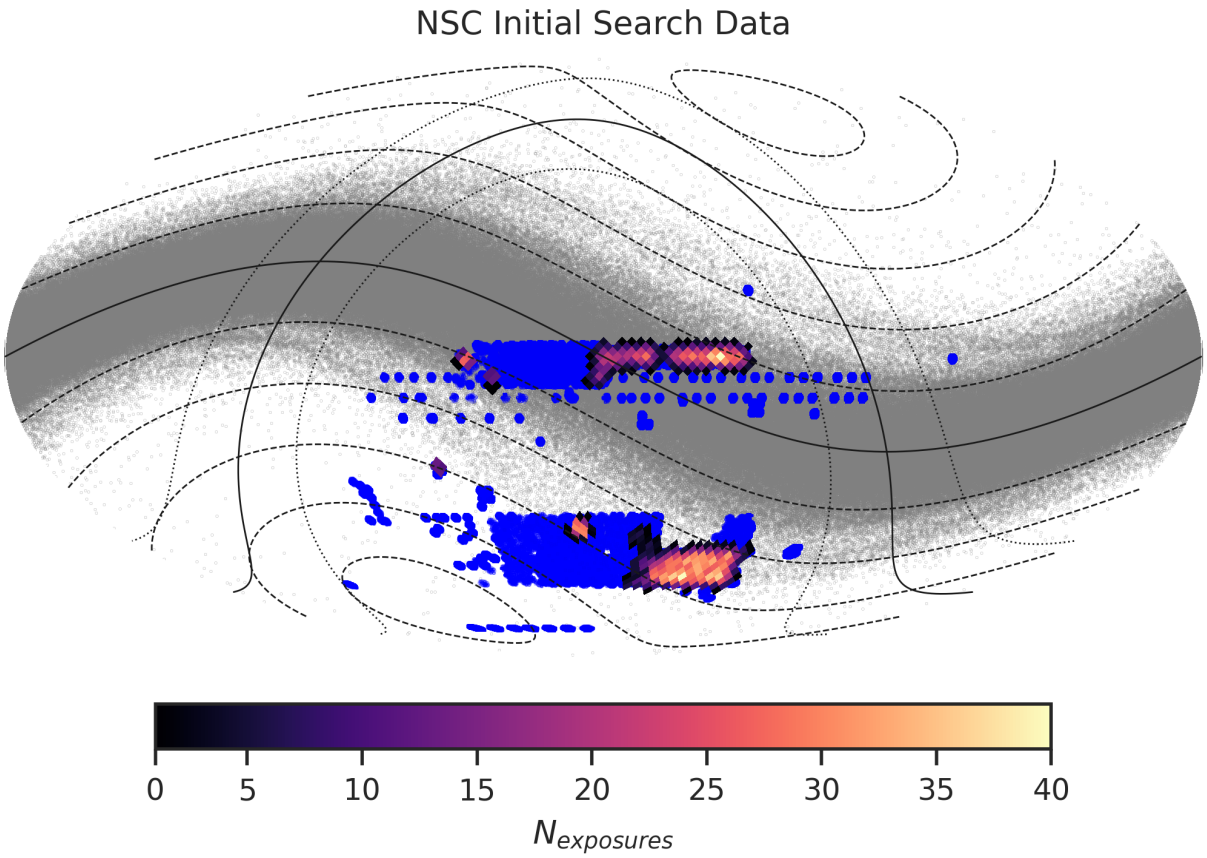

Figure 4.4 The NSC observations between 2013 August 30 and September 30 are plotted in blue. The number of visits to the reduced sample of observations are plotted on a HEALPix grid with characteristic parameter *nside*=32. The ecliptic plane is plotted as a solid gray line with corresponding dashed lines of constant ecliptic latitude at ±15, 30, 45, 60, 75 deg. The galactic plane is also plotted as a solid gray line, with corresponding dotted lines of galactic latitude at ±15 deg. This figure was generated using plots_nsc_dr2_initial.ipynb.

and the combination of the differential corrector and attribution, respectively. The maximum attribution distance for the latter component was set to one arcsecond. The minimum number of observations and minimum arc length for an orbit to be accepted were set to five observations and one day, respectively. Given that the September 2013 observations span a significantly smaller area of the sky than the ZTF alerts on which THOR was demonstrated, we elected to reduce the cell size of the gathered observations to ten deg^2. Lastly, we reduced the clustering radius from 0.005 deg or ∼18" to 10" and set the clustering algorithm to "hotspot_2d".

The NOIRLab Source Catalog observations do not have any published attributions or associations with known Solar System objects. To validate the effectiveness of THOR and to understand how well the algorithm performs, we constructed a series of labels that we could use to understand completeness. We used the attribution code included with THOR to build a set of labels for the September 2013 observations. We downloaded the 2021 June 6 catalog of known orbits from the Minor Planet Center and generated ephemerides for each orbit at each time of the observations in the 30-day window. The attribution code in THOR works as follows: in chunks of ten orbits per process, the attribution code constructs a BallTree (Omohundro, 1989; Pedregosa et al., 2011) of the predicted ephemerides and the observations for each exposure for each orbit. The BallTree can be queried for any observations that may fall to within a configurable amount of arcseconds from the predicted position of the orbit at the time of the observations. Due to the naive nature of this "brute-force" algorithm, the total processing time was 26 hours using 60 cores. To be able to process the entirety of the NOIRLab Source Catalog it became exceedingly obvious that we would need to build a more robust and significantly faster algorithm to perform attribution of this kind. Hence, the development and creation of the open-source Python package `precovery` described in Section 4.4.1.

We consider any known object with at least five observations throughout the 30 days of observations to findable, resulting in a total of 2261 such objects. Assuming a tracklet-based algorithm requiring at least a pair of observations on three individual nights, we find that

82 objects would be findable – a strong reminder of how the cadence of these observations is not suited for tracklet-based discovery. We consider any object recovered in a pure orbit with at least five observations to be found.

4.5.2 ADAM::THOR Performance

THOR, deployed and hosted on ADAM, recovered 17,464 orbits containing a mix of orbits of known objects, orbits of unknown objects, and erroneous linkages. As these orbits were aggregated by the THOR API from those tasks that completed, further cleanup and de-duplication was necessary. We do this with a two-step process: first, using the labels calculated using the attribution code, we remove all observations of known objects from orbits that contain a mix of unknown and known observations. Second, we remove any observations that occur within 30 minutes of another. This helps remove potentially static sources that remain in the observations. After both these filters are applied, any orbits that have less than five observations are removed. Next, we run the differential corrector to update the orbit fits accordingly. After the orbits have been differentially corrected, we remove any orbits that share subsets of observations in favor of the occurrence that has the lowest overall χ^2_ν. These series of filters reduce 17,464 orbits into two distinct groups: 1248 orbits containing only the observations of known objects, and 479 orbits containing only unknown observations for a total of 1727 orbits.

Of the 2261 objects with five or more observations, THOR recovered 1241 (54.8%) of these objects in pure orbits. Seven of the 1248 orbits contain a mix of known object observations, meaning the purity of the sample of known orbits is 99.4%. In the top panel of Figure 4.5, we plot completeness in bins of semi-major axis (a) and inclination (i). In the bottom panel of the same figure, we plot completeness in bins of semi-major axis (a) and eccentricity (e). In both panels, we plot the 626 test orbits used as red points, while red number density contours trace the number of objects that should be findable. The vertical dashed lines outline the bin edges in semi-major axis space used for the test orbit selection. Note that the outer Solar System bin stretches from 5.0 to 50 au, but is plotted in its entirety. In the top panel,

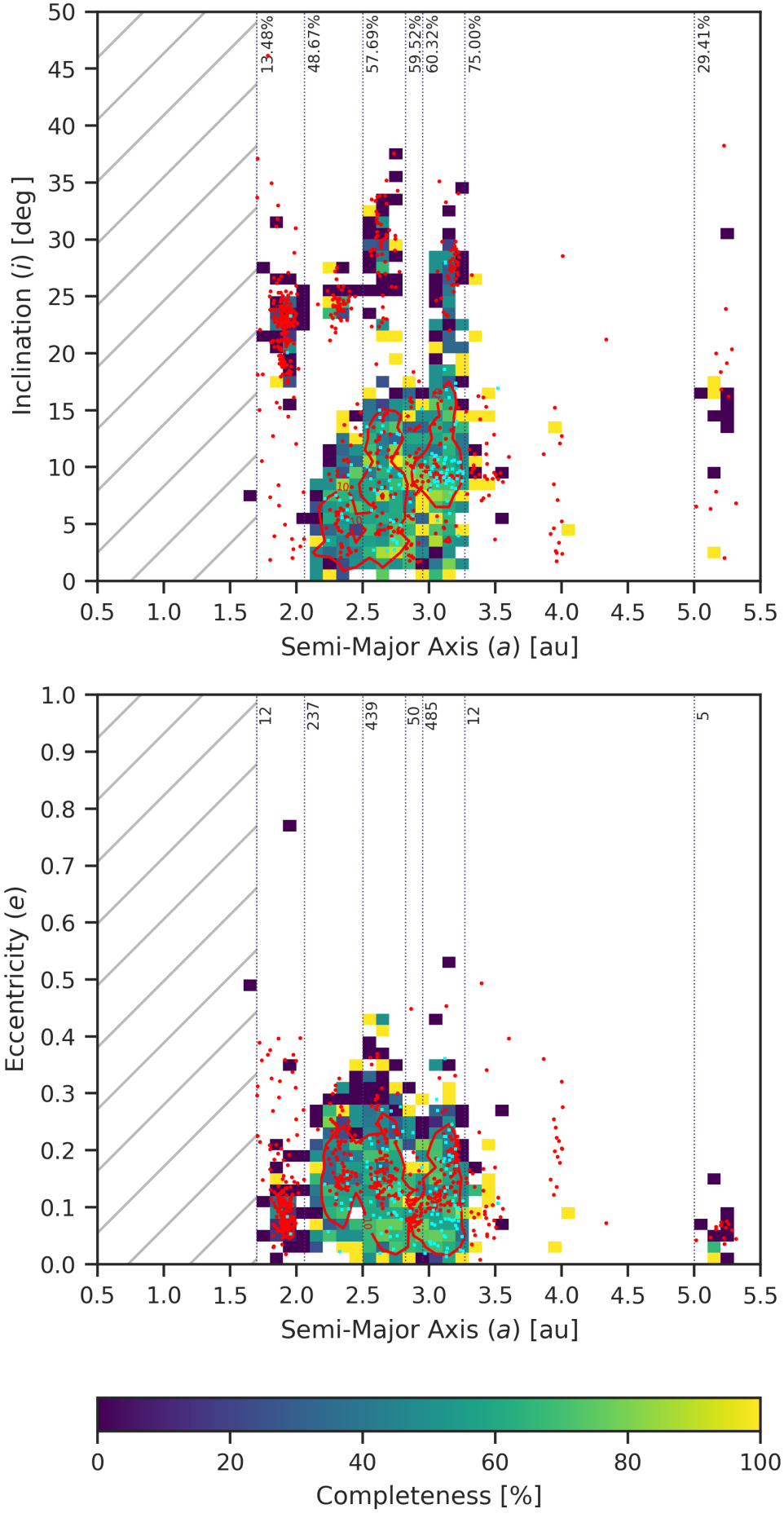
13.48%
48.67%
57.69%
59.52%
60.32%
75.00%
29.41%
Inclination (i) [deg]
Semi-Major Axis (a) [au]
12
237
439
50
485
12
5
Eccentricity (e)
Semi-Major Axis (a) [au]
Completeness [%]

Figure 4.5 In the top panel, the percentage of objects found in pure orbits is visualized in bins of semi-major axis (a) and inclination (i). Number density contours are drawn as red lines to show the number of objects findable (five or more detections throughout the four weeks of the initial NSC search window). The hatched area indicates the region of semi-major axis space where no test orbits were chosen. The 626 test orbits used are plotted as red points. The vertical dashed lines indicate the five semi-major axis bin edges, with the overall percentage completeness per bin written at the top. In the bottom panel, percent completeness in bins of semi-major axis (a) and eccentricity (e); contours and test orbits are plotted in the same style as in the top panel. Instead of percent completeness in each of the five bins, we now explicitly state the number of objects found. This figure was generated using plots_nsc_dr2_initial.ipynb.

the percentages at the top of each bin list the overall completeness in each bin, while in the bottom panel, we list the number of objects found in each bin. Note the stark difference between this plot and Figure 3.9 presented in Chapter 3.4 (Figure 9 in Moeyens et al. 2021). Note the significant drop in completeness in all bins of semi-major axis compared to the same plot as in the case of the ZTF data. The drop in completeness is in part due to the reduction of cell size by a factor of 100 and the factor of two decrease in clustering radius contribute compared to ZTF processing. These reductions were made to allow as much of the ADAM::THOR processing to finish.

In Figure 4.6, we plot the distribution of the 626 test orbits on the sky. We map the total number of observations to a HEALPix grid with *nside*=16 (the same size used when test orbits were selected). In the top panel, we plot in cyan the test orbits that successfully processed and returned orbits (243 of 626, 39%). In the same panel, we plot in red those orbits that successfully finished processing but returned no orbits (344 of 626, 55%). Notice how the population of those orbits is concentrated along the edges of the footprint and along the pixels with the fewest observations. This suggests additional work needs to be done to optimize the test orbit selection algorithm to account for regions of observation space where it is unlikely for objects to be found. Even more strikingly, in the bottom panel, we plot the test orbits that never finished processing within the 4 hr allowed per task (39 of 626, 6%). Note that these orbits are concentrated in the regions of the footprint with the greatest

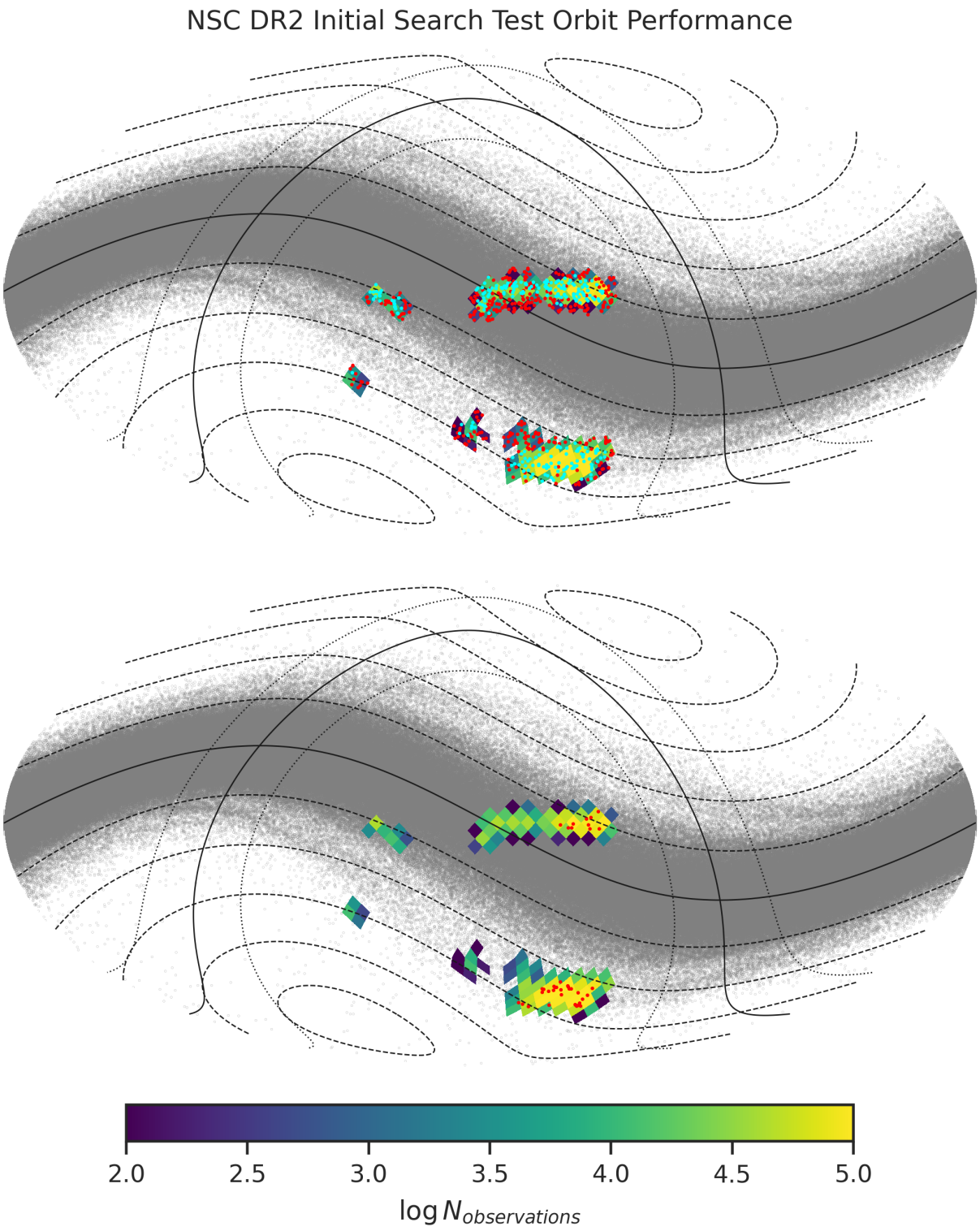
NSC DR2 Initial Search Test Orbit Performance
2.0
2.5
3.0
3.5
4.0
4.5
5.0
log $N_{observations}$

Figure 4.6 In both panels, the number of observations in different regions of the sky are visualized. NSC DR2 observation counts are mapped to a HEALPix scheme with characteristic parameter *nside*=16, the same as used for test orbit selection. The ecliptic plane is plotted as a solid cyan line with corresponding dashed lines of constant ecliptic latitude at ±15, 30, 45, 60, 75 deg. The galactic plane is also plotted as a solid cyan line, with corresponding dotted lines of galactic latitude at ±15 deg. In the top panel, test orbits that completed processing but found no known objects are plotted as red points, while test orbits that completed processing and successfully recovered known objects are plotted as cyan points. Note how the red points are primarily concentrated towards the regions where there are few observations and along the edges of the footprint. In the bottom panel, we plot in red those orbits that never completed processing due to timeout errors. This figure was generated using plots_nsc_dr2_initial.ipynb.

source density. On further inspection, these orbits failed to finish due to two reasons. First, unbeknownst to the authors at the time is that in regions where the source density is high "hotspot_2d" and `DBSCAN` can return more clusters than there are observations. Furthermore, for these tasks "hotspot_2d" can return up to 50% more clusters than `DBSCAN`. In the presence of more linkages, initial orbit determination will need to run for longer to filter the clusters. For a single representative test orbit task, when we increased timeout to 12 hr, we found that given the exact same number of observations (∼300,000) `DBSCAN` could complete clustering in 2.5 hr and generate ∼1 million clusters, while "hotspot_2d" could complete clustering in 11.5 hr and generate ∼1.4 million clusters. In regions with lower source density, "hotspot_2d" outperforms `DBSCAN` in terms of time required to complete clustering being on average at least a factor of ten faster. The second cause of incomplete tasks falls to errors with the Python wrapper of `oorb`. In certain circumstances, when given a particular orbit to propagate or for which to generate ephemerides `pyoorb` will unceremoniously crash. A crash of this type takes down the entire Python process. In a multiprocessed setting, such as is the case for THOR where a single test orbit is processed on a VM with 32 cores, the multiprocessing code will wait for a Python process that no longer exists. The job only finishes when it reaches the 4 hr timeout. On a cloud-based infrastructure, this can be an expensive error as VMs are requisitioned for use but end up sitting idle until timeout.

In an attempt to increase completeness, we processed the same observations but with an increased cell size and with the same clustering radius configuration as was used to process the ZTF observations in Moeyens et al. (2021). Due to the less conservative parameters, the number of orbits that timeout due to crashing increased where only a handful of orbits were completed before each available VM was working on a test orbit that was either stuck in clustering or had crashed due to `pyoorb`. We terminated this job to avoid unnecessarily using Google Cloud Compute credits.

Despite notable software engineering issues, we maintain that these results are still remarkably promising. While only achieving 54.8% completeness, ADAM::THOR which did not successfully process the full sample of test orbits still recovered 1241 known objects. A tracklet-based algorithm requiring three pairs of observations with each pair no more than 90 minutes apart could at best have recovered only 82 objects assuming *ideal* linking performance. This is purely due to the requirement of observing tracklets in a dataset that was not constructed with a tracklet-building cadence. In other words, by enabling tracklet-less linking, THOR was able to recover 15.1x as many objects as a tracklet-based algorithm.

4.5.3 Computational Performance

The total processing time of the ADAM::THOR run was ∼1.5 days using infrastructure as described in Section 4.2. However, as noted in the previous section, this computational time represents an upper bound due to the timeout errors accrued by sub-optimal performance in the presence of many observations and by the occurrence of process-ending crashes. Of 626 test orbits that were submitted for processing 39 tasks timed out.

4.5.4 Completeness and Purity Considerations

In Section 4.5.2, we described ADAM::THOR performance on the known population of asteroids observed during the 30-day window date between 2013 August 30 and 2013 September 30. We reported that the discovery pipeline recovered 1241 known objects of the 2261 with

five or more observations. ADAM::THOR also identified 479 orbits that contain only observations that are unknown or unassociated. It is among these orbits that there are likely undiscovered asteroids.

Given the novel nature of applying THOR to a dataset as deep as the NOIRLab Source Catalog, we conservatively elected to apply one final filter to the 479 discovery candidates to produce a sample of high confidence orbits. We require each orbit to have at least six observations that span a total arc length of at least six days and have $\chi^2_\nu < 5$. This final cut reduces the total number of discovery candidates to 228 orbits.

Each of the 228 orbits was run through the iterative combination of `precovery` and the differential corrector, hereafter the "arc extension" code. This was done with two different sets of parameter combinations based on some improvements made to the code and performance iterations. For the first pass, we ran the arc extension code to search for observations ± 1 yr of the discovery epoch for each orbit. The angular tolerance was set to 5", with a required χ^2_ν threshold of 5. The two years were searched in steps of ± 90 days over 4 iterations. The 228 orbits were reduced to 113 high-quality candidates after this step. Among the 228 orbits were many duplicate objects found by non-overlapping sub-arcs. After the arc extension was completed, we again removed any orbits that shared a subset of the same observations in favor of keeping the occurrence with the lowest overall goodness of fit.

By default, the Minor Planet Center does not accept non-traditional linkages (linkages that contain singleton observations such as the ones THOR can discover). To test that we were added to the appropriate lists and to test our submission format, we selected the 30 objects with the most observations as an initial submission batch. We felt this was a small enough sample that it would not be an undue burden over which to iterate with the MPC, and in the event of any issues, look at each object by eye.

For the second pass of the iterative combination of `precovery` and the differential corrector, we increased the angular tolerance to 10", reduced the χ^2_ν threshold to 3, and increased the search window to ± 1.5 yr from the discovery epoch of each orbit. As before, due to the combination of sub-arcs and several orbits not satisfying the acceptance threshold, the

number of candidates again reduced from 228 orbits to 113. We then removed the 30 orbits that were selected for the first batch, leaving a second batch of 83 high-quality orbits.

The 113 candidates comprise some 1356 observations. To ensure that we did not miss the identification of known objects among the discovery candidates, we ran each observation through SkyBot (Salvatier et al., 2016) to look for any known objects that were near or slightly offset from the observations of each candidate with a search radius of 100". There were five objects which appeared to be consistently offset from already known objects. The mean astrometric offset of these five objects from the recovered observations were 2, 2, 5, 6, and 14". While it seemed plausible that these objects were in fact known, we chose to keep them in the submitted sample to allow the MPC to make the final determination. We added a remark to each observation that was within 100" of a known object in the event that we did not correctly identify known objects among the discovery candidates. Of the 1356 observations of the discovery candidates, 135 were identified as being within 100" of a known object using SkyBot (or about $\sim$10%). We submitted the first batch of orbits on 2022 March 15, and the second on 2022 May 22.

Before the arc extension, the 113 orbits spanned 692 observations. The median number of observations per orbit was 6, with a range of 6 to 8 across all candidates. The median arc length was 10 days, with a range of 6 to 13 days. After arc extension, the median number of observations increased from 6 to 11, with a range of 6 to 24. The median arc length increased from 10 to 25 days, with a range of 8 to 495 days. In Section 4.5.6, we showcase the object with the longest submitted arc as an example of the capabilities of discovery searches using ADAM::THOR. The effect of arc extension on orbital uncertainty is significant. On average the 1-σ positional uncertainty decreased from 0.0030 to 0.0002 au (a 13x improvement), while the 1-σ velocity uncertainty decreased from 0.000025 to 0.000003 au day^{-1} (an 8x improvement).

4.5.5 MPC Submission

In total, we submitted 113 discovery candidates in two batches to the Minor Planet Center for final validation. Of the 113 submitted, 9 were identified as previously known minor planets by the MPC. Of the remaining 104 candidates, 92 were newly designated and were first observed by ADAM::THOR. Twelve were previously observed by other astronomical surveys but had insufficient observations to be identified as new minor planets. In particular, the observations of these 12 asteroids were placed in the MPC's Isolated Tracklet File (ITF), but were elevated out of the file by the observations found by the discovery pipeline. The observations of these 12 asteroids found by THOR in the NSC data would alone have been enough for them be newly designated on their own merit since we submitted fully-fit orbits with at least six observations. As a result of the submission to the MPC, 104 new asteroid orbits were added to the catalog of known asteroids.

Of the five objects we suspected were known, only two were identified as known by the MPC. The two correctly identified objects had astrometric offsets of 2 and 14 arcseconds from the observations found by THOR. The other three objects were newly designated. This means there were a total of seven known objects we failed to correctly identify. This could be due to errors in either the attribution code we used to build the labels for the observations and/or errors in SkyBot. In fact, after we received the results from the MPC, we re-ran SkyBot on all observations with a search radius of one deg. None of the seven objects that were known appeared within one degree of any of the submitted observations. Barring bugs in Skybot or the presence of an incomplete catalog, this likely suggests that the orbits of these objects were uncertain. Without support for the propagation of orbital uncertainty to do attribution, we were unable to appropriately identify these objects.

In Figure 4.5, we plot the orbits of the 104 asteroids in $a - i$ and $a - e$ space as cyan squares. While in Figure 4.7 we visualize the same orbits using OpenSpace[8]. In green are the trajectories of the discoveries during the September 2013 search window. Note that

[8]See https://www.openspaceproject.com/.

the location of the 104 asteroids along their orbits appears to match the footprint of the observations seen in Figure 4.4.

4.5.6 A Case Study: 2013 RR_{165}

We now briefly showcase the discovery of 2013 RR_{165} as it displays a variety of different aspects of the discovery pipeline. In Figure 4.8, we plot the mosaic of postage stamps of asteroid 2013 RR_{165}. The images in the mosaic are centered on the predicted position of the object in each image (represented by the crosshairs). The blue asterisk in the top left corner of the first six images indicates the initial six observations discovered by THOR. Note that the observations are six singletons with an arc length of nine days – a completely tracklet-less discovery. The orbit calculated from these six observations allowed precovery to find an additional 13 observations within 1.5 years of the discovery arc. These are shown by the remaining 13 observations that have green crosshairs.

After submission to the Minor Planet Center, an additional 21 observations of 2013 RR_{165} were extracted from the Isolated Tracklet File (ITF). In Table 4.1 we display the observations of 2013 RR_{165} post-submission and acceptance by the Minor Planet Center. The observations were extracted using the MPC Database Search[9] and represent, at the time of writing, all currently known observations of the object. In the rightmost column, we include an integer counter to delineate the tracklets contained among the 40 observations. All singleton observations represented with an asterisk (*) in addition to tracklets 4-7 were found using ADAM::THOR, while tracklets 1-3 and 8-11 which were observed by Pan-STARRS and Mt. Lemmon Survey were all extracted from the ITF. The observation times of the tracklets extracted from the ITF were too sparsely separated for them to have been identified as, and combined into, a single moving object. More importantly, since each ITF tracklet spanned at most four observations they alone were insufficient for the discovery to be made. Not until the 19 observations found by ADAM::THOR, which alone would have been sufficient

[9]See https://www.minorplanetcenter.net/db_search/.

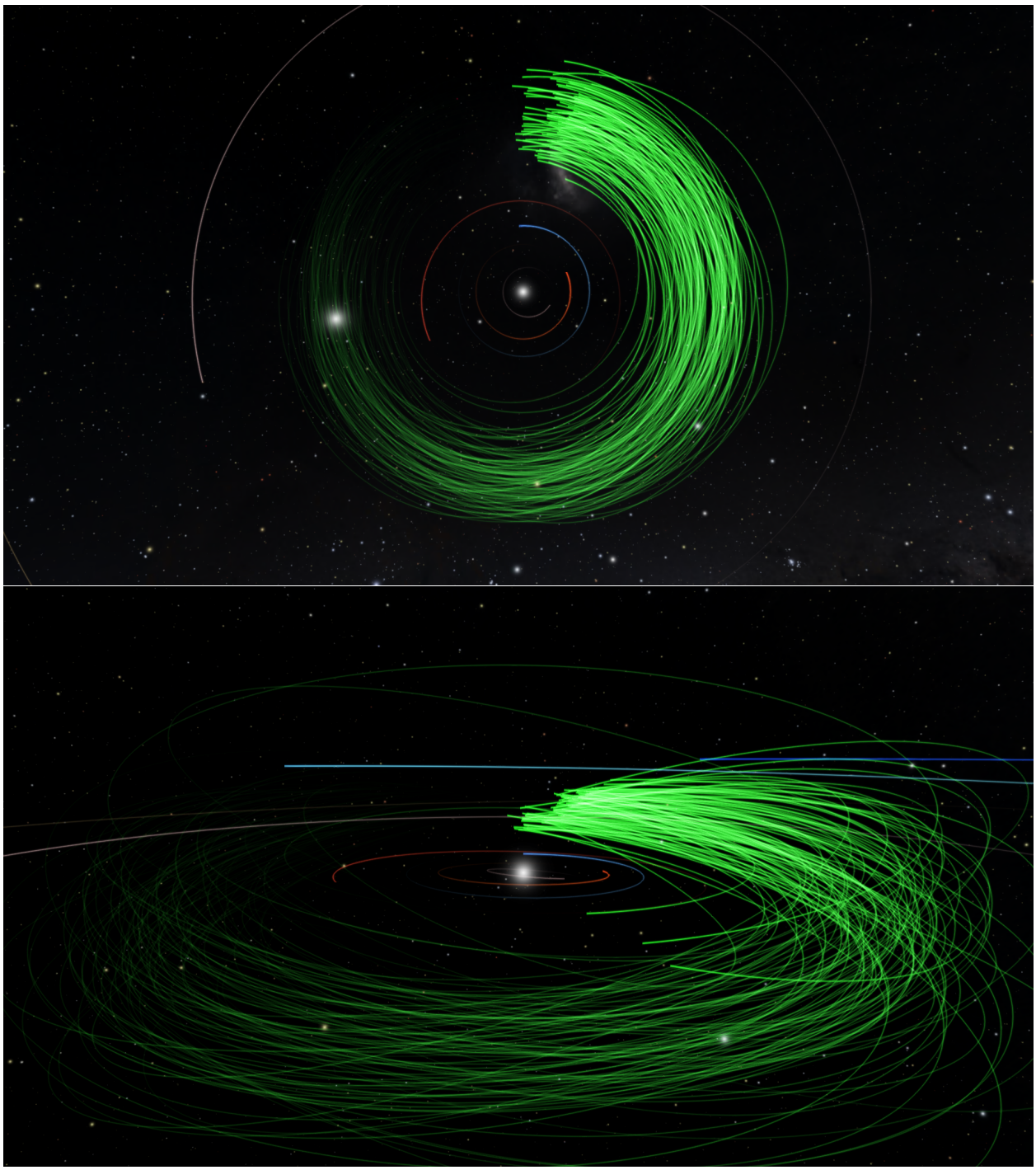

Figure 4.7 The 104 discoveries identified from ∼15% of observations dated between 2013 August 30 and 2013 September 30. This figure was generated using OpenSpace.

2013 RR_{165}

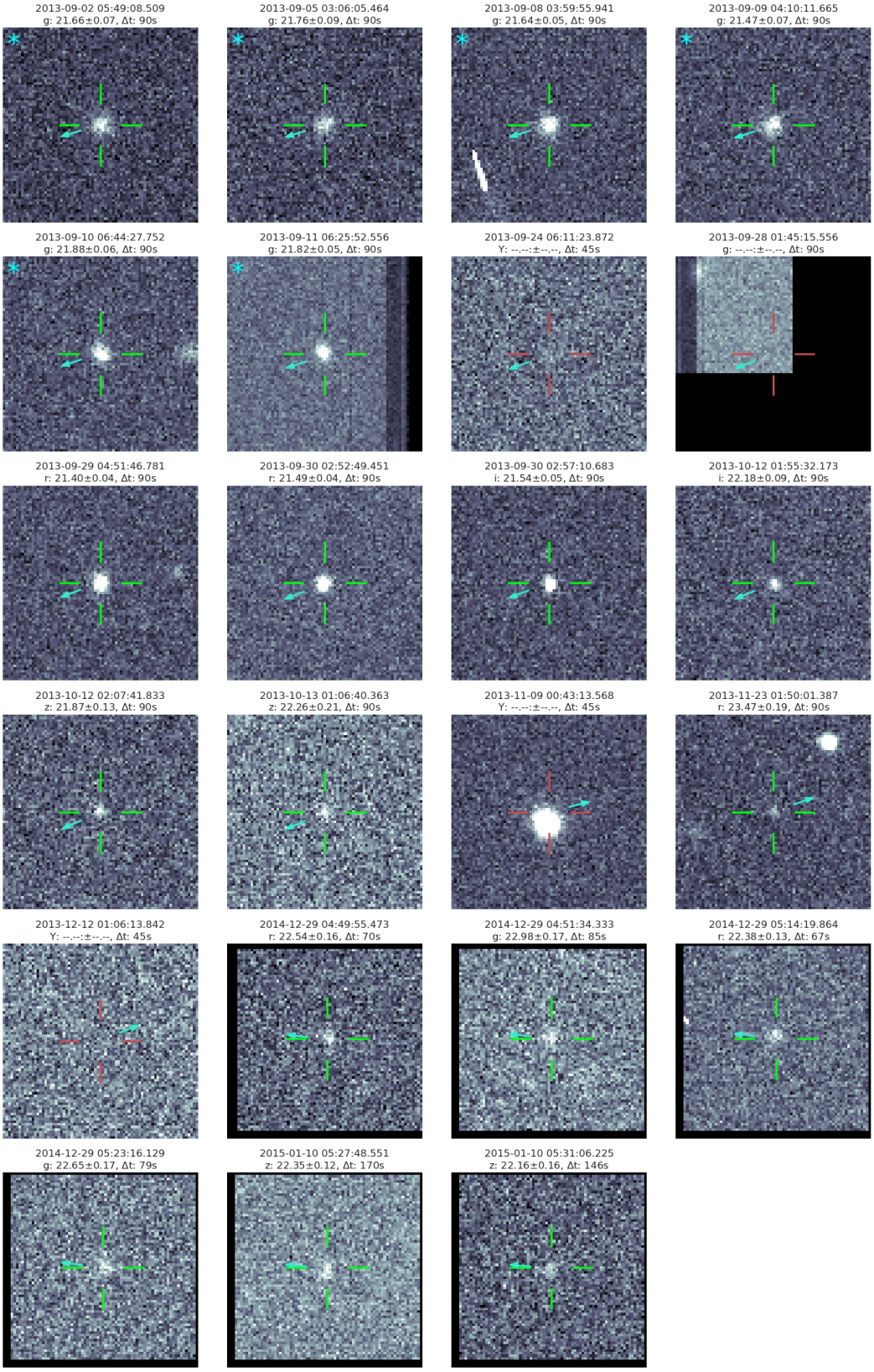
2013-09-02 05:49:08.509
g: 21.66±0.07, Δt: 90s
2013-09-05 03:06:05.464
g: 21.76±0.09, Δt: 90s
2013-09-08 03:59:55.941
g: 21.64±0.05, Δt: 90s
2013-09-09 04:10:11.665
g: 21.47±0.07, Δt: 90s
2013-09-10 06:44:27.752
g: 21.88±0.06, Δt: 90s
2013-09-11 06:25:52.556
g: 21.82±0.05, Δt: 90s
2013-09-24 06:11:23.872
Y: --.--:±--.--, Δt: 45s
2013-09-28 01:45:15.556
g: --.--:±--.--, Δt: 90s
2013-09-29 04:51:46.781
r: 21.40±0.04, Δt: 90s
2013-09-30 02:52:49.451
r: 21.49±0.04, Δt: 90s
2013-09-30 02:57:10.683
i: 21.54±0.05, Δt: 90s
2013-10-12 01:55:32.173
i: 22.18±0.09, Δt: 90s
2013-10-12 02:07:41.833
z: 21.87±0.13, Δt: 90s
2013-10-13 01:06:40.363
z: 22.26±0.21, Δt: 90s
2013-11-09 00:43:13.568
Y: --.--:±--.--, Δt: 45s
2013-11-23 01:50:01.387
r: 23.47±0.19, Δt: 90s
2013-12-12 01:06:13.842
Y: --.--:±--.--, Δt: 45s
2014-12-29 04:49:55.473
r: 22.54±0.16, Δt: 70s
2014-12-29 04:51:34.333
g: 22.98±0.17, Δt: 85s
2014-12-29 05:14:19.864
r: 22.38±0.13, Δt: 67s
2014-12-29 05:23:16.129
g: 22.65±0.17, Δt: 79s
2015-01-10 05:27:48.551
z: 22.35±0.12, Δt: 170s
2015-01-10 05:31:06.225
z: 22.16±0.16, Δt: 146s

Figure 4.8 A mosaic of "postage stamps" or "cutouts" centered on the predicted position of 2013 RR_{165} in each image. The predicted position of the asteroid in each image is indicated by the red or green crosshairs, where the color of the crosshair indicates a point source detection was found within the desired angular tolerance during the precovery search. A green crosshair indicates that a point source was found, whereas a red crosshair indicates a non-detection. The cyan arrow shows the predicted direction of motion in each image. The cyan asterisk in the top left corner of the first six postage stamps marks the six observations first recovered by the THOR algorithm. The 13 observations in the remaining postage stamps were identified using the iterative combination of precovery and the differential corrector. All 19 observations were submitted to, and accepted by, the MPC. After which, the MPC found an additional 21 observations from other discovery surveys in the Isolated Tracklet File. These observations were too sparsely separated to make the discovery but were elevated out of the Isolated Tracklet File by the observations found using ADAM::THOR. 2013 RR_{165}'s initial submission arc of 1.5 years was extended to 6 years with the observations extracted from the ITF. This figure was generated using plots_nsc_dr2_initial.ipynb.

for 2013 RR_{165} to be discovered, were the ITF tracklets correctly established as the same object. Until found by ADAM::THOR, 2013 RR_{165} had remained undetected in observations taken more than eight years prior.

Table 4.1. Observations of 2013 RR_{165}[a]

Date (UT)	RA (J2000)	Dec (J2000)	Magnitude Filter	Location	Ref	Tracklet Identifier[b]
2011 01 30.46905	08 15 32.323	+19 49 12.17	22.2 w	F51 – Pan-STARRS 1, Haleakala	MPS 1612034	1
2011 01 30.48429	08 15 31.384	+19 49 14.17	22.4 w	F51 – Pan-STARRS 1, Haleakala	MPS 1612034	1
2011 01 30.51568	08 15 29.444	+19 49 18.10	22.3 w	F51 – Pan-STARRS 1, Haleakala	MPS 1612034	1
2013 09 02.242460	23 37 59.390	+01 10 25.43	21.66 g	W84 – Cerro Tololo-DECam	MPS 1612034	*
2013 09 03.43078	23 37 02.215	+01 05 56.63	21.6 w	F51 – Pan-STARRS 1, Haleakala	MPS 1612034	2
2013 09 03.44566	23 37 01.446	+01 05 53.11	21.5 w	F51 – Pan-STARRS 1, Haleakala	MPS 1612034	2
2013 09 03.46058	23 37 00.689	+01 05 49.86	21.6 w	F51 – Pan-STARRS 1, Haleakala	MPS 1612034	2
2013 09 05.129230	23 35 38.646	+00 59 28.17	21.76 g	W84 – Cerro Tololo-DECam	MPS 1612034	*
2013 09 08.166620	23 33 04.798	+00 47 00.97	21.64 g	W84 – Cerro Tololo-DECam	MPS 1612034	*
2013 09 09.173746	23 32 12.902	+00 42 42.36	21.47 g	W84 – Cerro Tololo-DECam	MPS 1612034	*
2013 09 10.280877	23 31 15.259	+00 37 52.93	21.88 g	W84 – Cerro Tololo-DECam	MPS 1612034	*
2013 09 11.267969	23 30 23.867	+00 33 30.44	21.82 g	W84 – Cerro Tololo-DECam	MPS 1612034	*
2013 09 12.33407	23 29 28.26	+00 28 36.8	20.9 V	G96 – Mt. Lemmon Survey	MPS 1612034	3
2013 09 12.34315	23 29 27.79	+00 28 34.1	21.1 V	G96 – Mt. Lemmon Survey	MPS 1612034	3
2013 09 12.35223	23 29 27.21	+00 28 31.7	21.1 V	G96 – Mt. Lemmon Survey	MPS 1612034	3
2013 09 12.36127	23 29 26.77	+00 28 29.0	20.5 V	G96 – Mt. Lemmon Survey	MPS 1612034	3
2013 09 29.202625	23 15 28.802	-00 49 40.07	21.40 r	W84 – Cerro Tololo-DECam	MPS 1612034	*
2013 09 30.120017	23 14 48.945	-00 53 42.58	21.49 r	W84 – Cerro Tololo-DECam	MPS 1612034	**4**
2013 09 30.123040	23 14 48.809	-00 53 43.38	21.54 i	W84 – Cerro Tololo-DECam	MPS 1612034	**4**
2013 10 12.080233	23 07 43.253	-01 39 26.43	22.18 i	W84 – Cerro Tololo-DECam	MPS 1612034	**5**
2013 10 12.088679	23 07 43.012	-01 39 28.04	21.87 z	W84 – Cerro Tololo-DECam	MPS 1612034	**5**
2013 10 13.046300	23 07 17.921	-01 42 25.17	22.26 z	W84 – Cerro Tololo-DECam	MPS 1612034	*
2013 11 23.076405	23 12 32.668	-01 39 26.39	23.47 r	W84 – Cerro Tololo-DECam	MPS 1612034	*
2014 12 29.201336	07 37 44.833	+22 51 39.21	22.54 r	W84 – Cerro Tololo-DECam	MPS 1612034	**6**
2014 12 29.202481	07 37 44.765	+22 51 39.16	22.98 g	W84 – Cerro Tololo-DECam	MPS 1612034	**6**
2014 12 29.218285	07 37 43.827	+22 51 40.75	22.38 r	W84 – Cerro Tololo-DECam	MPS 1612034	**6**

4.6 Summary and Future Work

In Section 4.2, we discussed the Asteroid Institute's Asteroid Discovery, Analysis, and Mapping (ADAM) platform, and we described our efforts to build ADAM::THOR, a scalable cloud-based discovery service. In Section 4.3, we followed with a description of the enhancements that have been made to THOR. In Section 4.4, we introduced some of the additional software built to enhance the discovery pipeline: two open-source Python packages named `precovery` and `cutouts`. Finally, in Section 4.5 we described in detail the first results of processing ∼15% of the observations from a window of observations dated between 2013 August 30 and 2013 September 30. In just a 30-day window, THOR recovered 15x the number of objects a tracklet-based algorithm could have recovered. We also identified 113 high-quality discovery candidates that were submitted to the Minor Planet Center for final validation. Of the 113 candidates, 92 were first observed using ADAM::THOR, 12 were elevated out of the Isolated Tracklet File (ITF) by the observations THOR submitted, while nine objects were already known. In total 104 asteroids were added to the MPC's catalog of known objects. Despite discovering at least 104 asteroids in a small slice of the total dataset, ADAM::THOR achieved only 54.8% completeness. The drop in completeness relative to Moeyens et al. (2021) is primarily due to solvable issues of software engineering and a sub-optimal test orbit selection algorithm.

In Figure 4.9, we plot the total number of exposures processed thus far by ADAM::THOR. At the time of writing, only 0.2% of the exposures in the second data release of the NOIRLab Source Catalog have been searched. With 99.8% of the data remaining to be searched, the discovery potential of ADAM::THOR on these data is enormous. Successful and complete processing the NOIRLab Source Catalog is paramount in determining the feasibility of running ADAM::THOR on LSST data.

We now summarize some of the ongoing work to enable at-scale searches of the NSC data, in preparation for searches of LSST data. The early versions of the THOR code grew organically as details of the plausibility of a tracklet-less algorithm were being researched.

Table 4.1 (cont'd)

Date (UT)	RA (J2000)	Dec (J2000)	Magnitude Filter	Location	Ref	Tracklet Identifier[b]
2014 12 29.224492	07 37 43.436	+22 51 41.46	22.65 g	W84 – Cerro Tololo-DECam	MPS 1612034	**6**
2015 01 10.227645	07 25 25.319	+23 08 15.14	22.35 z	W84 – Cerro Tololo-DECam	MPS 1612034	**7**
2015 01 10.229933	07 25 25.179	+23 08 15.26	22.16 z	W84 – Cerro Tololo-DECam	MPS 1612034	**7**
2015 01 16.33421	07 19 01.978	+23 14 49.63	21.7 w	F51 – Pan-STARRS 1, Haleakala	MPS 1612034	8
2015 01 16.35861	07 19 00.421	+23 14 50.93	21.8 w	F51 – Pan-STARRS 1, Haleakala	MPS 1612034	8
2015 01 16.37085	07 18 59.626	+23 14 51.53	21.7 w	F51 – Pan-STARRS 1, Haleakala	MPS 1612034	8
2016 04 03.44182	13 30 42.958	-14 00 30.79	21.9 w	F51 – Pan-STARRS 1, Haleakala	MPS 1612034	9
2016 04 03.46656	13 30 41.612	-14 00 24.82	22.0 w	F51 – Pan-STARRS 1, Haleakala	MPS 1612034	9
2016 04 03.47894	13 30 40.963	-14 00 21.91	22.2 w	F51 – Pan-STARRS 1, Haleakala	MPS 1612034	9
2017 07 05.56689	22 34 06.176	-08 07 08.46	22.6 w	F51 – Pan-STARRS 1, Haleakala	MPS 1612034	10
2017 07 05.57794	22 34 06.219	-08 07 06.86	22.5 w	F51 – Pan-STARRS 1, Haleakala	MPS 1612034	10
2017 07 05.60007	22 34 06.291	-08 07 04.00	22.4 w	F51 – Pan-STARRS 1, Haleakala	MPS 1612034	10
2017 09 13.32757	21 52 17.229	-10 20 15.28	21.4 w	F51 – Pan-STARRS 1, Haleakala	MPS 1612034	11
2017 09 13.34117	21 52 16.676	-10 20 17.25	21.5 w	F51 – Pan-STARRS 1, Haleakala	MPS 1612034	11

[a]Observations from https://www.minorplanetcenter.net/db_search/show_object?utf8=%E2%9C%93object_id=2013+RR165

[b]A running counter to differentiate tracklets among the observations. An asterisk (*) indicates a singleton observation (non-tracklet). A bold identifier marks the observations submitted in this work.

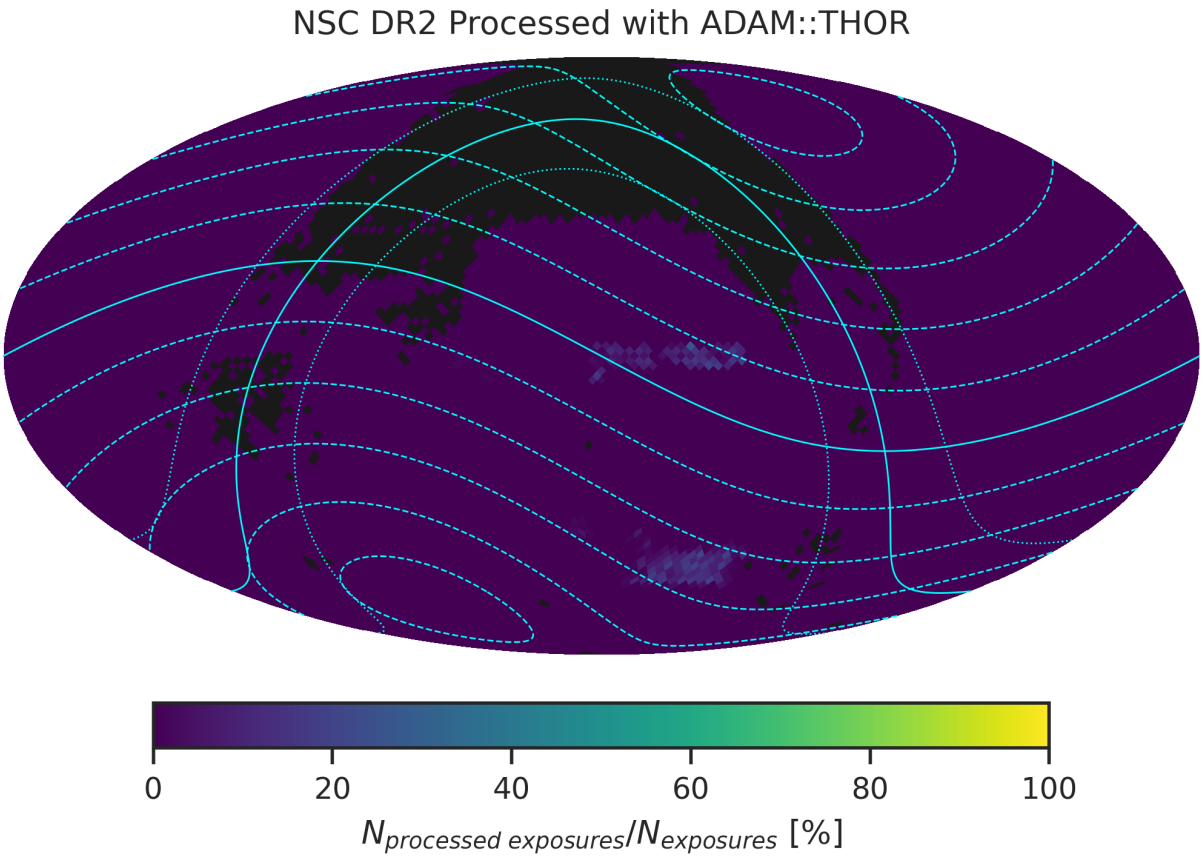

Figure 4.9 The percentage of exposures processed in different regions of the sky with ADAM::THOR are visualized. NSC DR2 exposures are mapped to a HEALPix scheme with characteristic parameter *nside*=32. The ecliptic plane is plotted as a solid cyan line with corresponding dashed lines of constant ecliptic latitude at ±15, 30, 45, 60, 75 deg. The galactic plane is also plotted as a solid cyan line, with corresponding dotted lines of galactic latitude at ±15 deg. This figure was generated using plots_nsc_dr2_initial.ipynb.

All the results mentioned here use THOR versions 1.1 or lower. We have been working on a new version of THOR, overhauled in anticipation of extending the algorithm to the $a < 1.7$ au population and in preparation for at-scale searches of the NSC. In the upcoming version of THOR, which is still in development, we have integrated `JAX` (Bradbury et al., 2018), a Python package developed and supported by Google, that enables composable function transforms. `JAX` combines `Autograd`[10], a package that allows for automatic differentiation of numpy code, with Accelerated Linear Algebra[11] (`XLA`), a linear algebra accelerator that compiles Python code down to C for use on GPUs (Graphics Processing Units) and TPUs (Tensor Processing Units). A Python function calling standard functions in `numpy` can be rewritten as a function of `JAX` primitives. Once written with `JAX` primitives, the function can be easily differentiated. The primary benefit of using `JAX` is that it makes error propagation through the THOR pipeline easier. Additionally, like `numba`, `JAX`-accelerated functions can be just-in-time (JIT) compiled down to C code, increasing computational performance on CPUs, GPUs, and TPUs. In the near future, we will investigate the feasibility of running THOR on GPUs as a method to increase the computational performance, particularly the clustering component.

In Moeyens et al. (2021), the authors described the desire to replace the notional "cell" of observations with the on-sky mapping of an assumed covariance matrix. Thanks to a refactoring of the underlying coordinate and orbits classes in combination with our transition to using the `JAX` code, we now are able to directly map propagated orbital covariances to observation space. The covariance mapping takes into account aberrations such as light travel time and stellar aberration, aberrations which typically are not accounted for when covariances are mapped analytically. Additionally, covariance transformations between different coordinate representations are implemented for Cartesian, Keplerian, cometary, and spherical representations. Observations have been extended to allow for up to six coordinate dimensions to support the information gained by existing tracklets and radar measurements.

[10]See https://github.com/hips/autograd.

[11]See https://www.tensorflow.org/xla.

With the ability to propagate covariance matrices, either via Monte Carlo sampling, sigma-point sampling, or by analytic 2-body methods, we are now able to map volumes of orbital phase space directly to observation space. The primary benefit of this robust orbit-to-observation space mapping is that it gives significantly deeper insight into algorithmic completeness, namely, any orbits that fall within the assumed volume of phase space surrounding a test orbit should be recovered by THOR. With an understanding of not only the mean state for each test orbit but also the volume of phase space probed, the test orbit selection algorithm can be further developed and optimized accordingly to fully cover regions of interest. A final benefit is that the mapping of orbital phase space has also been extended to the gnomonic tangent plane coordinates (the corotating frame of the test orbit) and can be used to enhance the orbit determination techniques which thus far have fallen back to using on-sky coordinates. We believe that completion of this overhaul with further software optimization should bring a factor of 10-100 increase in computational performance, which brings LSST data within computational reach.

The as-yet unanswered question regarding THOR, and by extension, ADAM::THOR is the extension to the $a < 1.7$ au population. The aforementioned upgrades will aid in the development of this work but the algorithmic development remains. In Figure 9 from Jones et al. (2018), reproduced below as Figure 4.10, are shown a variety of different hypothetical linking algorithm discovery criteria in the case of the LSST. With the baseline algorithm, LSST will discover 65.6% of all Potentially Hazardous Asteroids (PHAs) with $H \leq 22$. An algorithm capable of discovering asteroids using six observations over the course of 60 days, could increase that the cumulative discovery rate to 82.5% of PHAs with $H \leq 22$. In other words, a tracklet-less algorithm capable of linking observations over a two-month period would half the residual risk (decrease the remaining undiscovered population from 34.4% to 17.5%). Note that the observing strategy of six observations over 60 days is only a cumulative 1.7% completeness behind the observing strategy of being capable of discovering PHAs using a single observation. The extension of THOR to $a < 1.7$ au population could deliver this improvement in discovery. It is for this reason, that the B612 Foundation, a

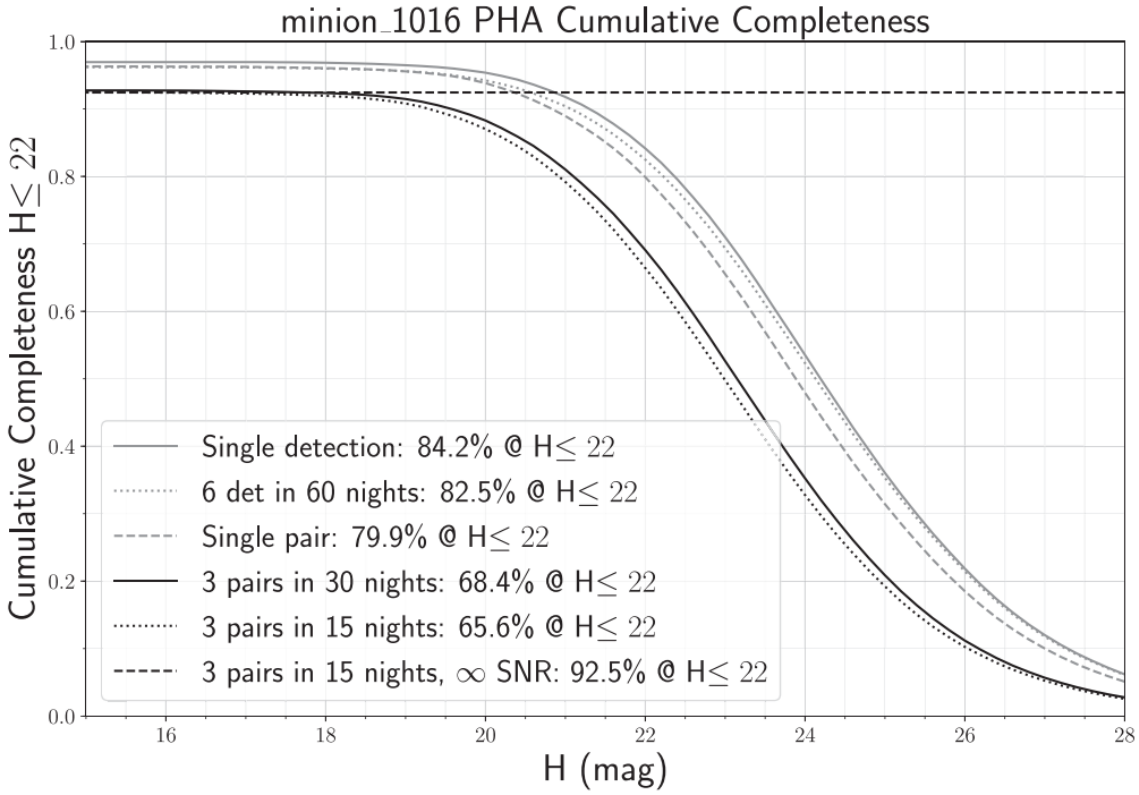

Figure 4.10 Figure 9 from Jones et al. (2018). The PHA cumulative completeness for a variety of different linking algorithmic detection techniques in the context of the Legacy Survey of Space and Time are plotted. Most notably, the tracklet-based approach of 3 pairs in 15 nights is the current plan for the LSST.

non-profit dedicated to planetary defense, has decided to support the development of THOR as the next-generation discovery algorithm and deliver the software, computational, and algorithmic expertise to enable discovery searches in the era of LSST.

The contributions of a NEO-capable THOR are not limited to the Legacy Survey of Space and Time. Recent announcements made by NASA have confirmed that NEO Surveyor (Sonnett et al., 2020; Mainzer et al., 2021), an infrared planetary defense mission, will proceed with a planned launch date in 2028. NEO Surveyor will operate in the L1 Earth-Sun Lagrange point and will focus on the characterization and discovery of Near-Earth Objects. While the cadence of the survey and details of the observing program are still to be determined, preliminary plans suggest NEO Surveyor will look to observe two $\sim$8 hr quads (tracklets made of at least four observations) nominally separated by about 12 days (Mainzer et al., 2021). A THOR-like algorithm capable of discovering NEOs as well – or better than – the current tracklet-based algorithms could allow NEO Surveyor to be extended into a multi-science driver survey, in addition, to making it a more efficient discovery machine.

BIBLIOGRAPHY

Abbott, T. M. C., Abdalla, F. B., Allam, S., et al. 2018, The Astrophysical Journal Supplement Series, 239, 18, doi: 10.3847/1538-4365/AAE9F0

Ballard, D. H. 1981, Pattern Recognition, 13, 111, doi: 10.1016/0031-3203(81)90009-1

Bellm, E. C., Kulkarni, S. R., Graham, M. J., et al. 2019, Publications of the Astronomical Society of the Pacific, 131, 018002, doi: 10.1088/1538-3873/aaecbe

Bernstein, G., & Khushalani, B. 2000, The Astronomical Journal, 120, 3323, doi: 10.1086/316868

Berthier, J., Vachier, F., Thuillot, W., et al. 2006, ASPC, 351, 367. https://ui.adsabs.harvard.edu/abs/2006ASPC..351..367B/abstract

Bowell, E., Hapke, B., Domingue, D., et al. 1989, in Asteroids II, ed. R. P. Binzel, T. Gehrels, & M. S. Matthews, 524–556

Bradbury, J., Frostig, R., Hawkins, P., et al. 2018, JAX: composable transformations of Python+NumPy programs, 0.3.21. http://github.com/google/jax

Chambers, K. C., Magnier, E. A., Metcalfe, N., et al. 2016. https://arxiv.org/abs/1612.05560

Coddington, O., Lean, J. L., Pilewskie, P., Snow, M., & Lindholm, D. 2016, Bulletin of the American Meteorological Society, 97, 1265 , doi: 10.1175/BAMS-D-14-00265.1

Denneau, L., Jedicke, R., Grav, T., et al. 2013, Publications of the Astronomical Society of the Pacific, 125, 357, doi: 10.1086/670337

Dowler, P., Bonnarel, F., & Tody, D. 2015, IVOA Simple Image Access Version 2.0, IVOA Recommendation 23 December 2015, doi: 10.5479/ADS/bib/2015ivoa.spec.1223D

Eggl, S., Juric, M., Moeyens, J., & Jones, L. 2020, in AAS/Division for Planetary Sciences Meeting Abstracts, Vol. 52, AAS/Division for Planetary Sciences Meeting Abstracts, 211.01

Ester, M., Kriegel, H.-P., Sander, J., & Xu, X. 1996, in (AAAI Press), 226–231

Gorski, K. M., Hivon, E., Banday, A. J., et al. 2005, The Astrophysical Journal, 622, 759, doi: 10.1086/427976

Graham, M. J., Kulkarni, S. R., Bellm, E. C., et al. 2019, Publications of the Astronomical Society of the Pacific, 131, 078001, doi: 10.1088/1538-3873/ab006c

Granvik, M., Virtanen, J., Oszkiewicz, D., & Muinonen, K. 2009, Meteoritics and Planetary Science, 44, 1853, doi: 10.1111/j.1945-5100.2009.tb01994.x

Harris, A. W. 1998, Icarus, 131, 291, doi: 10.1006/icar.1997.5865

Harris, A. W., Davies, J. K., & Green, S. F. 1998, Icarus, 135, 441, doi: 10.1006/icar.1998.6002

Harris, A. W., & Drube, L. 2014, Astrophysical Journal Letters, 785, L4, doi: 10.1088/2041-8205/785/1/L4

Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, Nature, 585, 357, doi: 10.1038/s41586-020-2649-2

Holman, M. J., Payne, M. J., Blankley, P., Janssen, R., & Kuindersma, S. 2018, The Astronomical Journal, 156, 135, doi: 10.3847/1538-3881/aad69a

Ivezić, Ž., Connolly, A., Vanderplas, J., & Gray, A. 2014, Statistics, Data Mining and Machine Learning in Astronomy (Princeton University Press)

Ivezić, Ž., Tabachnik, S., Rafikov, R., et al. 2001, The Astronomical Journal, 122, 2749, doi: 10.1086/323452

Ivezić, Ž., Lupton, R. H., Jurić, M., et al. 2002, The Astronomical Journal, 124, 2943, doi: 10.1086/344077

Ivezić, Ž., Kahn, S. M., Tyson, J. A., et al. 2019, The Astrophysical Journal, 873, 111, doi: 10.3847/1538-4357/ab042c

Jones, R. L., Slater, C. T., Moeyens, J., et al. 2018, Icarus, 303, 181, doi: 10.1016/j.icarus.2017.11.033

Jurić, M., Ivezić, Ž., Lupton, R. H., et al. 2002, The Astronomical Journal, 124, 1776, doi: 10.1086/341950

Kubica, J., Denneau, L., Grav, T., et al. 2007, Icarus, 189, 151, doi: 10.1016/j.icarus.2007.01.008

Larson, S., Beshore, E., Hill, R., et al. 2003, in AAS/Division for Planetary Sciences Meeting Abstracts, Vol. 35, AAS/Division for Planetary Sciences Meeting Abstracts #35, 36.04

Lebofsky, L. A., & Rieke, G. H. 1979, Icarus, 40, 297, doi: 10.1016/0019-1035(79)90074-5

Lebofsky, L. A., & Spencer, J. R. 1989, Radiometry and thermal modeling of asteroids., ed. R. P. Binzel, T. Gehrels, & M. S. Matthews, 128–147

Lebofsky, L. A., Sykes, M. V., Tedesco, E. F., et al. 1986, Icarus, 68, 239, doi: 10.1016/0019-1035(86)90021-7

Mainzer, A., Usui, F., & Trilling, D. E. 2015, Space-Based Thermal Infrared Studies of Asteroids, 89–106, doi: 10.2458/azu_uapress_9780816532131-ch005

Mainzer, A., Bauer, J., Grav, T., et al. 2011, Astrophysical Journal, 731, 53, doi: 10.1088/0004-637X/731/1/53

Mainzer, A., Masiero, J., Grav, T., et al. 2012, Astrophysical Journal, 745, 7, doi: 10.1088/0004-637X/745/1/7

Mainzer, A., Abell, P., Bauer, J., et al. 2021, Bulletin of the AAS, 53

Mainzer, A. K., Bauer, J. M., Cutri, R. M., et al. 2016, NASA Planetary Data System, EAR

Masci, F. J., Laher, R. R., Rusholme, B., et al. 2019, Publications of the Astronomical Society of the Pacific, 131, 018003, doi: 10.1088/1538-3873/aae8ac

Masiero, J. R., Grav, T., Mainzer, A. K., et al. 2014, Astrophysical Journal, 791, 121, doi: 10.1088/0004-637X/791/2/121

Masiero, J. R., Mainzer, A. K., & Wright, E. L. 2018a, The Astronomical Journal, 156, 62, doi: 10.3847/1538-3881/aacbd4

Masiero, J. R., Mainzer, A. K., Grav, T., et al. 2011, Astrophysical Journal, 741, 68, doi: 10.1088/0004-637X/741/2/68

Masiero, J. R., Redwing, E., Mainzer, A. K., et al. 2018b, The Astronomical Journal, 156, 60, doi: 10.3847/1538-3881/aacce4

Milani, A., Gronchi, G. F., Farnocchia, D., et al. 2008, Icarus, 195, 474, doi: 10.1016/j.icarus.2007.11.033

Milani, A., Sansaturio, M. E., & Chesley, S. R. 2001, Icarus, 151, 150, doi: 10.1006/icar.2001.6594

Minton, D. A., & Malhotra, R. 2009, Nature, 457, 1109, doi: 10.1038/nature07778

Moeyens, J., & Kiker, K. 2022, B612-Asteroid-Institute/cutouts: v0.1, v0.1, Zenodo, doi: 10.5281/zenodo.7035826

Moeyens, J., Jurić, M., Ford, J., et al. 2021, The Astronomical Journal, 162, 143, doi: 10.3847/1538-3881/ac042b

Mommert, M., Jedicke, R., & Trilling, D. E. 2018, The Astronomical Journal, 155, 74, doi: 10.3847/1538-3881/aaa23b

Myhrvold, N. 2018a, Icarus, 303, 91, doi: 10.1016/j.icarus.2017.12.024

—. 2018b, Icarus, 314, 64, doi: 10.1016/j.icarus.2018.05.004

Nelson, S., Moeyens, J., Kiker, K., Tellis, N., & Stetzler, S. 2022, B612-Asteroid-Institute/precovery: v0.1, v0.1, Zenodo, doi: 10.5281/zenodo.7080419

Nidever, D. L., Dey, A., Fasbender, K., et al. 2021, The Astronomical Journal, 161, 192, doi: 10.3847/1538-3881/abd6e1

Omohundro, S. M. 1989, Five balltree construction algorithms (International Computer Science Institute Berkeley)

Parker, A., Ivezić, Ž., Jurić, M., et al. 2008, Icarus, 198, 138, doi: 10.1016/j.icarus.2008.07.002

Patterson, M. T., Bellm, E. C., Rusholme, B., et al. 2019, Publications of the Astronomical Society of the Pacific, 131, 018001, doi: 10.1088/1538-3873/aae904

Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2011, Journal of Machine Learning Research, 12, 2825

Rozitis, B., & Green, S. F. 2011, Monthly Notices of the Royal Astronomical Society, 415, 2042, doi: 10.1111/j.1365-2966.2011.18718.x

Salvatier, J., Wiecki, T. V., & Fonnesbeck, C. 2016, PeerJ Computer Science, 2016, e55, doi: 10.7717/peerj-cs.55

Sonnett, S., Mainzer, A., Bauer, J., et al. 2020, Bulletin of the AAS, 52. https://baas.aas.org/pub/2020n6i208p02

Tonry, J. L., Denneau, L., Heinze, A. N., et al. 2018, Publications of the Astronomical Society of the Pacific, 130, 064505, doi: 10.1088/1538-3873/aabadf

Vallado, D. A. 2013, Fundamentals of Astrodynamics and Applications, 4th edn. (Microcosm Press)

van den Berg, J. 2013, Calculate Rotation Matrix to align Vector A to Vector B in 3d?, Mathematics Stack Exchange. https://math.stackexchange.com/q/476311

Vereš, P., & Chesley, S. R. 2017a, The Astronomical Journal, 154, 12, doi: 10.3847/1538-3881/aa73d1

—. 2017b, The Astronomical Journal, 154, 13, doi: 10.3847/1538-3881/aa73d0

Warner, B. D., Harris, A. W., Vokrouhlický, D., Nesvorný, D., & Bottke, W. F. 2009, Icarus, 204, 172, doi: 10.1016/j.icarus.2009.06.004

Weisstein, E. W. 2001, "Gnomonic Projection" From MathWorld–A Wolfram Web Resource, Wolfram Research, Inc. http://mathworld.wolfram.com/GnomonicProjection.html

Wright, E., Mainzer, A., Masiero, J., et al. 2018. https://arxiv.org/abs/1811.01454

Wright, E. L. 2007. https://arxiv.org/abs/0703085

Wright, E. L. 2013, in American Astronomical Society Meeting Abstracts, Vol. 221, American Astronomical Society Meeting Abstracts #221, 439.05

Wright, E. L., Eisenhardt, P. R., Mainzer, A. K., et al. 2010, Astronomical Journal, 140, 1868, doi: 10.1088/0004-6256/140/6/1868

Appendix A

A CORRECTION TO THE QUADRATURE FORMULA FOR THE W3 BAND

Wright (2013) has derived simple quadrature formulae that can be used to compute in-band fluxes for the four WISE bands from model flux F_ν^{ast}. It appears that the provided coefficients for the W3 band can be improved (N. Myhrvold, in prep.). The corrected coefficients are available in the ATM package, in method `bandpassLambda` that can be found in file atm/obs/wise.py. We validated the new coefficients by comparing the approximate integral obtained using the quadrature formula to the exactly integrated flux for a $T = 100$ K blackbody. The new coefficients match the exact integral to better than 0.2%, while the original coefficients result in a 35% smaller flux. We have verified that the original coefficients for the other three bands match exact integrals to sub-percent accuracy.

Appendix B

SELECTION OF M TYPE ASTEROIDS USING WISE-BASED BEST-FIT PARAMETERS

The joint analysis of optical and infrared properties discussed in Section 2.4 is focused on objects with "typical" properties. It shows a good correlation between WISE-based best-fit infrared albedo and optical colors measured by SDSS. Such correlations provide support that infrared emission models and best-fitting parameters are robust because the two datasets are essentially independent. Given this independence, we can also improve our understanding of outliers in each dataset. As discussed below, we select a judicious subsample of outliers using only IR parameters and show that their optical color distribution is different than that for the whole sample. This fact further demonstrates that infrared best-fit model parameters are robust – if instead IR outliers were random measurement or modeling failures, their optical color distribution would not differ from that for the whole sample.

Harris & Drube (2014) argued that the best-fit IR albedo (p_{IR} in the WISE context, p_{W1W2} in the ATM case) and beaming parameter (η) can be used to select metallic asteroids (M taxonomic type). Their main argument is that objects with high radar albedo values, indicative of metallic objects, display a very narrow distribution of IR albedo ($p_{W1W2} \sim 0.2$), while a larger fraction of objects with unusually high beaming parameter values are seen in the same albedo range. Therefore, objects with large η and $p_{W1W2} \sim 0.2$ are good candidates for metallic asteroids. Since WISE data are available for orders of magnitude more objects than radar observations, and metallic objects are interesting in many ways (for discussion see Harris & Drube (2014)), it is prudent to critically examine this method.

Due to the degeneracy discussed in Section 2.2.1, ATM fits only for temperature parameter T_1 and not for η. The relationship between T_1 and η is given by Equation 2.5; high η

corresponds to low T_1. Therefore, an implication of analysis from Harris & Drube (2014) is that low T_1 objects with $p_{W1W2} \sim 0.2$ are good candidates for metallic asteroids. We now examine whether IR data discussed here suggest that such outliers exist, and if so, whether they have distinct optical colors.

The right panel in Figure B.1 shows that the SDSS-WISE sample does contain objects at the low end of T_1 range that have $p_{W1W2} \sim 0.2$. By requiring $340 < T_1/K < 370$ and $0.1 < p_{W1W2} < 0.20$, we select 13 objects out of 1574 objects in the SDSS-WISE sample (there are 32 selected objects out of 2479 objects in the high-quality WISE sample), or about 1% of the sample. These candidates for metallic asteroids are listed in Table 4.

As the two-dimensional color scheme[1] in Figure B.1 illustrates, the IR-selected objects have significantly different optical a color distribution than the full sample: the a color mean and standard deviation for selected objects are – 0.02 and 0.027, respectively (see the left panel in Figure B.1). This difference demonstrates that these objects are not random outliers in the p_{W1W2} vs. T_1 diagram and provides support to the hypothesis about metallic asteroids advanced by Harris & Drube (2014). A Kolmogorov-Smirnoff test using the optical a colors of both the metallic asteroid candidates and the remainder of the full SDSS-WISE sample yields only a 1.7% probability that both samples are drawn from the same distribution (1.9% if the optical a color selection is limited to $a < 0.4$). Yet, the optical colors of these metallic candidates are not sufficiently unique for an efficient selection using only optical data – for example, a restrictive selection based on a and $i - z$ colors that selects only eight out of 13 WISE-selected candidates (a selection completeness of ∼50%) still results in only 10% sample selection purity (that is, there are about 10 times as many other objects in the SDSS subsample selected by the cut).

Out of the 13 candidates, only five have measured taxonomies. One of which has been classified as metal (M), while three are possibly metal-like (X, Xe). There is one potential interloper with classification Sq. Given the simple selection criteria using only modeled

[1]For Python code, see http://www.astroml.org/book_figures/chapter1/fig_moving_objects_multicolor.html.

Table B.1. Candidates for M type (metallic) asteroids[a]

Designation	$g-r$	$r-i$	$logD$	σ_{logD}	$logT_1$	σ_{logT_1}	p_{W1W2}	$\sigma_{p_{W1W2}}$	Tax. T.	Tax. B-D.
(497)	0.48	0.15	4.717	0.024	2.542	0.009	0.141	0.015	M	-
(844)	0.50	0.19	4.715	0.029	2.541	0.011	0.146	0.019	-	X
(1349)	0.51	0.16	4.466	0.040	2.539	0.015	0.175	0.033	-	-
(1546)	0.52	0.19	4.468	0.029	2.546	0.011	0.180	0.022	-	-
(1670)	0.62	0.14	4.365	0.016	2.549	0.006	0.170	0.012	-	-
(1730)	0.58	0.13	4.213	0.025	2.538	0.010	0.166	0.018	-	Xe
(1732)	0.63	0.21	4.397	0.028	2.560	0.011	0.155	0.019	-	-
(1860)	0.52	0.20	4.270	0.012	2.554	0.004	0.130	0.009	-	X
(1977)	0.56	0.15	4.284	0.021	2.552	0.008	0.164	0.016	-	Sq
(2294)	0.61	0.15	4.206	0.017	2.557	0.006	0.158	0.014	-	-
(2573)	0.61	0.17	4.325	0.020	2.564	0.007	0.110	0.011	-	-
(2904)	0.49	0.19	4.209	0.017	2.557	0.006	0.188	0.015	-	-
(4813)	0.56	0.24	4.256	0.027	2.560	0.011	0.147	0.018	-	-

[a] $g-r$ and $r-i$ are SDSS colors. Diameter D is in meters, T_1 in Kelvin. Tholen and Bus-DeMeo (SMASSII) taxonomies drawn from the JPL Small-Body Database Browser.

albedo and characteristic temperature parameter T_1, to have four out of five classified objects be metal or possibly metal-like shows promise.

Infrared data are required to efficiently select candidates for metallic asteroids. As the right panel in Figure B.1 shows, T_1 selection is more restrictive than p_W1W2 selection. Therefore, in the context of selecting metallic candidates with a hypothetical two-band survey, having W3 and W4 data would be more useful than W2 and W3 data. On the other hand, for studies requiring p_{W1W2}, the W1 or W2 band would be a more useful addition to the W3 band than the addition of the W4 band.

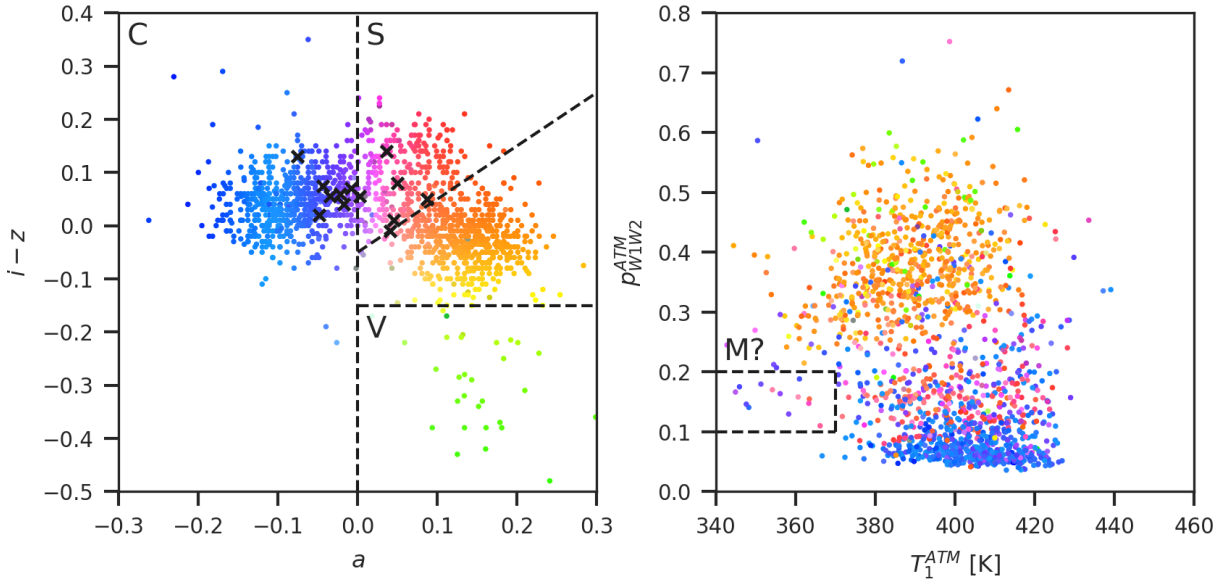

Figure B.1 The colored symbols in the left panel show the a vs. $i-z$ SDSS color-color diagram for the same objects as in Figure 2.15. The symbols' color code is two-dimensional, according to a and $i-z$ colors (for algorithmic details see Ivezić et al. 2002). The same color coding is used in the right panel to visualize the correlation of optical colors and WISE-based best-fit values of IR albedo (p_{W1W2}) and temperature parameter T_1. It is easy to discern that, for example, objects with $i-z<-0.15$ have high IR albedo, while objects with $a<0$ have predominantly low IR albedo. The dashed lines in the right panel outline selection of 13 candidates for metallic asteroids. Their distribution of optical colors is different from the color distribution for the full sample, as further visualized by showing them as crosses in the left panel. The dashed lines in the left panel outline the distribution of main taxonomic classes, as marked in the panel (C, S and V). This figure was generated using analysis_SDSS.ipynb.

Appendix C

NEOWISE PDS 2016 VS. 2019

Since the processing of our results, the NEOWISE team released a newer version of their diameters and albedos for some 160,000 minor planets. In Table C.1, we compare the ratios of diameters from each dataset to best-fit diameters and emissivity in bands W1 and W2 from our "gold" sample. We regenerated the top half of Figure 2.10 with diameters and albedos, or equivalently, emissivities from the 2019 dataset and summarize the results in Table C.1. Comparing ATM diameters to those published in 2019 still yields sub-percent bias (from +0.4% in the 2016 version to −0.4% in the 2019 version) with the same scatter (5.5%). The ratio of emissivities improved with even smaller sub-percent bias (from 0.3% to 0.2%) and a slightly smaller scatter (from 11% to 10%).

Table C.1. NEOWISE Diameters and Albedos 2016 vs. 2019

Dataset	$D^{ATM}/D^{NEOWISE}$ $[\mu \pm \sigma_G]$	$\epsilon^{ATM}_{W1W2}/\epsilon^{NEOWISE}_{W1W2}$ $[\mu \pm \sigma_G]$	Number[b]
NEOWISE Diameters and Albedos v1.0 (2016)	1.004 ± 0.055[a]	1.003 ± 0.11[a]	2656
NEOWISE Diameters and Albedos v2.0 (2019)	0.996 ± 0.055	1.002 ± 0.10	3072

[a]These values are the same as those in the top panels of Figure 2.10.

[b]We used the same selection criteria for both comparisons: NEOWISE fit code "DVBI", $\chi^2 < 3$ (ATM), and at least 28 WISE observations. The number of crossmatched minor planets in the 2019 comparison sample increased relative to its 2016 counterpart in part due to more fits being published and some designations being updated.

Appendix D

TRUE SOLAR SPECTRUM VS. IDEALIZED BLACKBODY

In the ATM modeling framework, the incident solar flux on a model asteroid is described by the Planck function (see Equation 2.9). This assumes that the Sun's spectral energy distribution can be idealized as a blackbody. We now describe the effect of using a more accurate solar spectrum on best-fit parameters for the three asteroids discussed in Section 2.2.4 and for several outliers with unusually strong reflected components from the "gold sample" described in Section 2.3.

We took the three asteroids described in Section 2.2.1.5 and changed their reflected fluxes by the ratio of a more accurate solar spectrum to that of an idealized blackbody spectrum. We used the best-fit diameter, characteristic temperature T_1, and emissivity from the Model 1 case in Table 2.1. We then calculated the NEATM flux with and without the solar component. This allowed us to isolate the thermal and reflected components. We selected the yearly-averaged solar spectrum for 2010 (when the majority of the observations in our "gold" sample were made) from Coddington et al. (2016)[1]. The ratios of this solar spectrum to that of an idealized blackbody integrated over each WISE band are 0.98, 0.88, 0.85, 0.81, in W1, W2, W3, W4, respectively. We then modified the isolated reflected components using these ratios and added them back to the emitted component to create a set of modified fluxes. Using these modified fluxes, MCMC fitting was re-run to examine the effect on the best-fit parameters for the three asteroids in question. We found that the diameters are at worst within 0.2% of previously calculated diameters.

However, the three asteroids selected do not have strong reflected components compared to their thermal emission components. We decided to find a sample of asteroids with un-

[1]See http://lasp.colorado.edu/lisird/data/nrl2_ssi_P1Y/.

usually dominant reflected components, ie, outliers. We selected the fits from our "gold" sample with $\chi^2 < 3$ and then filtered for asteroids where the difference between their median magnitude between W1 and W3 is low number using (W1 - W3) < 3. This is a simple way to find those asteroids whose observations suggest strong reflected fluxes. This yielded only 14 asteroids of which we randomly selected ten and then we repeated the modeling process applied to the three asteroids chosen for the earlier case study. The mean difference in diameter between those in our gold sample and those in this experiment is 1.99% with the highest difference being 3.85%. While this is a not-insignificant difference, the NEATM model itself assumes simple blackbody emission and it seems a more consistent approach to carry the same approximation to incident solar radiation by modeling the Sun as a blackbody. Another key aspect is that the approximation of the Sun as a blackbody follows NEOWISE's modeling assumptions and allows for a better comparison our results.

Appendix E

TRANSFORMATIONS, ROTATIONS AND PROJECTIONS

Here, we summarize the key equations and relations used to carry out the heliocentric transformation and projection into the corotating frame of the test orbit.

E.1 Equatorial and Ecliptic Coordinates

For convenience, we reproduce a subset of the equations presented in Bernstein & Khushalani (2000) that outline the coordinate transformations between the equatorial and ecliptic systems. A set of observations in equatorial coordinates (α, δ) can be transformed into Cartesian equatorial coordinates as follows

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix}_{eq} = \begin{bmatrix} \cos\delta\cos\alpha \\ \cos\delta\sin\alpha \\ \sin\delta \end{bmatrix}. \tag{E.1}$$

Similarly, a set of ecliptic coordinates (λ, β) can be transformed into Cartesian ecliptic coordinates:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix}_{ec} = \begin{bmatrix} \cos\beta\cos\lambda \\ \cos\beta\sin\lambda \\ \sin\beta \end{bmatrix}. \tag{E.2}$$

Cartesian equatorial coordinates can be transformed to ecliptic Cartesian coordinates. Here, ϵ is the obliquity of the ecliptic ($\epsilon = 84381.448''$):

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix}_{ec} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\epsilon & \sin\epsilon \\ 0 & -\sin\epsilon & \cos\epsilon \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}_{eq} . \tag{E.3}$$

Likewise, ecliptic Cartesian coordinates can be converted to equatorial Cartesian coordinates

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix}_{eq} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\epsilon & -\sin\epsilon \\ 0 & \sin\epsilon & \cos\epsilon \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}_{ec} . \tag{E.4}$$

Using equations E.1 - E.4, coordinates in either the angular or Cartesian form for both the ecliptic and equatorial systems can be trivially transformed between the respective systems and the different forms. To proceed with the heliocentric transformation which is described in the next section, THOR transforms all equatorial coordinates to ecliptic coordinates.

E.2 Heliocentric Transformation

Before proceeding, it is useful to define certain quantities in relation to the assumed observing geometry. The location of the observer in the heliocentric frame we denote as $\vec{r}_o$ and is assumed to be known. The distance from the Sun to the observer is then simply the magnitude of said vector ($|\vec{r}_o| = r_o$). The unit vector n_{ao} lies in the direction of an observed asteroid from the position of the observer. Unless the asteroid's orbit is well-constrained and known beforehand, we do not know the distance from the observer to the asteroid. We express the observer to asteroid position vector as $\vec{r}_{ao}$ which has magnitude Δ and is equivalent to $n_{ao} \cdot \Delta$. Last, we denote the unknown heliocentric position vector of the asteroid as $\vec{r}_a$ which has magnitude r.

A fiducial test orbit has heliocentric position vector $\vec{r}$ and velocity $\vec{v}$ given at some epoch, t. The location in space and time of an object on the test orbit can be precisely predicted. Assuming there are other minor planets on orbits similar to the test orbit we can propagate

the test orbit to all times where it could have been detected in a survey and gather actual detections near its predicted location (Sections 3.2.1.1 - 3.2.1.2). Nearby detections likely belong to moving objects on similar orbits. If we assume those lie at the same heliocentric distance from the Sun, we have enough information to fully constrain the observing geometry for each individual detection allowing the unknown quantities (Δ and $\vec{r}_a$) to be calculated.

From our assumed observing geometry, we know the direction $\hat{n}_{ao}$ for every gathered detection. Assuming each detection lies at the distance r of the test orbit, we can then calculate Δ:

$$\Delta = -\vec{n}_{ao} \cdot \vec{r}_o + \sqrt{(\vec{n}_{ao} \cdot \vec{r}_o) + r^2 - r_o^2}. \tag{E.5}$$

Once we know Δ, we can solve for $\vec{r}_a$, where

$$\vec{r}_a = \vec{r}_{ao} + \vec{r}_o. \tag{E.6}$$

At this point, we can attribute a heliocentric position vector to every gathered detection at its specific epoch, again under the assumption that the detection belongs to an orbit similar to the test orbit.

E.3 Transformation into Frame of the Test Orbit

Once the heliocentric coordinates of each observation have been calculated, the observations of other minor planets can be transformed into the frame of the test orbit. This is accomplished using a double rotation. First, the detections are rotated so that the test orbit lies in the x-y plane. Second, the detections are rotated so the test orbit lies along the x-axis.

To perform the first rotation which we denote as $\vec{R}_1$, we transform the detections so that $\vec{r}$ lies in the x-y plane. This is equivalent to rotating the vector normal to the plane of the orbit ($\hat{n}$) towards the z-axis as defined in heliocentric space ($\hat{z}$). We now follow the method outlined in van den Berg (2013). The rotation axis, $\hat{v}$, can be found by taking the cross

product of $\hat{n}$ and $\hat{z}$,

$$\hat{v} = \hat{n} \times \hat{z}. \tag{E.7}$$

The rotation matrix, $\vec{R}_1$, is then

$$\vec{R}_1 = I + [\hat{v}]_\times + [\hat{v}]_\times^2 \frac{1}{1+c}, \tag{E.8}$$

where $c = \hat{n} \cdot \hat{z}$ and $[\hat{v}]_\times$ is the skew-symmetric cross product matrix of $\hat{v}$,

$$[\hat{v}]_\times \equiv \begin{bmatrix} 0 & -\hat{v}_3 & \hat{v}_2 \\ \hat{v}_3 & 0 & -\hat{v}_1 \\ -\hat{v}_2 & \hat{v}_1 & 0 \end{bmatrix}. \tag{E.9}$$

Finally, we need to rotate the detections so that $\vec{r}$ lies along the x-axis. We denote this rotation as $\vec{R}_2$. Let $\hat{r}'$ be a unit vector in the direction of a detection or test orbit rotated into the x-y plane by $\vec{R}_1$. The angle of rotation towards the x-axis, α, is equivalent to $-\hat{x} \cdot \hat{r}'$,

$$\vec{R}_2 = \begin{bmatrix} \cos\alpha & \sin\alpha & 0 \\ -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{E.10}$$

The complete rotation matrix is $\vec{M} = \vec{R}_2 \cdot \vec{R}_1$.

E.4 Gnomonic Projection

After the observations have been rotated, they can be projected into a gnomonic tangent plane (Weisstein, 2001). The gnomonic tangent plane preserves great circle distances and so offers a natural projection in which to discover minor planets relative to a test orbit. The gnomonic projection is also used by both Holman et al. (2018) and Bernstein & Khushalani (2000). The gnomonic projection can be accomplished using either spherical or Cartesian coordinates. For completeness, we list both sets of equations. In our implementation, we directly project the transformed and rotated Cartesian coordinates onto the gnomonic tangent plane.

To project a set of spherical coordinates (λ, ϕ) onto a gnomonic plane (θ_X, θ_Y) tangent about point (λ_0, ϕ_0) (in our case, the location of the test orbit at each epoch in which detections were found and gathered):

$$\theta_X = \frac{\cos\phi \sin(\lambda - \lambda_0)}{\sin\phi_0 \sin\phi + \cos\phi_0 \cos\phi \cos(\lambda - \lambda_0)}, \tag{E.11}$$

$$\theta_Y = \frac{\cos\phi_0 \sin\phi - \sin\phi_0 \cos\phi \cos(\lambda - \lambda_0)}{\sin\phi_0 \sin\phi + \cos\phi_0 \cos\phi \cos(\lambda - \lambda_0)}. \tag{E.12}$$

The same projection can be achieved using the rotated Cartesian coordinates described in the previous section. In this case, by convention, the test orbit lies at $(\lambda_0, \phi_0) = (0, 0)$, making the projection simply

$$\theta_X = \frac{x''}{y''}, \tag{E.13}$$

$$\theta_Y = \frac{z''}{x''}. \tag{E.14}$$

The double prime ($''$) is used to denote the transformed Cartesian coordinates after they have been rotated as described in Section E.3. After completing the procedure outlined in Sections E.1 - E.4, we now have a set of detections and their locations in the corotating reference frame of a test orbit. From this reference frame, other orbits can be recovered as described in Section 3.2.1.4.