Publications of the Astronomical Society of the Pacific, 133:034503 (17pp), 2021 March https://doi.org/10.1088/1538-3873/abd4bc

The TAOS II Survey: Real-time Detection and Characterization of Occultation Events

Chung-Kai Huang[1,2], Matthew J. Lehner[1,3,4], Agueda Paula Granados Contreras[1], Joel H. Castro-Chacón[5], Wen-Ping Chen[2], Charles Alcock[4], Fernando I. Alvarez-Santana[6], Kem H. Cook[1], John C. Geary[4], Carlos Alberto Guerrero Peña[6], Joannes Bosco Hernández-Águila[6], Benjamín Hernández-Valencia[6], Jennifer Karr[1], J. J. Kavelaars[7,8], Timothy Norton[4], Juan Manuel Nuñez[6], Diego Ochoa[6], Mauricio Reyes-Ruiz[6], Edilberto Sánchez[6], José Sergio Silva[5], Andrew Szentgyorgyi[4], Shiang-Yu Wang[1], Wei-Ling Yen[1], and Zhi-Wei Zhang[1]

[1] Institute of Astronomy and Astrophysics, Academia Sinica, No.1, Sec. 4, Roosevelt Rd, Taipei 10617, Taiwan; kevinhuang@asiaa.sinica.edu.tw

[2] Institute of Astronomy, National Central University, 300 Zhongda Road, Zhongli 32001, Taiwan

[3] Department of Physics and Astronomy, University of Pennsylvania, 209 South 33rd Street, Philadelphia, PA 19125, USA

[4] Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA

[5] CONACYT—Instituto de Astronomía, Universidad Nacional Autónoma de México, Ensenada, B.C., México

[6] Instituto de Astronomía, Universidad Nacional Autónoma de México, Ensenada, B.C., México

[7] Herzberg Astronomy and Astrophysics Research Centre, National Research Council of Canada, 5071 West Saanich Road, Victoria, British Columbia V9E 2E7, Canada

[8] Department of Physics and Astronomy, University of Victoria, Building, 3800 Finnerty Road, Victoria, BC V8P 5C2, Canada

Received 2020 October 19; accepted 2020 December 17; published 2021 February 17

Abstract

The Transneptunian Automated Occultation Survey (TAOS II) is a blind occultation survey with the aim of measuring the size distribution of Trans-Neptunian Objects with diameters in the range of $0.3 \lesssim D \lesssim 30$ km. TAOS II will observe as many as 10,000 stars at a cadence of 20 Hz with all three telescopes simultaneously. This will produce up to ∼20 billion photometric measurements per night, and as many as ∼6 trillion measurements per year, corresponding to over 70 million individual light curves. A very fast analysis pipeline for event detection and characterization is needed to handle this massive data set. The pipeline should be capable of real-time detection of events (within 24 hours of observations) for follow-up observations of any occultations by larger TNOs. In addition, the pipeline should be fast and scalable for large simulations where simulated events are added to the observed light curves to measure detection efficiency and biases in event characterization. Finally, the pipeline should provide estimates of the size of and distance to any occulting objects, including those with non-spherical shapes. This paper describes a new data analysis pipeline for the detection and characterization of occultation events.

Key words: Trans-Neptunian objects – Stellar occultation – Astronomy data modeling – Time domain astronomy – Light curves

1. Introduction

The size distribution of Trans-Neptunian Objects (TNOs) has been accurately measured down to diameters of $D \gtrsim 25$ km (Luu & Jewitt 2002; Bernstein et al. 2004; Fuentes & Holman 2008; Fraser et al. 2008; Fraser & Kavelaars 2008, 2009; Fuentes et al. 2009). A measurement of the size distribution of even smaller objects (down to $D \gtrsim 0.3$ km) is needed because it would help constrain models of both the dynamical evolution of the solar system (Duncan et al. 1995; Stern & Campins 1996; Davis & Farinella 1997; Benz & Asphaug 1999; Kenyon & Luu 1999b, 1999a; Kenyon & Bromley 2001, 2004; Pan & Sari 2005; Benavidez & Campo Bagatin 2009; Kenyon & Bromley 2009) and the origin of short-period comets (Holman & Wisdom 1993; Duncan & Levison 1997; Levison & Duncan 1997; Morbidelli 1997; Tancredi et al. 2006; Volk & Malhotra 2008). The detection of such objects is difficult because they are extremely faint, with typical magnitudes $r' > 28$, and are thus invisible to surveys using even the largest telescopes. However, a small TNO will induce a detectable drop in the measured brightness of a distant star when it passes across the line of sight (Bailey 1976; Roques et al. 1987; Brown & Webster 1997; Roques & Moncuquet 2000; Cooray 2003; Cooray & Farmer 2003; Roques et al. 2003, 2006; Chang et al. 2006, 2007; Nihei et al. 2007; Bickerton et al. 2008, 2009; Liu et al. 2008; Zhang et al. 2008, 2013; Bianco et al. 2009; Schlichting et al. 2009, 2012; Wang et al. 2009, 2010; Bianco et al. 2010; Arimatsu et al. 2019a, 2019b).The goal of the TAOS II project (Lehner et al. 2014) is to detect such occultation events and measure the size distribution of TNOs with diameters $0.3 \text{ km} < D < 30$ km.

TAOS II will operate at the Observatorio Astronómico Nacional at San Pedro Mártir (SPM) in Baja California, México. The survey will operate three 1.3 m F/4 telescopes, with separations ranging from 130 m to 323 m. Each telescope will be equipped with a custom CMOS camera, which will be capable of simultaneous 20 Hz imaging on more than 10,000 stars by reading out small subframes (around 9×9 pixels, depending on the brightness of the star and the seeing) around each of the target stars. Multiple telescopes are used in order to reduce the false positive rate. The three telescopes will image the same stars simultaneously, and any candidate events will need to be detected coincidentally in all three telescopes. The wide separations of the telescopes ensure that any false positive event signatures due to scintillation in the upper atmosphere will not affect all three telescopes simultaneously.

The TAOS II telescopes are installed and functional, and the site development is complete. The cameras are currently still being built, and should be completed in early 2021. SPM is currently closed due to COVID-19, but once it reopens and the cameras are delivered, the survey can begin after a short period of commissioning is completed.

TAOS II will typically collect 20 Hz observations on a field for a nominal time of 2 hr. Each of the 10,000 stars in the field will thus have 144,000 photometric measurements on three telescopes in this time span. Assuming an average observing time of 10 hr per night (5 fields observed for 2 hours), and 300 nights per year of observations, TAOS II will collect a total of 15 million light curves in a single year, or 75 million light curves over the nominal 5 yr planned for the survey. This corresponds to 6.5 trillion photometric measurements per year, or 32.5 trillion measurements over 5 yr. Assuming an average window size of 9×9 two byte pixels, the size of the raw photometric data set will be about one petabyte per year.

A fast, efficient, and scalable analysis pipeline is required to handle this massive data set. First, the pipeline should be capable of finding and characterizing any candidate occultation events in one night of observations before the next night of observations begins. TAOS II will possibly detect a few TNOs that are large enough to be observed with either space-based or large ground-based telescopes, and if any of these objects are successfully imaged directly it would both confirm the occultation events and present the possibility to learn more valuable information about the occulting objects. Given the very limited orbital information (i.e., a rough estimate of distance) a blind occultation survey can provide, the error window on the location of the TNO will increase rapidly after the event. Nearly immediate follow-up of any such detections will thus be required in order maximize the probability of successful recovery of the object. Second, the detection efficiency of the survey will be estimated by analyzing the recovery rate of simulated events added to the raw light curve data collected each night. This will be repeated many times for each light curve, and while this will not be expected to finish overnight, the pipeline should be fast enough and scalable to enough CPUs to finish the task in a reasonable amount of time (i.e., months, rather than years).

This paper describes an analysis pipeline that has been developed to meet these requirements. This pipeline consists of two stages: the event trigger and event parameter estimation. The data rate is extremely high and the event rate is very low, so the detection algorithm needs the capability of finding occultation events very efficiently while minimizing the false positive rate. The event trigger for this pipeline is discussed in Section 2. For candidate events that are found by the trigger algorithm, a quick estimate of diameter of and distance to the occulting TNO is needed to both help confirm the detection and facilitate any follow-up campaign. These estimates will also be used in the efficiency analysis, and will be used to investigate any biases in the pipeline. The event shape parameter estimation algorithm used in this step is discussed in Section 3 and its application to simulated occultation light curves is presented in Section 4.

We do not discuss photometric analysis in any detail in this paper. While this process is under investigation, the default plan is to use simple aperture photometry. Given the bright target stars and high cadence observations, dark current will be negligible, and sky background will be estimated from a number of subframes with no stars placed strategically around each imager. Moments of the point-spread function for each star will be monitored to ensure no light from the target star falls outside of any individual imager subframe, and the stellar intensity will be estimated by summing up the photons counted in each of these subframes.

2. Event Trigger

The light curves collected by the survey need to be searched for significant deviations over small timescales that are coincident in time in the data from all three telescopes. The trigger for event detection must be sensitive to different event durations (5 to 60 data points at opposition, corresponding to 250 ms to 3 s), and it should be insensitive to short period trends. Given the vast quantity of data that will be collected by the survey and the extremely low expected event rate, the event detection algorithm must also be fast and efficient. We have developed an event detection pipeline that meets these criteria, and this pipeline is presented in the following subsections.

For the remainder of this paper, we define the variable I as the measured intensity of a target star. This intensity is measured periodically at a high cadence (nominally 20 Hz, but the possibility exists to change this if necessary) to generate a light curve. The measurements are taken at times t_n, where n is the index of the measurement in the time series. For a given light curve, $1 \leqslant n \leqslant N_{\mathrm{LC}}$, where N_{LC} is the number of measurements in the light curve. For a two hour time series of exposures collected at a cadence of 20 Hz, $N_{\mathrm{LC}} = 144{,}000$.

In order to simplify our notation, we define

$$I_n = I(t_n). \quad (1)$$

2.1. Event Window

To detect a candidate event, we search over a range of timescales for significant variations relative to an otherwise flat light curve. We define an *event window* as a series of points in a light curve over which we look for this variation. On average, the event detection is optimized when the width of this window is nearly equal to the duration of the event and if the window is centered on the event. We thus search for events in different windows with a series of widths corresponding to the range of expected occultation event durations. Furthermore, we will search for variation with the window centered on every point n in a given light curve, with the exception of some points on either end of the light curve (this will be discussed in Section 2.2).

In order to cover the range of sensitivity of the survey, we use event window sizes of $s_{\mathrm{evt}} = 3$, 5, 7, 9, 11, 15, 19, 23, 27, 31, 35, 41, 47, 53, 59, 67, 75, 83, 91, 99, 109, 119, 129, 139, 149, 161, 173, 185, and 197. This large number of windows is due to the fact that it is desirable to have the window size closely match the event duration, since the optimum window size is used to bootstrap the parameter estimation for any detected events. We note that this list might change slightly once we test the pipeline on actual survey data, where the trade-offs between extra processing time versus event detection efficiency will be considered.

2.2. Light Curve Filtering

Experience from TAOS I has shown (Lehner et al. 2010) that light curves collected at high cadence will often exhibit short term variations in measured intensity. These variations are seen in both the mean and variance of the light curves. As long as these variations are on somewhat larger timescales than the event window, they can be filtered out. If not, then the data are likely too noisy to search for events on the desired timescale and should not be used.

To remove these short term variations in the light curves, we follow a method similar to that described in Lehner et al. (2010) and define the quantities μ_n and σ_n, the mean and standard deviation of the light curve in the region of the measurement at time t_n. The values of μ_n and σ_n must be calculated directly from the data using a set of points in the region around point n. We thus define a pair of *filter windows* on either side of the event window at each point in the light curve.

The points in the event window should be excluded from the filter windows since, in the event of an occultation inside of the event window, including these points will reduce the value of μ_n and thus lessen the measured variability. Furthermore, a set of points immediately adjacent on either side of the event window should also be excluded. A large fraction of the occultation events we expect to detect will have gradual drops at the start of the events and correspondingly slow rises at the ends of the events, so in many cases the optimum window size will be somewhat smaller than the actual event duration. We thus put a buffer between the event window and each filter window with the same size s_{evt} of the event window itself. The size of the filter windows should be large enough to obtain accurate estimates of μ_n and σ_n, but small enough that any short term variability in the light curve is not averaged out. The selection of the value is somewhat arbitrary, and a default value of $s_{\mathrm{fltr}} = 30$ was chosen as this works well in our testing. We may adjust this value after testing with actual data from the survey.

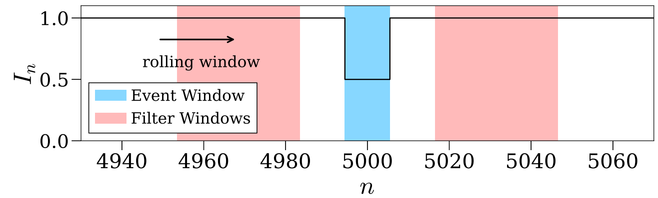

Figure 1. Illustration of event and filter windows. The blue region is the event window and the two red regions, one on each side of the event window, are the filter windows. We place a buffer with the same size as the event window between the event window and each filter window. The black line indicates a light curve with an inverse top hat event with a width of 11 points and a depth of 0.5. The set of windows "rolls" along the light curve from left to right and the values of μ_n and σ_n are calculated using the points in the filter windows at each step. The event widow is centered on the event in this figure, which is the point of maximum variability in the event window.

The rolling light curve filter and event windows are illustrated in Figure 1. For a given event window size s_{evt} and a given point n in the light curve, we center the event window at point n, and define the two filter windows as containing the points

$$\begin{aligned} n - s' - s_{\mathrm{fltr}} &\leqslant m \leqslant n - s', \\ n + s' &\leqslant m \leqslant n + s' + s_{\mathrm{fltr}}, \end{aligned} \quad (2)$$

where $s' = \frac{1}{2}(3s_{\mathrm{evt}} + 1)$.

At each point n, the mean μ_n and standard deviation σ_n are calculated using the points m in the filter windows given in Equation (2). 3σ-clipping is applied in these calculations in order to keep a single large random deviation from affecting the estimates of μ_n and σ_n over a large range of points. Note that we do not search for events (that is, we do not center the event window on points) in the ranges

$$\begin{aligned} &n < s' + s_{\mathrm{fltr}} \\ &\text{and} \\ &n > N_{\mathrm{LC}} - s' - s_{\mathrm{fltr}} \end{aligned} \quad (3)$$

since we need the points at either end of the light curve for the filter windows and buffers.

2.3. Three-telescope Event Detection

The metric we use to define variability is based on the p-value resulting from the χ^2 of the light curve over the window using the null hypothesis of no event (flat light curve), that is

$$\chi_n^2 = \sum_m \left(\frac{I_m - \mu_n}{\sigma_n}\right)^2, \tag{4}$$

where I_m is the measured light curve intensity measured at point m in the light curve, μ_n and σ_n are calculated as described in the previous subsection, and the sum is over all points m in the event window centered on point n. The variability metric used to trigger our event detection algorithm is

$$p_{\chi^2}(\chi_n^2;\, s_{\rm evt}), \tag{5}$$

where p_{χ^2} is the p-value calculated with a value of χ_n^2, using the window size $s_{\rm evt}$ as the number of degrees of freedom.

It should be noted that the p-value used in Equation (5) can only be used as a true probability estimate in the case of Gaussian errors with no σ-clipping. While such errors are used in the simulations described in this paper, the actual data collected by the survey will certainly have non-Gaussian tails in the error distribution. However, after a large amount of testing, these estimates have proven perfectly adequate for use as an event trigger.

While the description above is applied to a single light curve, TAOS II will measure light curves for each star using all three telescopes. Thus, for each telescope j, we have a time series I_{jn}, which is the brightness of a target star measured at time t_n by telescope j. The measurements by the three TAOS II telescopes are taken simultaneously, so t_n is identical for all three measurements. After running the rolling filter over the light curve for each telescope, we have a pair of values μ_{jn} and σ_{jn} for each telescope at every point in the light curve. Similarly, we calculate the value of χ_{jn}^2 at each point n using Equation (4), and our variability test value is thus

$$p_{jn} = p_{\chi^2}(\chi_{jn}^2;\, s_{\rm evt}). \tag{6}$$

We can set a variability threshold for each telescope on p_{jn}, and in addition, we can combine all three test statistics to create a single threshold for the complete three-telescope event. Using Fisher's method (Fisher 1925), we calculate the combined χ^2 and resulting p-value as

$$\chi_{{\rm C}n}^2 = -2\sum_{j=1}^{N_{\rm tel}} \ln p_{jn} \tag{7}$$

and

$$p_{{\rm C}n} = p_{\chi^2}(\chi_{{\rm C}n}^2,\, 2 \times N_{\rm tel}), \tag{8}$$

where $N_{\rm tel} = 3$ is the number of telescopes.

In order to keep our false positive rate to a manageable level, we will set a high threshold on $p_{{\rm C}n}$ for a detection of a candidate event. In addition, we will set lower thresholds on p_{jn} in order to ensure that the simultaneous variability is seen in the light curves from all three telescopes, and not dominated by a large, spurious event from a single telescope. In order to set trigger levels in the ranges of small, easy to remember numbers, we define two event significance metrics as

$$\begin{aligned} T_{jn} &= -\log p_{jn} \\ T_{{\rm C}n} &= -\log p_{{\rm C}n}, \end{aligned} \tag{9}$$

and for the detection of a candidate event at point n, we thus require that

$$\begin{aligned} &T_{jn} > T_{\rm 1tel} \qquad \text{for each telescope } j, \\ &T_{{\rm C}n} > T_{\rm 3tel}, \end{aligned} \tag{10}$$

where we adopt trigger threshold values $T_{\rm 1tel} = 3$ and $T_{\rm 3tel} = 10$. Note that these values are used in our tests of the pipeline, but they are subject to change as the survey begins collecting data and we have better estimates of the false positive event rate as a function of the different detection thresholds.

As an example to illustrate this algorithm, we define a "test event" which is an inverted top hat function (hereinafter *U-shaped event*) with a width of 11 points and a depth of 0.5 on an otherwise flat light curve. Assuming a SNR of 10 (Gaussian noise), we generate the three telescope light curves shown in Figure 2. In Figure 3 we display the results of the event trigger search, assuming an event window size of $s_{\rm evt} = 11$, matching the duration of the event. In Figure 3(a), the values for χ_{jn}^2 are displayed at points n around the event shown in Figure 2. Similarly, Equation (3)(b) shows the values of $\chi_{{\rm C}n}^2$ versus n around the event. In both panels, the values of χ^2 increase as the event window moves over the event, reaching a maximum when the window is centered on the event. Figures 3(c) and (d) show the corresponding values of T_{jn} and $T_{{\rm C}n}$ from the values of χ^2 shown in the above panels. These values also reach peaks when the event window is centered on the event.

As discussed above, in order to be sensitive to events with a wide range of durations (3 points to nearly 200), multiple values of $s_{\rm evt}$ are used to search for events in each light curve. However, while the algorithm is on average most sensitive when $s_{\rm evt}$ is nearest to the event duration, it is possible that the same event will be detected using other values of $s_{\rm evt}$ as well. Furthermore, for longer events, the same event may be detected at multiple epochs. For the event parameter estimation discussed in the next section, it is necessary to have a single trigger for each event. The epoch of the trigger, which we define as the point $n_{\rm T}$ in the three telescope light curve set, as well as the window size $s_{\rm T}$ corresponding to the largest value of $T_{{\rm C}n}$, are needed to bootstrap the parameter estimation process. Therefore, a step has been added to the algorithm to not only ensure that any event only passes the event trigger one time,

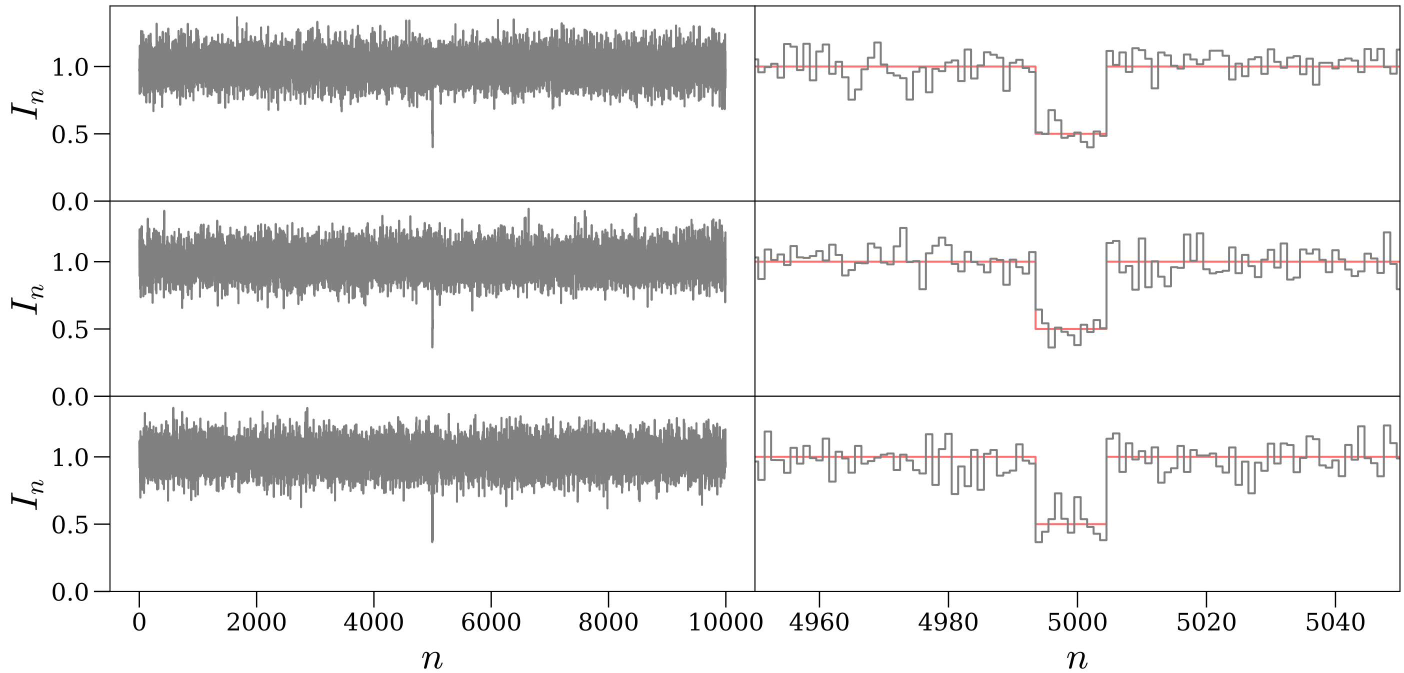

Figure 2. Simulated three-telescope light curve set with a simple inverted top hat test event. The width of the event is 11 points, the depth is 0.5 and the SNR is 10. The light curves have length $N_{\rm LC} = 10,000$ points. Left panels: Full 10,000 point light curves. Right panels: detailed view of the event in the center of the light curves shown in the left panels.

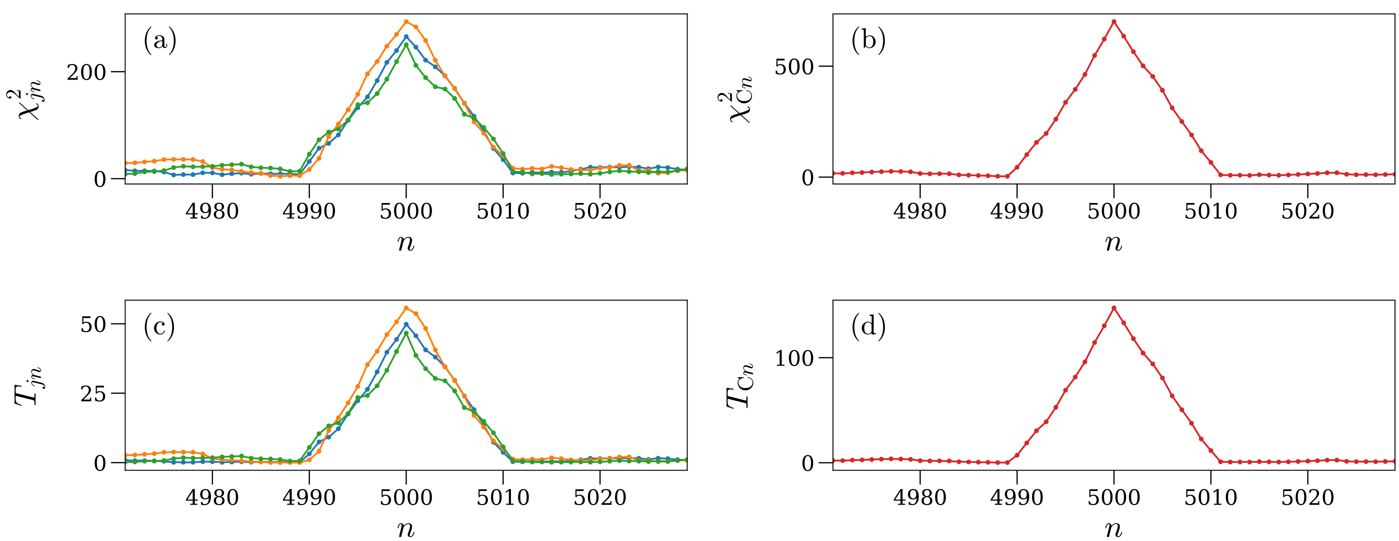

Figure 3. Examples of rolling χ^2 and p-values for the test event shown in Figure 2. Panel (a): χ^2_{jn} for each telescope j vs. index n near the event. Panel (b): $\chi^2_{{\rm C}n}$ vs. n for points n around the event center. Panel (c): T_{jn} from the values of χ^2_{jn} shown in Panel (a). Panel (d): $T_{{\rm C}n}$ from the values of $\chi^2_{{\rm C}n}$ shown in Panel (b).

but also to provide the values of $n_{\rm T}$ and $s_{\rm T}$ for the most significant trigger.

First, for each value of $s_{\rm evt}$, for any pair of triggers separated by $s_{\rm evt}$ points or fewer, the trigger with the lower value of $T_{{\rm C}n}$ is discarded. Second, for every pair of consecutive values of $s_{\rm evt}$ in the list of possible event window sizes, if two triggers are found within the sum of the two values of $s_{\rm evt}$, the trigger with the lowest value of $T_{{\rm C}n}$ is also discarded, and the most significant event is added to the list of triggers for the larger value of $s_{\rm evt}$. This process is repeated from the smallest to the largest window size, and we are then left with a single trigger for any given event. The values of n and $s_{\rm evt}$ for the most significant trigger are saved as $n_{\rm T}$ and $s_{\rm T}$, which are used in the event shape parameter estimation described below. Note that if there are multiple occultation events within the maximum window size of 197 points, the less significant events will be discarded by this algorithm. However, given that the event rate is $\lesssim 10^{-3}$ per star per year, the odds of this happening are negligible.

3. Event Parameter Estimation

The characteristics of occultation event light curves are determined by a significant number of parameters, and are dominated by diffraction for TNOs with diameters $D \lesssim 10$ km (Roques et al. 1987; Nihei et al. 2007; Castro-Chacón et al. 2019). Simulated light curves are thus somewhat complicated to calculate, and once a candidate event passes the trigger, estimating all of the free parameters is a very CPU intensive and time consuming process. However, a rough estimate of TNO size and distance is actually all that is needed for the detection efficiency simulation, where we just need to see if any triggered event looks reasonably close to the corresponding simulated event. Furthermore, for any event detections, rough estimates of these parameters could be used to bootstrap any likelihood analysis performed to find better parameter estimates.

Nihei et al. (2007) demonstrated that the width and depth of an occultation event can be estimated directly from the TNO diameter D and distance Δ. Therefore, a reasonable method would be to estimate the width and depth of the occultation event, and use the method shown in to reverse calculate D and Δ. Given that the "width" and "depth" of an event in a noisy light curve are not well defined, these parameters must be estimated by some sort of fitting routine. However, occultation event light curves have a variety of shapes, so the width and depth would only be accurate of the correct shape were used in the fit. Furthermore, our early attempts to perform simultaneous two-parameter fits found significant covariance between the width and depth, depending on the noise level and event shape, giving rise to large uncertainties in both parameters.

See Figure 4 for an example of a significant event in a very noisy light curve. The width and depth of this event are not well defined, and fitting different shapes would give significantly different results.

3.1. Equivalent Width

A useful parameter that does not have the problems outlined above is the *equivalent width* (Sicardy et al. 1991), which is defined as

$$w_{\rm eq} \equiv \sum_n (1 - F_n), \tag{11}$$

where

$$F_n \equiv \frac{I_n}{\langle I \rangle} \tag{12}$$

is the normalized light curve, such that $F_n = 1$ indicates the nominal brightness of the star. For our analysis pipeline, we will set

$$\langle I \rangle = \mu_n, \tag{13}$$

where μ_n is the rolling mean calculated from the filter window as described in Section 2.2.

Roques et al. (2003) used this parameter as a filter for the detection of occultation events, and while we have opted for a different method (see Section 2), this value is useful to parameterize the shape of an event because it can be estimated independently of the event shape itself. However, obtaining an accurate estimate of $w_{\rm eq}$ is not as easy as it may appear from Equation (11). Most importantly, note that the set of points n over which the summation is performed in Equation (11) is not defined. Determining which points go into the sum runs into one of the same problems outlined above, namely that the width of the event is not well defined. In order to determine $w_{\rm eq}$ as well as possible, all of the points in the event itself should be included (that is, all of the points with any significant deviation from a flat light curve). For a noisy light curve like that shown in Figure 4, the larger the number points outside of the event are included in the sum, the larger the uncertainty in the value of $w_{\rm eq}$ itself. Furthermore, the noisier the light curve, the more likely that the value of $s_{\rm T}$ from the trigger will be farther from the actual width of the event, and we have found that using this value does not work well in many cases.

To estimate $w_{\rm eq}$, we perform an iterative calculation over an increasing number of points u, starting at the center of the event. We define the function

$$W(u) \equiv \sum_m (1 - F_m), \tag{14}$$

where the sum is over the range

$$n_{\rm T} - \frac{u-1}{2} \leqslant m \leqslant n_{\rm T} + \frac{u-1}{2}, \tag{15}$$

where

$$u \in \{1, 3, 5, \cdots\} \tag{16}$$

is the number of points in the sum, and $n_{\rm T}$ is the index n at the center of the event as determined by the trigger (see Section 2.3). We define the function $W(u)$ as the *Equivalent Width Curve* (EWC).

This is explained more clearly in Figure 5, where the calculation of $W(u)$ is illustrated with the (noiseless) test event described in Section 2.3. Starting with $u = 1$ (top panel), the sum is simply equal to $1 - F_{n_{\rm T}}$. For each subsequent panel, u is increased by two and the additional point on each side of the current set of points is added to the sum. Eventually u is large

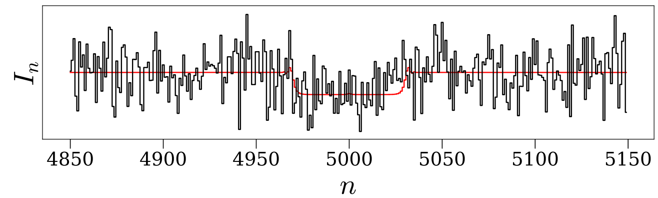

Figure 4. A simulated occultation event in a noisy light curve. The simulated event is from a $D = 30$ km TNO at $\Delta = 43$ au and an opposition angle $\phi = 60°$ (corresponding to a relative velocity between the observer and TNO of about 14 km s^{-1} (Nihei et al. 2007, Equation (8))) occulting a $R = 12$ G0V star. The noiseless simulated event is shown with the red line, while the black line is a simulation of the measured data assuming a SNR of 1.5. While this is a very low value of SNR, when $s_{\mathrm{evt}} = 59$ is used, we find a maximum $\chi_n^2 = 107$, and a corresponding trigger value of $T_n = 3.9$. This event is significant and passes the single telescope trigger outlined in Section 2.

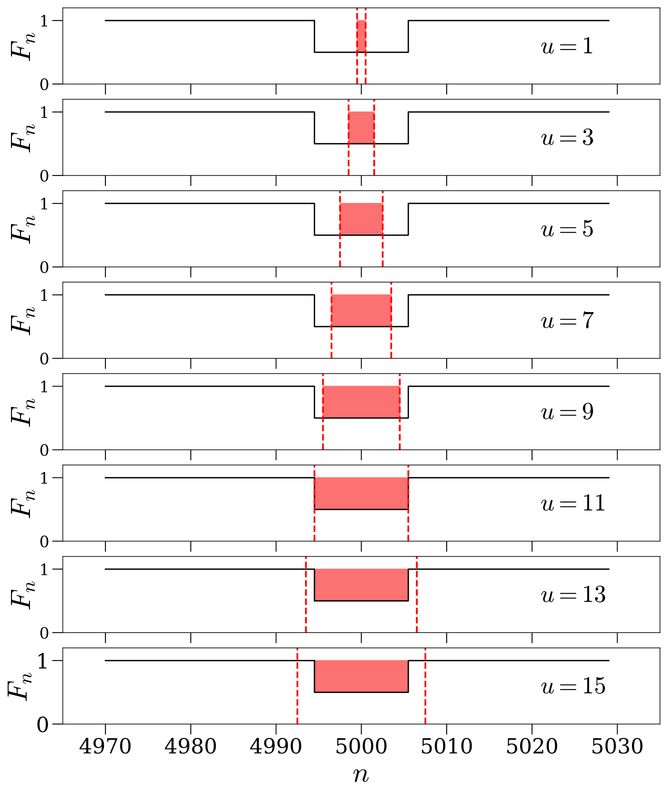

Figure 5. Illustration of the calculation of $W(u)$ for increasing values of u for the noiseless test event described in Section 2.3, assuming $n_{\mathrm{T}} = 5000$. The red area indicates value of W, and the dashed red lines show the limits of the sum in Equation (14).

enough that all of the points in the occultation event are added to the sum.

Example EWCs are shown in Figure 6 for the standard test event defined in Section 2.3. Figure 6(a) shows the light curve of the event with a SNR of 100, while Figure 6(b) shows the light curve with a SNR of 6. Figures 6(c) and (d) show the corresponding EWCs. As one would expect,for the high SNR light curve W increases linearly with u until the edge of the event is reached at $u = 11$. At that point, the EWC levels off to a constant at the correct value of w_{eq} for this event. On the other hand, for the noisy event shown, W increases with u, and then flattens near the edge of the event. However, given the fact that the EWC is noisy itself due to the noise in the original light curve, there is no convergence to a clear value of w_{eq} from this curve. Nevertheless, a reasonable estimate of w_{eq} can be obtained by performing a simple least squares fit to to the EWC.

3.2. Equivalent Width Curve Fitting

We model the EWC using two components: a line with a slope d_{eq} and intercept at 0, and a line with a slope of 0 and an intercept at w_{eq}. To fit these functions, we need to define the point v, the point at which the function changes from one form to the other. We are thus fitting a simple function of the form

$$\begin{aligned} W &= d_{\mathrm{eq}} \times u(u < v) \\ W &= w_{\mathrm{eq}}(v < u \leqslant u_{\max}), \end{aligned} \tag{17}$$

where $u_{\max}$ is the maximum value of u used in the fit. Since u is an odd integer, we set the possible values of v to the set of even integers

$$v \in \{4, 6, 8, \cdots, v_{\max}\}, \tag{18}$$

where we set

$$v_{\max} = u_{\max} - 15. \tag{19}$$

Remembering that u increases by two intensity measurements for each point, a minimum value of 4 for v ensures that at least two values of u are used to determine d_{eq}, and the value of $v_{\max}$ is chosen to ensure that there are at least 8 values of u beyond v to get an accurate estimate of w_{eq}. We begin the process with

$$u_{\max} = \max(29, 3.5 s_{\mathrm{T}}), \tag{20}$$

rounding up to the nearest odd integer. We set a minimum value of $u_{\max} = 29$ in order to ensure that at least 16 intensity measurements are used in the fit to estimate w_{eq} for shorter duration events while simultaneously covering a reasonable range of values for v. The second argument to max() in the above equation is used to find the nearest odd integer that is at least 3.5 times the event width we roughly estimate from the value of s_{T} from the trigger. This should cover both the event itself and enough points beyond the event for a good estimate of w_{eq} for longer duration events. However, the value of $u_{\max}$ can be increased if necessary, as discussed later in this section.

To perform the least squared fit, the variance σ_{eq}^2 must be calculated correctly. Recall that every point in the EWC is a sum of all of the points up until the current value of u, so the variance for each term must be added in. The variance is thus

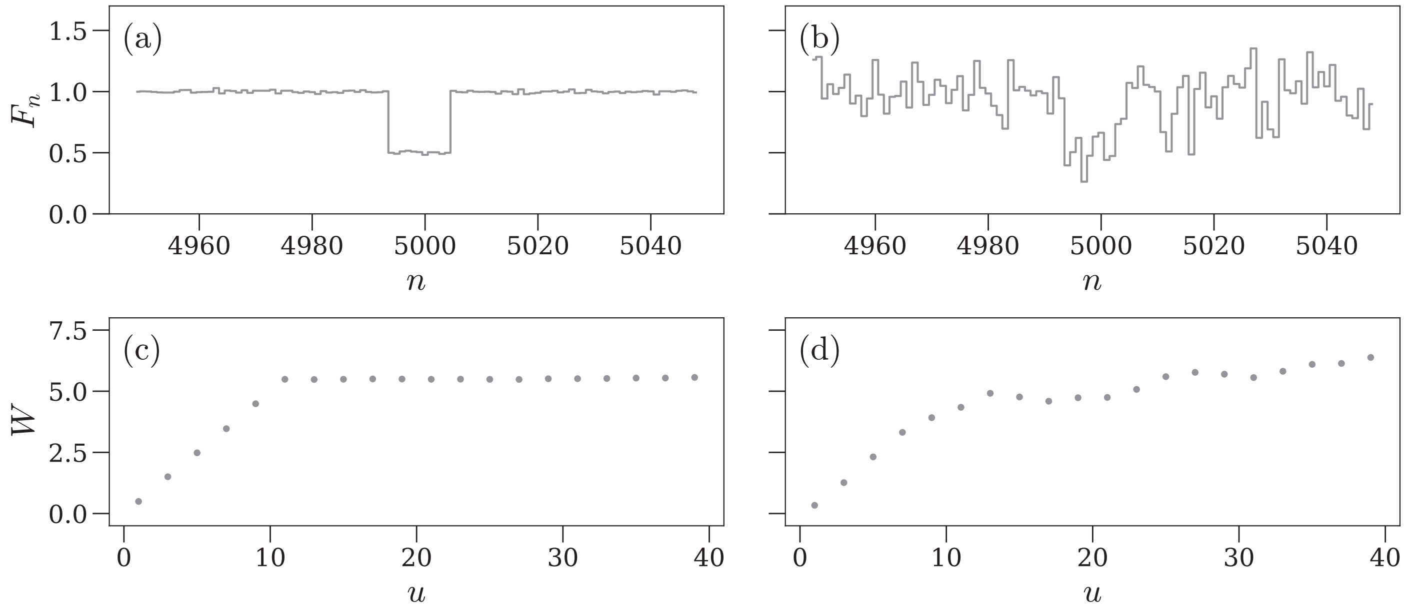

Figure 6. Example light curves and EWCs for the test event defined in Section 2.3. Panel (a) shows the event with a SNR of 100, and panel (b) shows the same with a SNR of 6. Panel (c) shows the EWC for the light curve shown in panel (a), and panel (d) shows the same but for the noisier light curve shown in panel (b). In both EWC plots, $W(u)$ is flattens when u is larger than the event width, but there is not a clear value of w_{eq} for the noisier event.

given by

$$\sigma_{\mathrm{eq}}^2(u;\, n) \equiv \sum_m \sigma_m^2, \tag{21}$$

where the σ_m^2 is the rolling variance calculated from the filter windows (see Section 2.2) and the sum is over the range given by Equation (15). The best fit parameters to the EWC are then given by

$$\langle d_{\mathrm{eq}}(v)\rangle = \frac{\sum_{m=0}^{\frac{1}{2}v-1} \frac{(2\,m+1)\times W(2\,m+1)}{\sigma_{\mathrm{eq}}(2\,m+1)}}{\sum_{m=0}^{\frac{1}{2}v-1} \frac{(2\,m+1)^2}{\sigma_{\mathrm{eq}}(2\,m+1)}},$$

$$\langle w_{\mathrm{eq}}(v)\rangle = \frac{\sum_{m=\frac{1}{2}v}^{\frac{1}{2}(u_{\max}-1)} \frac{W(2\,m+1)}{\sigma_{\mathrm{eq}}(2\,m+1)}}{\sum_{m=\frac{1}{2}v}^{\frac{1}{2}(u_{\max}-1)} \frac{1}{\sigma_{\mathrm{eq}}(2\,m+1)}}. \tag{22}$$

Of course, the correct value of v is not known a priori, so a χ^2 minimization is performed. We calculate χ^2 for each part of the EWC, and then use Fisher's method again to combine them into a single value χ_{C}^2 (see Equations (6) and (7)). We find the value $\langle v\rangle$ that minimizes χ_{C}^2, and use the corresponding values of $\langle w_{\mathrm{eq}}\rangle$ and $\langle d_{\mathrm{eq}}\rangle$ as the event shape parameters. Figure 7 shows two examples of this EWC fitting algorithm using the high and low SNR light curves shown in Figure 6. In both cases, the best fit values are fairly close to the correct values.

It is important to ensure enough points are used to get an accurate estimate of $\langle w_{\mathrm{eq}}\rangle$, so $u_{\max}$ may be increased if necessary. If we find the minimum value of χ_{C}^2 is at the largest allowed value of $v \geqslant v_{\max} - 4$, then we cannot be confident that a true minimum as been reached. In this case, we enlarge $u_{\max}$ by a factor of 1.2 (rounding up to the nearest odd integer), and the fitting process is repeated. We iterate this until $\langle v\rangle < v_{\max} - 4$, at which point the values of $\langle w_{\mathrm{eq}}\rangle$ and $\langle d_{\mathrm{eq}}\rangle$ corresponding to the last best fit are returned. We set a hard upper limit of $u_{\max} < 1000$, and if we do not find a minimum value of $\chi_{\mathrm{C}}^2(v)$ by this point we label the event as a false trigger and flag it for further inspection.

3.3. EWC Slope as an Estimate of Event Depth

As discussed earlier in this section, Nihei et al. (2007) have shown that the event width and depth can be used to estimate D and Δ. For the example of a U-shaped event, the equivalent width is simply the product of the width and depth of the event. Given that

$$w_{\mathrm{eq}} = \mathrm{width} \times d_{\mathrm{eq}}, \tag{23}$$

(see Figure 7(c) and Equation (17)) it follows that the value of d_{eq} corresponds to the depth of the event. The slope was initially used as part of the fitting routine designed to improve the estimate of w_{eq}. However, it is clear the both d_{eq} and w_{eq} can be used to parameterize an event in a way similar to using the width and depth.

However, it turns out that getting an accurate estimate of d_{eq} depends on starting the EWC fit at the correct center point of the event. Recall that we are using the point n_{T}, the point n with the highest value of $T_{\mathrm{C}n}$ as an estimate of the event center. However, in many cases (particularly in the case of a long duration U-shaped event), the real center of the event will very

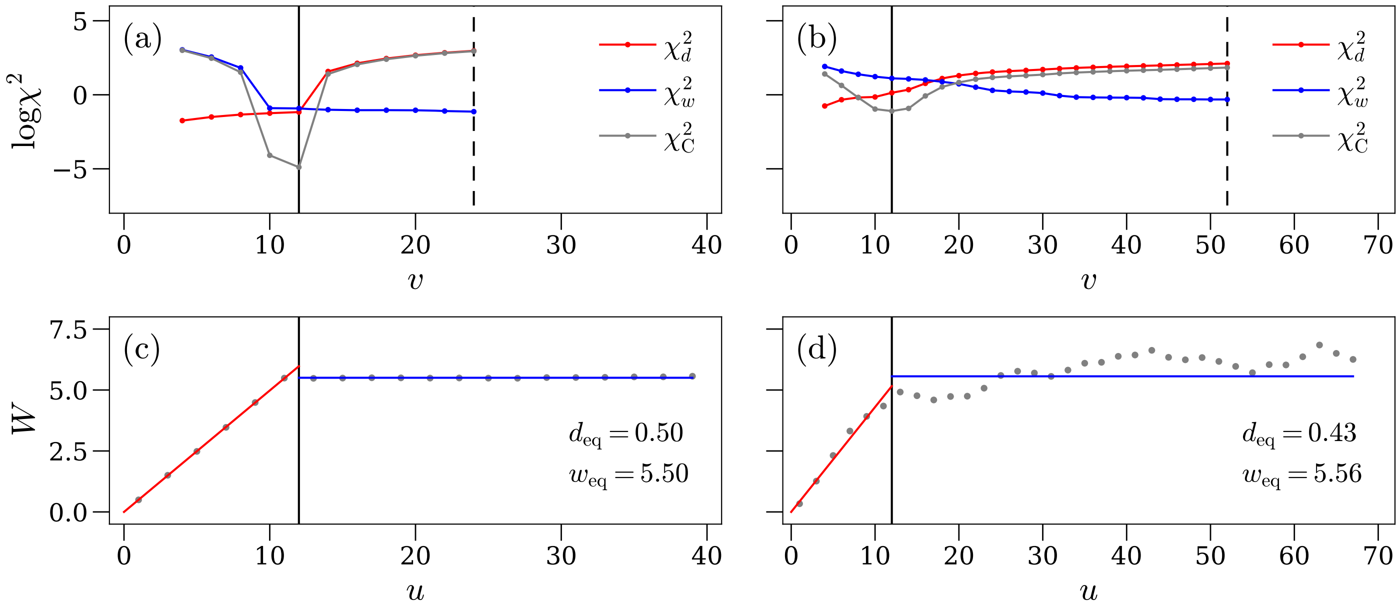

Figure 7. Shape parameterization example for U-shaped test event. Panel (a) shows $\log \chi^2(v)$ vs. v for the event high SNR light curve shown in Figure 6(a). The values of χ^2 are shown both for the fits to the slope and equivalent width, as well as the combined χ^2. The solid vertical line shows v at the minimum χ^2_C, and the dashed vertical line indicates v_{max}. Panel (b) shows the same for the low SNR light curve shown in Figure 6(b). Panel (c) shows the EWC and the best fit for the event for the high SNR event, and panel (d) shows the same for the low SNR light curve. In both cases, the fit values are very close to the known values of $d_{eq} = 0.5$ and $w_{eq} = 5.5$. Note that v_{max} is larger for the noisier light curve. This is because the best trigger was found with s_{evt} quite a bit larger than the event duration due to some low points on right side of the light curve.

possibly be in a different location. In order to improve our estimate of the event center, we have developed an event centering algorithm, which will be described in Section 3.4.

3.4. Event Centering Algorithm

As shown in Figure 8, the value of d_{eq} will change with the event center. This is because if the center used in the fit n_{fit} is away from the true center n_c, then a point u will be reached where, instead of adding two values to $W(u)$ from inside the event in subsequent steps, one would add a point from inside the event which contributes to the sum w_{eq} and another point outside of the event which does not. One would thus expect that the maximum value for d_{eq} should be found when the true center is used. This is illustrated in Figure 9, which shows d_{eq} versus n_{fit}. In both plots, the curve is peaked near the true event center n_c, but in the case of the noisier event it is shifted from the peak, indicating that finding the maximum value of d_{eq} is not the best way to find the event center.

We have found that a better way to estimate the value of n_c is by taking the weighted mean of the distribution of d_{eq} versus n_{fit}, that is

$$\langle n_c \rangle = \frac{\sum_{m=n_c'-N_c}^{n_c'+N_c} n_{fit\,m} d_{eq\,m}}{\sum_{m=n_c'-N_c}^{n_c'+N_c} d_{eq\,m}}, \tag{24}$$

where

$$N_c = \max(5, s_T), \tag{25}$$

and n_c' is initialized to the initial center estimate n_T from the trigger. Since we are doing a mean over a finite number of points, if we start far from the true event center then the distribution would be skewed, possibly leading to an inaccurate estimate of $\langle n_c \rangle$. We thus iterate this algorithm, replacing n_c' with $\langle n_c \rangle$ at each step, until we converge on a value of n_c such that

$$|\langle n_c \rangle - n_c'| < 0.5. \tag{26}$$

At this point, we set the true event center n_c to the nearest integer value.

As seen in Figure 9, this method works very well in finding the correct event center. Once the new event center is found, we use the values of w_{eq} and d_{eq} calculated using the best estimate of n_c as the event shape parameters.

This algorithm converges very quickly, 99% of the time in only two iterations in the testing we have performed. However, on occasion the algorithm does not converge. We have found that when it takes more than ten iterations, the value of $\langle n_c \rangle$ oscillates between two adjacent points as the same two points are added and removed from either end of the sum in Equation (24). When this occurs, we simply include both points in the sum and use the resulting value of $\langle n_c \rangle$ for the final estimates of w_{eq} and d_{eq}.

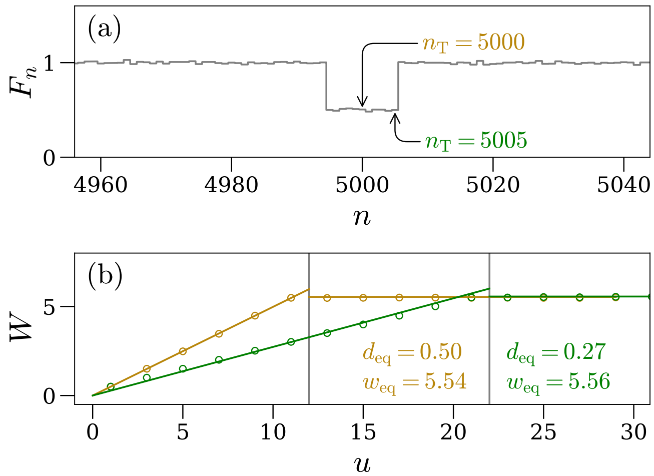

Figure 8. Two EWC fits starting from two different event centers. The top plot shows the event, and the two points used as event centers are indicated. The bottom plot shows the two fits resulting from the two values of n_T. The brown points are calculated assuming $n_T = 5000$ (correct event center), and the green points are calculated assuming $n_T = 5005$. If n_T is away from the true event center, the slope d_{eq} is underestimated. However, the effect of the choice of event center on the estimate of w_{eq} is minimal.

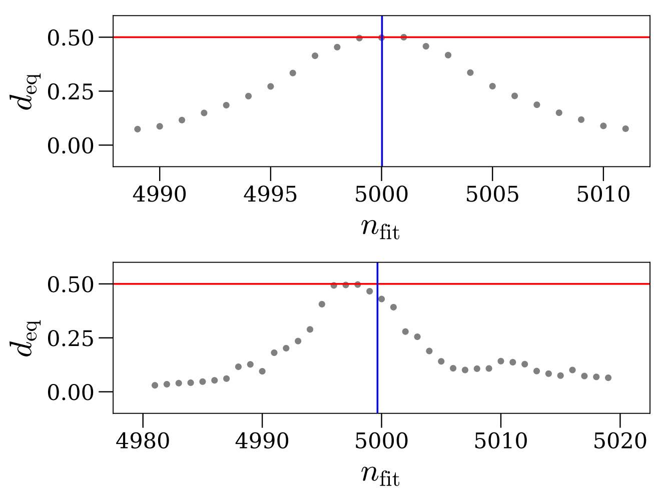

Figure 9. Plots of d_{eq} vs. event center used in the EWC fit for the same events shown in Figure 6. The top panel shows the curve from the high SNR event (Figure 6(a)), where the event is centered at $n_c = 5000$. The lower panel shows the curve from the lower SNR event (Figure 6(b)). In both panels, the blue line indicates the location of the weighted mean. For the top panel, the weighted mean gives $\langle n_c \rangle = 5000.02$, leading to the correct estimate of n_c. For the bottom panel, the weighted mean is $\langle n_c \rangle = 4999.66$, also giving the correct value. Note that while the peak of the curve is fairly far left of the correct value, the tail on the right-hand side is larger, offsetting the weight of the peak.

4. Shape Parameters of Occultation Events

So far we have only considered the simple U-shaped test event. However, occultation light curves will have a variety of shapes, so it must be demonstrated that our parameter estimation algorithm will work with actual events.

4.1. Spherical TNOs

The shape of an occultation light curve is determined by a large number of parameters; however, most of these are known from the survey design and observation target. In the simplest case of spherical TNOs, the two free parameters of interest are the diameter D and the distance Δ between the observer and the occulting object. The other two free parameters are the impact parameter b (the distance of closest approach between the center of the occultation shadow and observer) the time offset t_{off} (the time difference between the epochs at the point of closest approach and the center of the closest exposure). The parameters b and t_{off} are simple geometric variables with uniform distributions, and are only of interest in how they affect the shapes of the light curves.

A selection of light curves covering a range of D and Δ are shown in Figure 10, along with their corresponding EWCs and fits. For the larger, closer objects presented in this figure, our model of two straight lines for the EWC works quite well, but for the other events this model is not nearly as good. However, we are not trying to accurately model the EWC, we are calculating simple shape parameters that should correspond to the physical parameters of interest. Note that we use $\mathrm{SNR} = 10^6$ for the light curves in Figure 10 to approximate noiseless light curves. A finite value of SNR is needed in order for Fisher's method to work during the χ^2 minimization for the EWC fit.

Plots of w_{eq} versus d_{eq} for a variety of sizes and distances are shown in Figure 11. Figure 11(a) shows contours of w_{eq} versus d_{eq} for constant diameters D and varying Δ, assuming no noise, observations at opposition, and both $b = 0$ and $t_{off} = 0$. Figure 11(b) shows the same, but with an opposition angle of $\phi = 60°$, where the relative velocity at $\Delta = 43$ au is about $14\ \mathrm{km\ s^{-1}}$, compared with about $25\ \mathrm{km\ s^{-1}}$ at opposition (Nihei et al. 2007, Equation (8)). With a lower relative velocity, the events have longer durations, increasing the equivalent widths while the slopes remain roughly the same. Contours corresponding to two target stars with different angular sizes are shown with the dashed and solid lines in this plot. The Poisson spots in the diffraction shadows are averaged out in the cases with the star with the larger angular size (see Nihei et al. 2007, Figures 7–11), giving rise to measurable differences in the contours. Figure 11(c) shows the same as panel (a), but for both D and Δ constant with varying b. The dashed lines show contours of constant D with varying Δ, with $b = 0$. The intersections of the solid and dashed curves indicate degeneracy among the shape parameters when the diameter, distance, and impact parameter are all considered. Finally, Figure 11(d) shows the same as panel (a), but for both D and Δ constant with varying t_{off}. Clearly, the time offset has negligible impact on the parameter estimation.

4.2. Non-spherical TNOs

(Castro-Chacón et al. 2019, hereinafter C19) presented a simulator to generate occultation shadows and light curves for

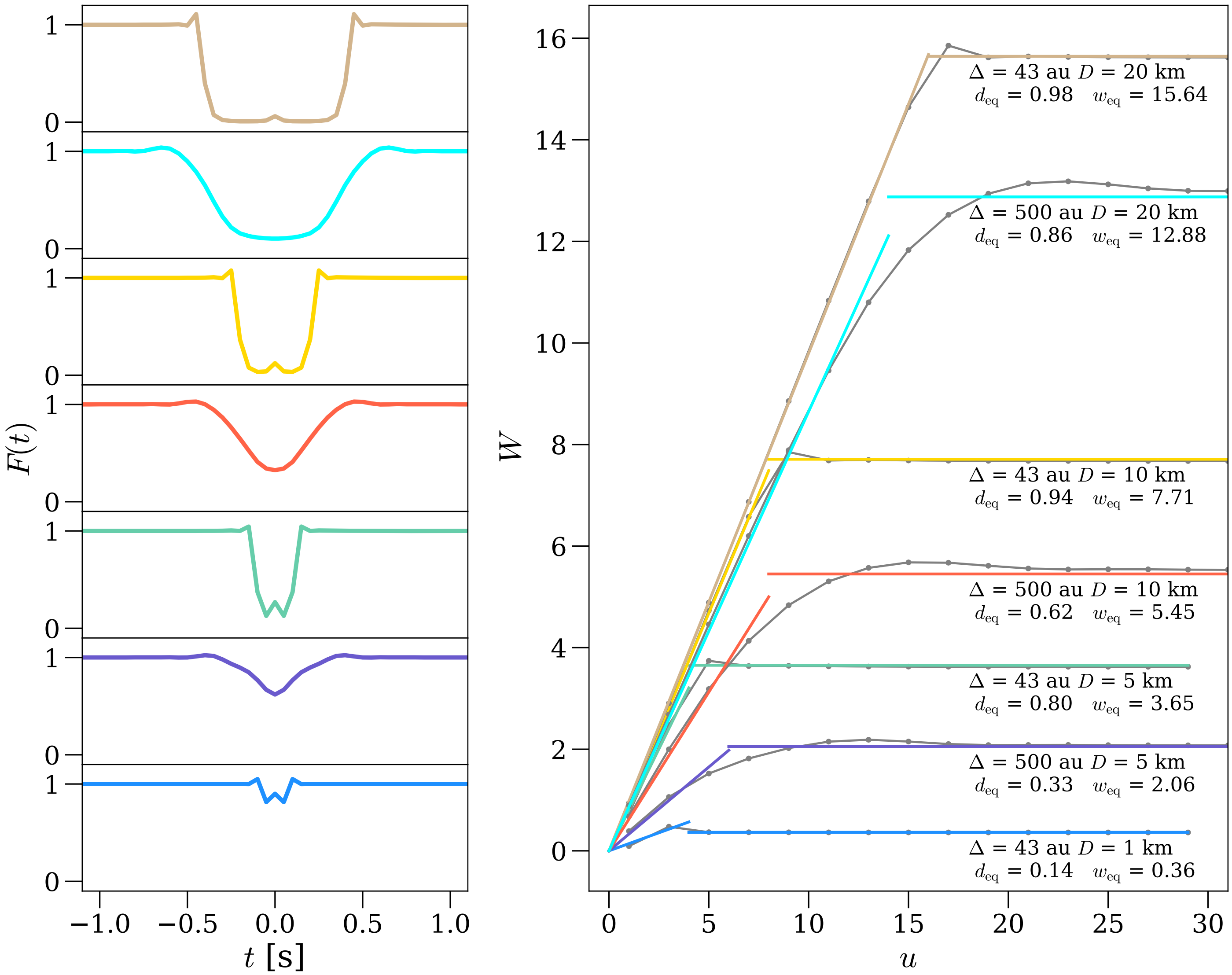

Figure 10. EWC fits for simulated occultation events, assuming a $R = 12$ G0V target star at an opposition angle of $\phi = 0°$ and SNR $= 10^6$. Ranges of both D and Δ are covered. The left panel shows the light curves of the simulated events, and the right panel shows the corresponding EWCs and fits using the same colors. Note that each line for the fits has an endpoint at $u = v$, and the lines to not necessarily intersect because EWC fits to the two curves are performed independently.

non-spherical TNOs. The simulator can work with any TNO shape, but for our testing purposes we will use the compact binary shape shown in C19, Figure 6. Furthermore, we note that the term *diameter* is not well defined for non-spherical objects, so we adopt the C19 definition of the effective diameter

$$D_{\text{eff}} = \sqrt{\frac{A}{4\pi}}, \tag{27}$$

where A is the projected area of the occulting TNO. Finally, due to the fact that the occultation shadows do not have axial symmetry, C19 introduced a new parameter $\hat{b}$ which they labeled the *reading direction*, which is the angle of the chord across the diffraction shadow along which a light curve is measured.

C19 considered occultations for three different regimes, defined in terms of the Fresnel scale

$$F_{\text{s}} = \sqrt{\frac{\lambda\Delta}{2}}, \tag{28}$$

where λ is the nominal observing wavelength. For a TNO at a distance of $\Delta = 43$ au and $\lambda = 600$ nm (the average value for the TAOS II filters), $F_{\text{s}} = 1.4$ km. The first regime is the *far-field* domain, where $D_{\text{eff}} \lesssim F_{\text{s}}$. In this regime, C19 showed that the effects of TNO shape on the observed light curves are minimal, so we ignore this case in our testing. However, for the *mid-field* domain ($D_{\text{eff}} \gtrsim F_{\text{s}}$), the occultation shadow is strongly

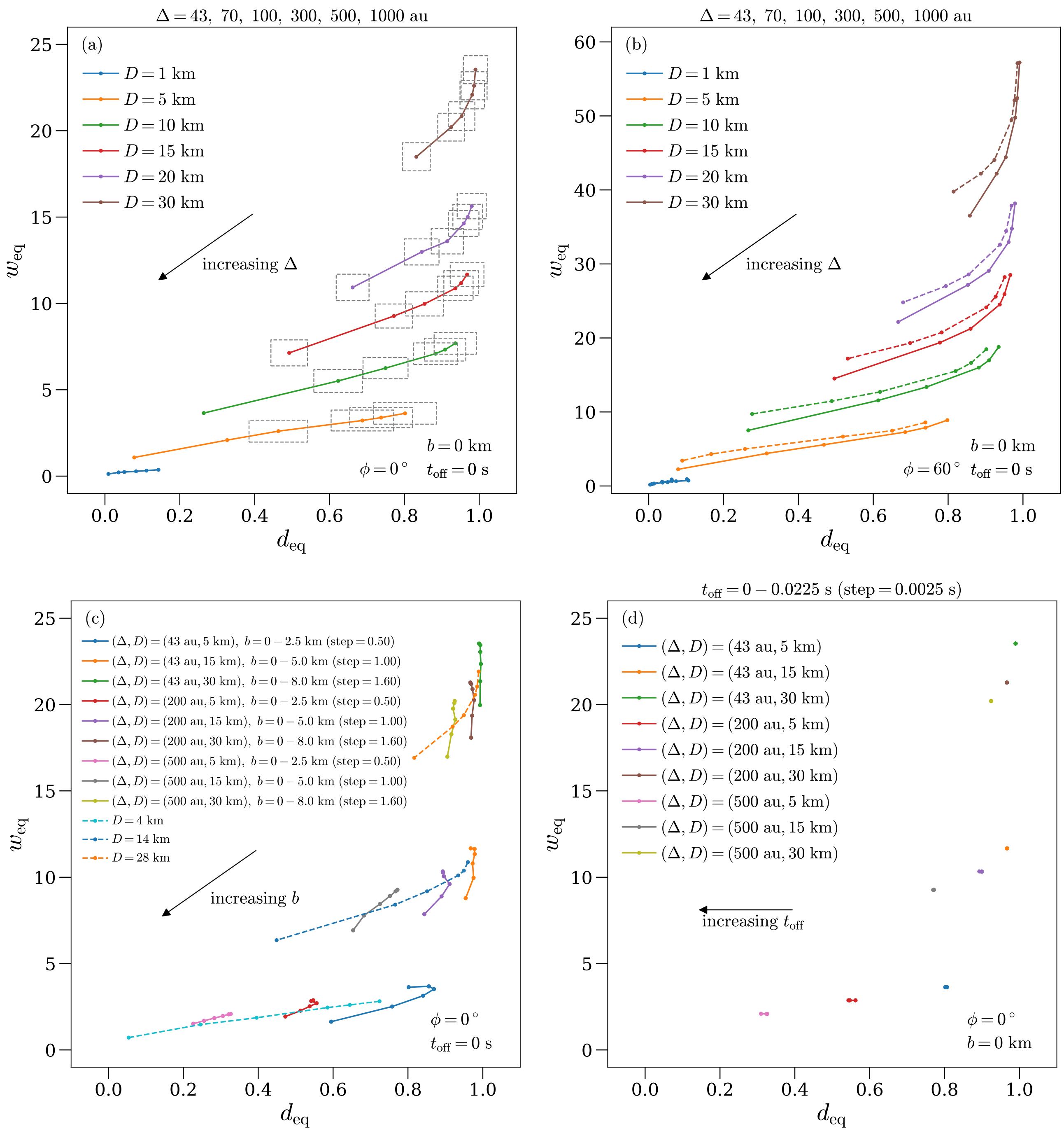

Figure 11. Contour plots of $w_{\rm eq}$ vs. $d_{\rm eq}$ over a range of parameters for occultations of a $R = 12$ G0V star. Panel (a) shows contours of $w_{\rm eq}$ vs. $d_{\rm eq}$ for constant diameters D over a range of distances Δ, assuming target star is at opposition. The boxes around each point indicate the standard deviations of the measurements of $w_{\rm eq}$ and $d_{\rm eq}$ at each point, assuming a SNR of 5. (Some points have no box since no events pass the event trigger at this SNR). Panel (b) shows the same, but assuming an opposition angle of $\phi = 60^\circ$. The events get wider at larger values of ϕ due to the lower relative velocity between the TNO and observer (Nihei et al. 2007, Equation (8)), leading to larger values of $w_{\rm eq}$. The solid lines show contours for the $R = 12$ mag star used in panel (a), while the dashed lines show contours for a $R = 18$ mag target star. The angular radius of the $R = 18$ mag star is 4.7×10^{-12} rad, in comparison to 7.5×10^{-11} rad for the $R = 12$ mag star, leading to differences in the expected values of $w_{\rm eq}$ and $d_{\rm eq}$. Panel (c) shows contours (solid lines) of $w_{\rm eq}$ vs. $d_{\rm eq}$ for ranges of constant D and Δ, while varying the impact parameter b. Contours of constant D with varying Δ and $b = 0$ are shown with dashed lines. These lines intersect the contours with varying b, indicating that the introduction of non-zero impact parameters introduces some degeneracy in the dependence of D and Δ on $w_{\rm eq}$ and $d_{\rm eq}$. Panel (d) shows the same as panel (c), but with the offset time $t_{\rm off}$ varying rather than b. Note that there are ten points in each contour, but the differences between them are so small that the points cannot be resolved in this plot, indicating that the parameters $w_{\rm eq}$ and $d_{\rm eq}$ are practically independent of $t_{\rm off}$.

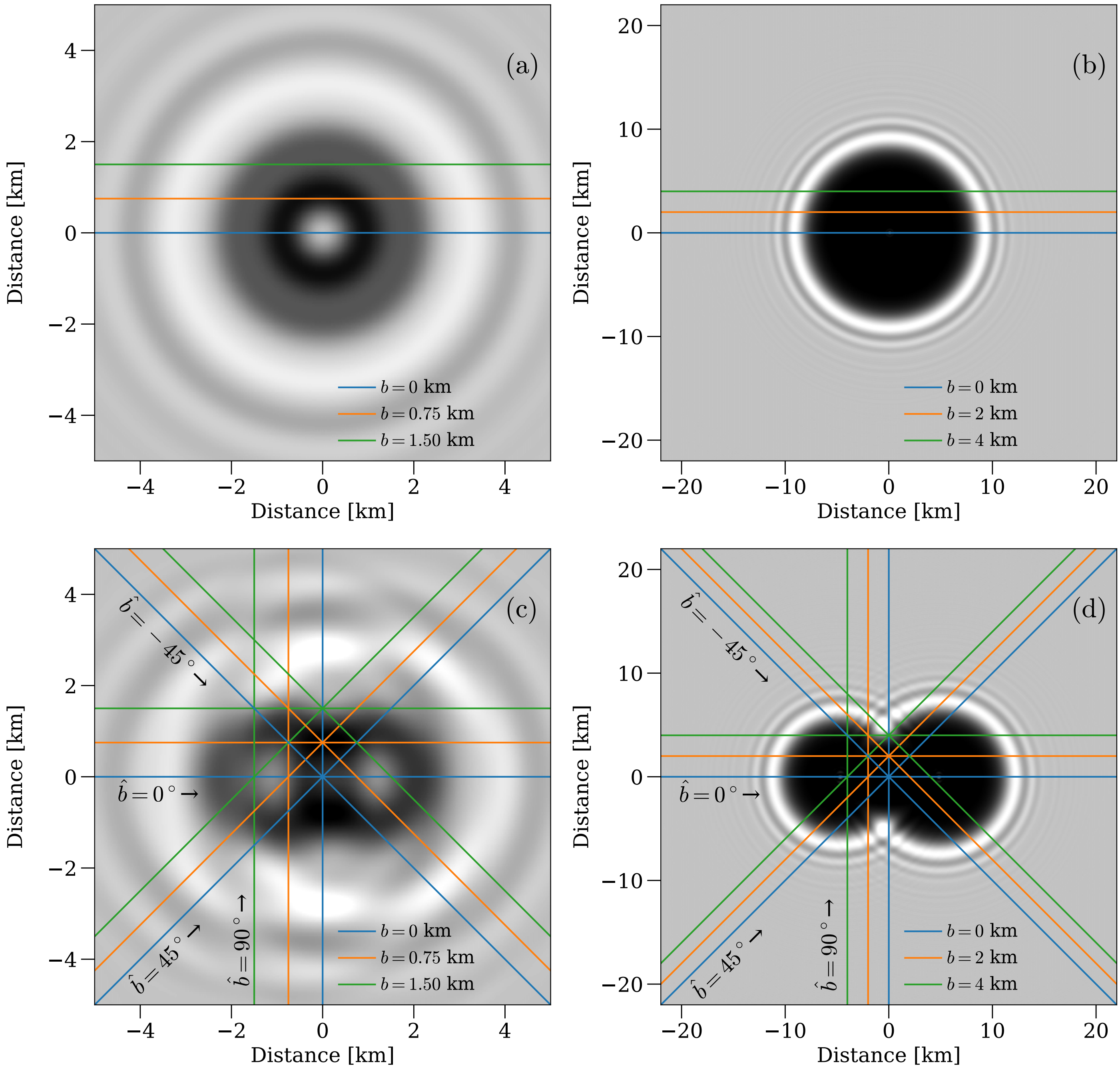

Figure 12. Diffraction shadows produced by compact binary TNOs, with the shadows produced by spherical TNOs also shown for comparison. All events are generated assuming a $R = 15$ G0V target star and TNO distance $\Delta = 45$ au. Panel (a) shows the shadow for a spherical $D_{\rm eff} = 3$ km TNO in the mid-field regime, and panel (b) shows the same for a spherical $D_{\rm eff} = 15$ km TNO in the near-field regime. Panels (c) and (d) show the same, but for TNOs with compact binary shapes. The colored lines indicate possible chords along which individual light curves could be measured, covering a range of impact parameters b, as well as a range of reading directions $\hat{b}$ for the compact binary cases.

dependent on the shape, with a large number of intersecting fringes. Finally, in the *near-field* regime ($D_{\rm eff} \gg F_s$), diffraction effects are minimal and the occultation shadow can be roughly approximated by the geometric shadow.

Example occultation shadows for non-spherical TNOs are shown in Figure 12, along with their spherical counterparts for comparison. Figures 12(a) and (b) show occultation shadows for $D = 3$ km (mid-field regime) and $D = 15$ km (near-field

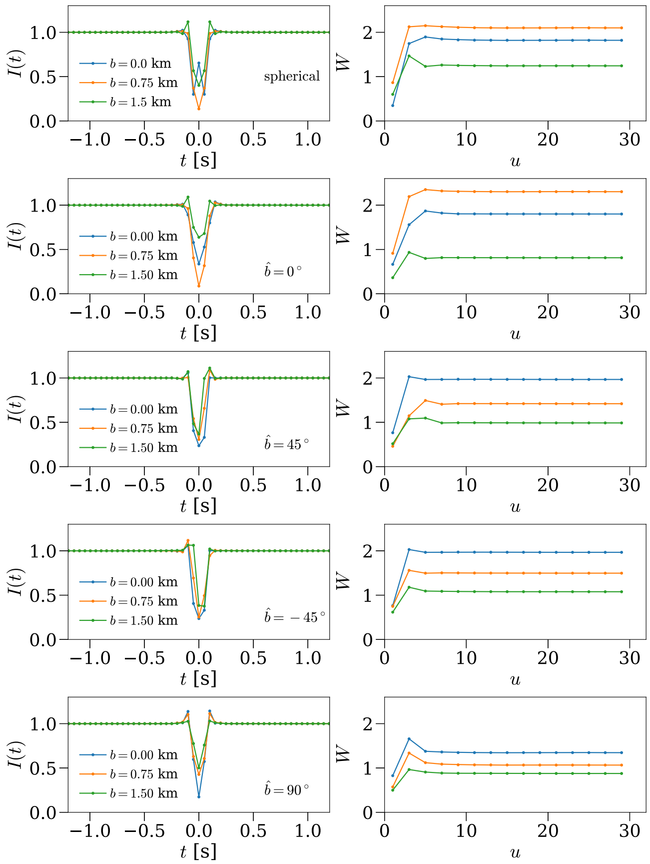

Figure 13. Example light curves (SNR $= 10^6$) and EWCs for $D_{\mathrm{eff}} = 3$ km spherical and compact binary TNOs corresponding to the events shown in Figures 12(a) and (c). The top panel on the left shows light curves for a spherical TNO, while the four lower panels on the left show light curves for different values of the reading direction $\hat{b}$. In each plot, light curves are shown with three different values of the impact parameter b. The chords across the occultation shadow are indicated in Figures 12(a) and (c) using lines of the same colors. The panels on the right show the EWCs for the light curves shown in the left panels, using lines of the same colors.

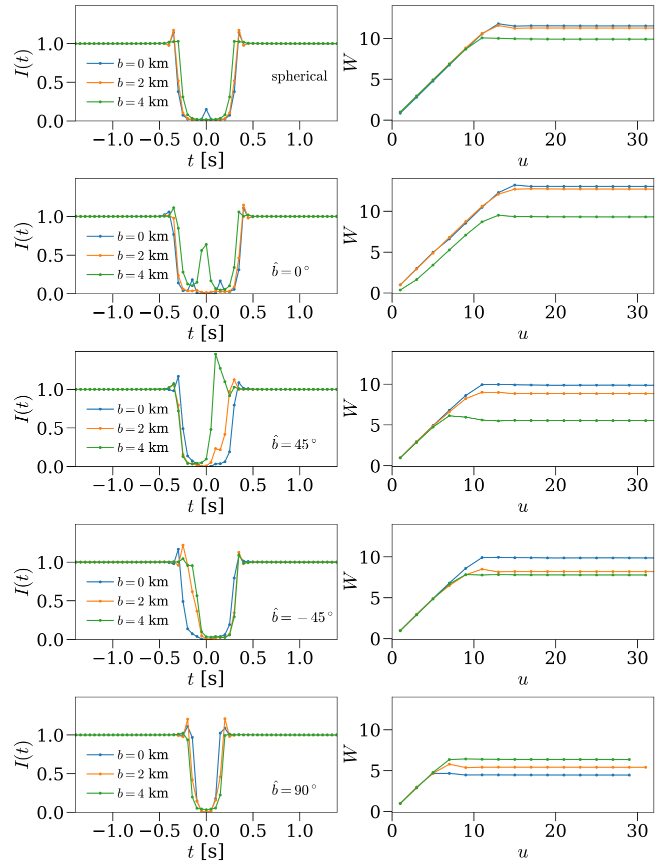

Figure 14. Example light curves (SNR $= 10^6$) and EWCs for $D_{\mathrm{eff}} = 15$ km spherical and compact binary TNO corresponding to the events shown in Figures 12(b) and (d). The top panel on the left shows light curves for a spherical TNO, while four lower panels on the left show light curves for different values of the reading direction $\hat{b}$. In each plot, light curves are shown with three different values of the impact parameter b. The chords across the occultation shadow are indicated in Figures 12(b) and (d) using lines of the same colors. The panels on the right show the EWCs for the light curves shown in the left panels, using lines of the same colors.

regime) spherical TNOs. Figure 12(c) shows the case of a $D_{\mathrm{eff}} = 3$ km TNO in the mid-field regime, and the shadow for a $D_{\mathrm{eff}} = 15$ km TNO is shown in Figure 12(d). For the mid-field compact binary event, the pattern is very complicated, while in the near-field regime, the patterns are very similar in shape and size to those of the TNOs for both the non-spherical and compact binary cases.

In the case of the $D_{\mathrm{eff}} = 3$ km compact binary event, for each of the chords across the occultation shadows shown in Figures 12(a) and (c), the corresponding light curve and EWC are shown in Figure 13. Despite the complicated nature of the occultation shadow, the finite exposure time smooths out most of the oscillations, and the resulting light curves are not very complex. The same plots for the $D_{\mathrm{eff}} = 15$ km compact binary are shown in Figure 14. With the exception of a couple points where the exposure is centered on a crossing of the bright fringe, the light curves are very close to what one would expect when modeling the shadow geometrically with the size and shape of the TNO.

Scatter plots of w_{eq} versus d_{eq} are shown in Figures 15 and 16 from the EWCs for the $D_{\mathrm{eff}} = 3$ km and $D_{\mathrm{eff}} = 15$ km light curves, respectively, which are shown in Figures 13 and 14. In both figures, with the exception of a couple outliers, the points follow the trend of increasing b for spherical TNOs with the same D_{eff}. Increasing the impact parameter tends to reduce the values of w_{eq}, as well as the values of d_{eq} for the smaller object, which subsequently provide better matches to smaller and closer spherical TNOs with $b = 0$. This is expected, since larger values of b will typically lead to smaller portions of the occultation shadow being sampled. However, in the case of the spherical 3 km object, note that in Figure 15 both w_{eq} and d_{eq} increase as b increases from 0, and then drop down to lower

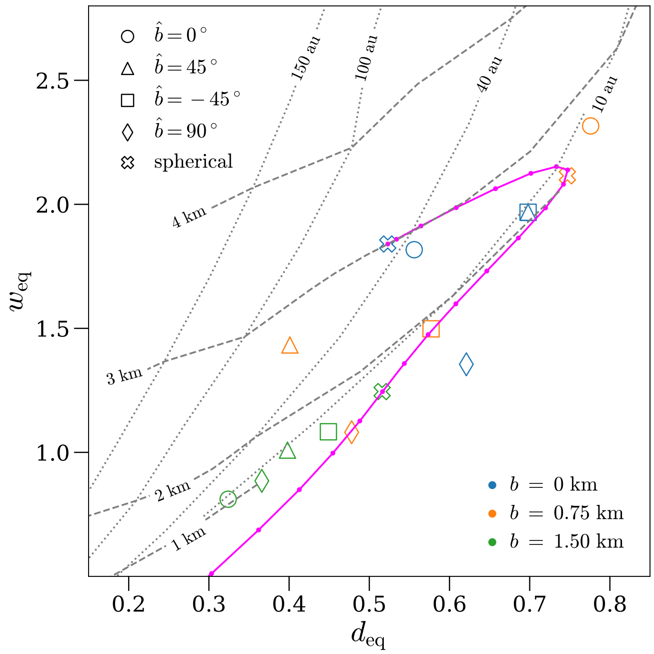

Figure 15. Scatter plot of w_{eq} vs. d_{eq} from the light curves shown in Figure 13. Colors of the points indicate the impact parameter b, while the shapes of the points indicate the either the value of $\hat{b}$ or that it is for a spherical TNO. The dashed gray lines show contours of constant D with varying Δ, assuming spherical TNOs and $b = 0$. The dotted gray lines show the same, but for constant Δ and varying D. The magenta line and points indicate the contour of varying b (in steps of 100 m) for a spherical TNO with $D = 3$ km and $\Delta = 45$ au.

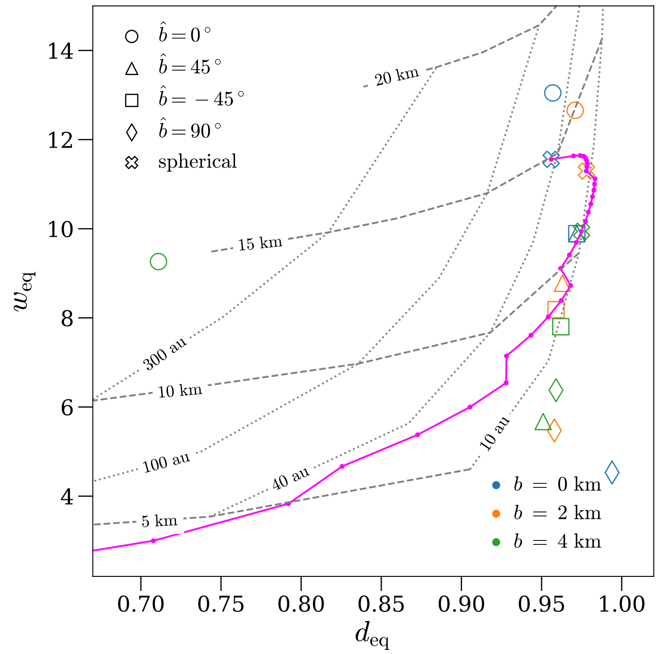

Figure 16. Scatter plot of w_{eq} vs. d_{eq} from the light curves shown in Figure 14. Colors of the points indicate the impact parameter b, while the shapes of the points indicate the either the value of $\hat{b}$ or that it is for a spherical TNO. The dashed gray lines show contours of constant D with varying Δ, assuming spherical TNOs and $b = 0$. The dotted gray lines show the same, but for constant Δ and varying D. The magenta line and points indicate the contour of varying b (in steps of 250 m) for a spherical TNO with $D = 15$ km and $\Delta = 45$ au.

values as b increases further. This is due to the Poisson spot (see Figure 12(a)) at the center of the diffraction pattern. The event gets deeper as the chord moves off of the spot, increasing both w_{eq} and d_{eq}. However, at larger b the length of the chord gets smaller, leading to the eventual decrease in w_{eq} and d_{eq}. It can be seen in Figure 15 that the parameters increase in some cases for the compact binary as well.

5. Conclusion

We have developed an event detection algorithm to search for simultaneous variability due to occultations by TNOs in light curves collected on a target star synchronously with three telescopes. We have also developed an algorithm to estimate the physical parameters of the occulting TNO which uses information from the event detection to bootstrap the process. For the analysis of survey data, the event trigger will need to be run on every three-telescope light curve set collected throughout a night of observations before the next night of observing begins, and the event parameter estimation algorithm will be run on any possible detections (only a few at most, if any, on a single night) to help determine whether or not we should attempt any follow-up observations. For our detection efficiency simulations, simulated events will be added to real data and run through the event detection process. In this case a large number of simulated events will pass our cuts, and we will then run the parameter estimation algorithm on all of these to see if they match the values used to simulate the events. Both algorithms will need to run very quickly for both cases.

We have performed timing tests on both the trigger and parameter estimation algorithms using a single core of an Intel Xeon E5-2660 2.6 GHz CPU. Our tests indicate that running the trigger algorithm on a typical full night of observations (50,000 three-telescope light curve sets, each with 144,000 photometric measurements per telescope) will take about 100 hr on a single core. Given that we have 192 cores on more powerful CPUs installed at SPM, we can search for events in an entire night of data in ∼45 minutes. If each light curve set had an event, the event parameter estimation would take 36 hr on a single core, or just over 10 minutes using all 192 cores. This will provide ample time for preliminary detection efficiency analysis every day in addition to the original trigger search. Finally, we note that the trigger and parameter estimation processes are easily scalable to multiple cores since each core can analyze a different light curve set in parallel.

The event parameter estimation algorithm, while a bit complicated, offers four distinct advantages over the standard maximum likelihood methods. First, it is *fast*. A χ^2 minimization process would entail the generation of a large number of simulated events as the four free parameters (D, Δ, b and $t_{\rm off}$) are modified. It takes a significant amount of time to generate these events, driven mostly by the time it takes to integrate the point source diffraction pattern over the finite angular size of the target star (Nihei et al. 2007; Castro-Chacón et al. 2019). The accuracy that could be achieved by such a process in the parameter estimation is not necessary when measuring our detection efficiency, and given that we will repeatedly add a simulated event to each light curve set collected over the entire survey hundreds of times in order to accurately cover the entire parameter space, speed is a much higher priority than accuracy.

Second, the algorithm is *robust*. Given the degeneracy between D, Δ, and b (assuming spherical TNOs), as well as the complicated dependence of the light curve shapes on these parameters, it is likely that there will be local minima in the four dimensional χ^2 space, complicating any attempt to have an automated high-speed analysis pipeline. While we will do our best to fit any real events detected by the survey, the number of such detections will be small enough that this can be done on an ad hoc basis. On the other hand, the parameter estimation algorithm presented here only involved calculating two simple values involving four sums (see Equation (22)) on a simple data set. The only fitting done is to find the division point between the two parts of the fit to the EWC, and this only involves calculating a χ^2 over a small range of the single, discrete parameter v. We have set a limit on the maximum value of v, but in all of our testing we have never had a case where an event that passed our detection criteria and reached the maximum value without finding a minimum χ^2. Once the best v is found, the algorithm will always return values of $w_{\rm eq}$ and $d_{\rm eq}$, no matter how badly behaved the light curve actually is.

Third, the estimates of $w_{\rm eq}$ and $d_{\rm eq}$ are *accurate*. The error boxes in around the data points in Figures 11(a) and (b) show that the parameters can be calculated accurately for light curves with SNR as low as 5. Furthermore, in the process of developing an algorithm for parameter estimation, we made many different attempts (which we do not document here) to perform simultaneous two-parameter fits of simple shapes to the raw light curves. Each of these attempts failed due to the significant covariance between the different shape parameters we were attempting to calculate. The parameters $w_{\rm eq}$ and $d_{\rm eq}$ are calculated using two completely different sets of data points, so there is no covariance at all between them. In addition, the parameter $w_{\rm eq}$ is completely independent of the shape of the event, and while the part of the EWC that is used to find $d_{\rm eq}$ is shape dependent, the simple straight line model we use works well enough.

Finally, this method works well for most non-spherical events, with the results being nearly as accurate as what we get for spherical events, given the degeneracy when including finite values of b. Also, even ignoring the additional free parameter of the reading direction, the large range of possible TNO shapes would make it very difficult to do much better than the method outlined here, especially in the near-field case where the light curves are mostly measuring chords across approximately geometric shadows. Any number of TNOs with different shapes and sizes could give rise to the light curves shown in Figure 14 if the light curve covers the right chord across the object.

Regarding the degeneracy between D, Δ, and b, this is not an issue for our efficiency simulations. We will know all three parameters for any simulated event we add to our light curves, so we can easily compare the measured values of $w_{\rm eq}$ and $d_{\rm eq}$ with the expected values. However, for purposes of triggering rapid follow-up observations after real time detections of events in the survey data, we can easily set a range of possible values which would trigger such observations. Note that any object that would be a target for follow-up observations will be near enough and large enough that the event will be in the near-field case, where the event duration approximately corresponds to the angular size of a chord across the object, meaning that the object is at least that big in one direction. Unless the object is somewhat ellipsoidal and the chord happens to correspond roughly with the semimajor axis, the TNO will be larger and/or closer to Earth (and thus easier to detect via direct observations) than an object giving rise to the same values of $w_{\rm eq}$ and $d_{\rm eq}$ with $b = 0$. In any case, for any events detected in the survey, every effort will be made to find the most accurate estimates as possible for D and Δ. However, this work is beyond the scope of this paper.

6. Future Work

A database will be created to store the physical event parameters D, Δ, and b, along with the resulting shape parameters $w_{\rm eq}$ and $d_{\rm eq}$, for large ranges of the physical parameters. The database will be indexed for searches on $w_{\rm eq}$ and $d_{\rm eq}$, so a list of combinations of the three physical event parameters which could result in the measured shape parameters can be used for further analysis of any detected events. Given that moving away from opposition leads to larger values of $w_{\rm eq}$, separate tables will be created for different values of the opposition angle ϕ, covering the range over which we will observe. Since the opposition angle is known at the time of observation, we will know which table to use for the search. Separate tables will also be generated covering a range of stellar types and angular sizes since these also give rise to different values of $w_{\rm eq}$ and $d_{\rm eq}$. We will obtain spectra of any target stars where an occultation event is observed in order to calculate the angular size, and we have made reasonably good estimates of the sizes of all of our target stars which will be sufficient for our detection efficiency simulations (J. Karr et al. 2021, in preparation).

For the full three-telescope light curve set, the telescopes are widely separated enough that on occasion there will be significant differences in the light curves measured at the different telescopes during an occultation event. We will conduct an investigation into whether or not this can give rise to different estimates of w_{eq} and/or d_{eq} for the different telescopes, and how such differences could help improve the characterization of any events we may find.

A maximum likelihood analysis pipeline will be developed for the survey, but this will only work under the assumption of spherical TNOs (although a small set of other simple shapes may be considered in the future). However, we expect that the degeneracy between D, Δ, and b will manifest itself in showing significant covariance in the error contours of D versus Δ, similar to the way in which increasing the impact parameter b will lead to lower estimates of D and Δ from the parameters of w_{eq} and d_{eq}.

Finally, it is possible that if enough occultation events are discovered, then the distribution of these parameter estimates as a function of observing parameters can be used to exclude the null hypothesis that all of the detected events are false positives, thus strengthening the case that the events are actual TNO occultations. Of particular interest is the opposition angle ϕ, which along with Δ determines the relative velocity between the observer and TNO and thus the event duration (Nihei et al. 2007, Equation (8)). It is expected that w_{eq} will be larger as ϕ increases (see Figures 11(a) and (b)), while the distribution of w_{eq} should be independent of ϕ for the false positive events. This work will involve a large simulation of the survey itself combined with a simulation of the Trans-Neptunian region.

This work was supported in part by Academia Sinica and the Universidad Nacional Autónoma de México. J.H.C.C.H. and M.R.R. acknowledge support from DGAPA-PAPIIT projects IN107316/IN110217, Cátedras-CONACYT project 219 and from CONACYT through CB project 283800.

ORCID iDs

Chung-Kai Huang https://orcid.org/0000-0003-0728-1493
Matthew J. Lehner https://orcid.org/0000-0003-4077-0985
Agueda Paula Granados Contreras https://orcid.org/0000-0001-8214-5147
Joel H. Castro-Chacón https://orcid.org/0000-0002-2583-076X
Wen-Ping Chen https://orcid.org/0000-0003-0262-272X
Charles Alcock https://orcid.org/0000-0002-7892-3636
Carlos Alberto Guerrero Peña https://orcid.org/0000-0002-1113-7489
Jennifer Karr https://orcid.org/0000-0003-3095-4772
J. J. Kavelaars https://orcid.org/0000-0001-7032-5255
Andrew Szentgyorgyi https://orcid.org/0000-0002-0255-2525
Shiang-Yu Wang https://orcid.org/0000-0001-6491-1901
Wei-Ling Yen https://orcid.org/0000-0002-1417-6264
Zhi-Wei Zhang https://orcid.org/0000-0001-7805-9334

References

Arimatsu, K., Ohsawa, R., Hashimoto, G. L., et al. 2019a, AJ, 158, 236
Arimatsu, K., Tsumura, K., Usui, F., et al. 2019b, NatAs, 3, 301
Bailey, M. E. 1976, Nature, 259, 290
Benavidez, P. G., & Campo Bagatin, A. 2009, P&SS, 57, 201
Benz, W., & Asphaug, E. 1999, Icarus, 142, 5
Bernstein, G. M., Trilling, D. E., Allen, R. L., et al. 2004, AJ, 128, 1364
Bianco, F. B., Protopapas, P., McLeod, B. A., et al. 2009, AJ, 138, 568
Bianco, F. B., Zhang, Z. W., Lehner, M. J., et al. 2010, AJ, 139, 1499
Bickerton, S. J., Kavelaars, J. J., & Welch, D. L. 2008, AJ, 135, 1039
Bickerton, S. J., Welch, D. L., & Kavelaars, J. J. 2009, AJ, 137, 4270
Brown, M. J. I., & Webster, R. L. 1997, MNRAS, 289, 783
Castro-Chacón, J. H., Reyes-Ruiz, M., Lehner, M. J., et al. 2019, PASP, 131, 064401
Chang, H.-K., King, S.-K., Liang, J.-S., et al. 2006, Natur, 442, 660
Chang, H.-K., Liang, J.-S., Liu, C.-Y., & King, S.-K. 2007, MNRAS, 378, 1287
Cooray, A. 2003, ApJL, 589, L97
Cooray, A., & Farmer, A. J. 2003, ApJL, 587, L125
Davis, D. R., & Farinella, P. 1997, Icarus, 125, 50
Duncan, M. J., & Levison, H. F. 1997, Science, 276, 1670
Duncan, M. J., Levison, H. F., & Budd, S. M. 1995, AJ, 110, 3073
Fisher, R. 1925, Statistical Methods For Research Workers (Edinburgh: Cosmo Publications)
Fraser, W. C., & Kavelaars, J. 2008, Icarus, 198, 452
Fraser, W. C., & Kavelaars, J. J. 2009, AJ, 137, 72
Fraser, W. C., Kavelaars, J. J., Holman, M. J., et al. 2008, Icarus, 195, 827
Fuentes, C. I., George, M. R., & Holman, M. J. 2009, ApJ, 696, 91
Fuentes, C. I., & Holman, M. J. 2008, AJ, 136, 83
Holman, M. J., & Wisdom, J. 1993, AJ, 105, 1987
Kenyon, S. J., & Bromley, B. C. 2001, AJ, 121, 538
Kenyon, S. J., & Bromley, B. C. 2004, AJ, 128, 1916
Kenyon, S. J., & Bromley, B. C. 2009, ApJL, 690, L140
Kenyon, S. J., & Luu, J. X. 1999a, ApJ, 526, 465
Kenyon, S. J., & Luu, J. X. 1999b, AJ, 118, 1101
Lehner, M. J., Coehlo, N. K., Zhang, Z. W., et al. 2010, PASP, 122, 959
Lehner, M. J., Wang, S.-Y., Alcock, C. A., et al. 2014, Proc. SPIE, 9145, 914513
Levison, H. F., & Duncan, M. J. 1997, Icarus, 127, 13
Liu, C.-Y., Chang, H.-K., Liang, J.-S., & King, S.-K. 2008, MNRAS, 388, L44
Luu, J. X., & Jewitt, D. C. 2002, ARA&A, 40, 63
Morbidelli, A. 1997, Icarus, 127, 1
Nihei, T. C., Lehner, M. J., Bianco, F. B., et al. 2007, AJ, 134, 1596
Pan, M., & Sari, R. 2005, Icarus, 173, 342
Roques, F., Doressoundiram, A., Dhillon, V., et al. 2006, AJ, 132, 819
Roques, F., & Moncuquet, M. 2000, Icarus, 147, 530
Roques, F., Moncuquet, M., Lavillonière, N., et al. 2003, ApJL, 594, L63
Roques, F., Moncuquet, M., & Sicardy, B. 1987, AJ, 93, 1549
Schlichting, H. E., Ofek, E. O., Sari, R., et al. 2012, ApJ, 761, 150
Schlichting, H. E., Ofek, E. O., Wenz, M., et al. 2009, Natur, 462, 895
Sicardy, B., Roques, F., & Brahic, A. 1991, Icarus, 89, 220
Stern, A., & Campins, H. 1996, Natur, 382, 507
Tancredi, G., Fernández, J. A., Rickman, H., & Licandro, J. 2006, Icarus, 182, 527
Volk, K., & Malhotra, R. 2008, ApJ, 687, 714
Wang, J. H., Lehner, M. J., Zhang, Z. W., et al. 2009, AJ, 138, 1893
Wang, J. H., Protopapas, P., Chen, W. P., et al. 2010, AJ, 139, 2003
Zhang, Z. W., Bianco, F. B., Lehner, M. J., et al. 2008, ApJL, 685, L157
Zhang, Z. W., Lehner, M. J., Wang, J. H., et al. 2013, AJ, 146, 14