

The TAOS Project: Statistical Analysis of Multi-Telescope Time Series Data

M. J. LEHNER,^{1,2,3} N. K. COEHLO,⁴ Z.-W. ZHANG,^{1,5} F. B. BIANCO,^{2,3,6,7} J.-H. WANG,^{1,5} J. A. RICE,⁴ P. PROTOPAPAS,^{3,8}
 C. ALCOCK,³ T. AXELROD,⁹ Y.-I. BYUN,¹⁰ W. P. CHEN,⁵ K. H. COOK,¹¹ I. DE PATER,¹² D.-W. KIM,^{3,8,10}
 S.-K. KING,¹ T. LEE,¹ S. L. MARSHALL,^{11,13} M. E. SCHWAMB,¹⁴ S.-Y. WANG,¹ AND C.-Y. WEN¹

Received 2010 February 18; accepted 2010 June 2; published 2010 June 30

ABSTRACT. The Taiwanese-American Occultation Survey (TAOS) monitors fields of up to ~ 1000 stars at 5 Hz simultaneously with four small telescopes to detect occultation events from small (~ 1 km) Kuiper Belt Objects (KBOs). The survey presents a number of challenges, in particular the fact that the occultation events we are searching for are extremely rare and are typically manifested as slight flux drops for only one or two consecutive time series measurements. We have developed a statistical analysis technique to search the multi-telescope data set for simultaneous flux drops which provides a robust false-positive rejection and calculation of event significance. In this article, we describe in detail this statistical technique and its application to the TAOS data set.

1. INTRODUCTION

The Taiwanese-American Occultation Survey operates four small telescopes (Bianco et al. 2010; Wang et al. 2009; Zhang et al. 2008; Lehner et al. 2009) at Lulin Observatory in central Taiwan to search for occultations by small (~ 1 km diameter) KBOs (Schlichting et al. 2009; Wang et al. 2010; Bianco et al. 2009; Bickerton et al. 2009, 2008; Nihei et al. 2007; Chang et al. 2007; Roques et al. 2006; Cooray 2003; Cooray & Farmer 2003; Roques et al. 2003). Occultation surveys are the only method available to detect these objects, as objects smaller than about 20 km in diameter have magnitudes $R > 30$, which is

beyond the limit of direct observation. Occultation events are extremely rare (estimated rates range from 10^{-4} to 10^{-2} events $\text{star}^{-1} \text{yr}^{-1}$), they are very short in duration ($\lesssim 200$ ms), and at the 5 Hz observing cadence used by TAOS, they result in measured flux drops of typically $\lesssim 30\%$ in one or two consecutive points. This presents a number of challenges, in particular the identification of false-positive events of statistical origin and candidate events which are in fact of terrestrial origin (e.g., birds, airplanes, and extreme scintillation events). We reject these false-positive events by requiring simultaneous detection in multiple telescopes.

A second challenge is finding a robust method to determine the statistical significance of any candidate events. The noise distribution of each light curve is not known a priori, due to non-Poisson and non-Gaussian processes on the tails of the flux distributions. The typical stars in our fields have magnitudes $R \sim 13$ and a signal-to-noise ratio (S/N) of ~ 10 . Moreover, TAOS monitors fields for durations of up to 1.5 hr, and changes in atmospheric transparency and air mass introduce further uncertainties into the flux measurements.

To overcome these difficulties, we have developed nonparametric techniques using *rank statistics*. Rank statistics facilitate a simultaneous analysis of multi-telescope photometric measurements to enable a robust determination of event significance and false-positive rejection, which are independent of the underlying noise distributions of the light curves being analyzed. Occultation events and the application of rank statistics to detect such events in the TAOS data are described in the following sections. In § 2 we review the characteristics of occultation events and describe how such events would appear in the data. In § 3 we discuss the rank product statistical test used to calculate event significance and the false-positive rate. In § 4 we describe the light-curve filtering techniques and diagnostic tests used to ensure that the rank product statistical test is valid,

¹ Institute of Astronomy and Astrophysics, Academia Sinica. P.O. Box 23-141, Taipei 106, Taiwan; mlehner@asiaa.sinica.edu.tw.

² Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, PA 19104.

³ Harvard-Smithsonian Center for Astrophysics, Cambridge, MA 02138.

⁴ Department of Statistics, University of California Berkeley, Berkeley, CA 94720.

⁵ Institute of Astronomy, National Central University, No. 300, Zhongda Rd, Jhongli City, Taoyuan County 320, Taiwan.

⁶ Department of Physics, University of California, Santa Barbara CA 93106-9530

⁷ Las Cumbres Observatory Global Telescope Network, Inc., Santa Barbara, CA 93117.

⁸ Initiative in Innovative Computing, Harvard University, Cambridge, MA 02138.

⁹ Steward Observatory, Tucson, AZ 85721.

¹⁰ Department of Astronomy, Yonsei University, 134 Shinchon, Seoul 120-749, Korea.

¹¹ Institute of Geophysics and Planetary Physics, Lawrence Livermore National Laboratory, Livermore, CA 94550.

¹² Department of Astronomy, University of California, Berkeley, CA 94720.

¹³ Kavli Institute for Particle Astrophysics and Cosmology, Menlo Park, CA 94025.

¹⁴ Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA 91125.

and in § 4.4 we describe a new and more robust set of diagnostics tests.

The following definitions apply throughout the remainder of this article. We define a *data run* as a consecutive series of multi-telescope observations of a given star field made at a cadence of 5 Hz. Typical data runs last 1.5 hr, comprising 27,000 time series images on each telescope. We define a *light-curve set* as a set of multi-telescope light curves of a single star during a given data run. There are typically 300–500 stars in an image, and hence 300–500 light-curve sets in a data run. We adopt the standard statistical notation wherein we denote a random variable with an upper case letter, and use the corresponding lower case letter for an actual value for that variable (e.g., Z is a random variable which could take on a value of z). We use the function $p()$ to describe a probability density distribution, and $\mathbb{P}()$ to describe an actual probability. Finally, we note that the four telescopes are labeled TAOS A, TAOS B, TAOS C, and TAOS D. TAOS C came online in 2008 August, and to date no data from this telescope have been analyzed. All example light curves shown in this article come from telescopes A, B, and D.

2. OCCULTATIONS BY KUIPER BELT OBJECTS

An occultation event occurs when an object passes between the telescope and a distant star (Bickerton et al. 2009; Nihei et al. 2007; Roques et al. 2003). The Earth and the occulting object are in relative motion, inducing a variation in the measured stellar flux over time. The target population for TAOS is small

(~1 km diameter) KBOs, whose sizes are on the order of the *Fresnel scale*, which is given by

$$F = \sqrt{\frac{\lambda \Delta}{2}},$$

where λ is the wavelength of observation and Δ is the observer–KBO distance. For TAOS, the median wavelength of observation is $\lambda \approx 600$ nm, and the typical distance to KBOs is 43 AU, resulting in $F = 1.4$ km. Occultation events by KBOs with diameters $D \lesssim 10$ km thus show significant diffraction effects. This is illustrated in the left panel of Figure 1, which shows a simulated occultation “shadow” from a 3 km diameter KBO projected onto the surface of the Earth.

The timescale of an occultation event is set by the relative velocity between the KBO and observer, the size of the occultation shadow, and the impact parameter (minimum distance between the KBO and the line of sight to the target star). Assuming a circular orbit, the relative velocity between the Earth and KBO in the plane of the sky is given by

$$v_{\text{rel}} = v_E \left[\cos \phi - \left(\frac{1 \text{ AU}}{\Delta} \right)^{\frac{1}{2}} \left(1 - \frac{1 \text{ AU}^2}{\Delta^2} \sin^2 \phi \right)^{\frac{1}{2}} \right], \quad (1)$$

where ϕ is the angle of observation between the occulted star and opposition, and $v_E = 29.8 \text{ km s}^{-1}$ is the velocity of the Earth around the Sun. The event width (the length of the chord across the occultation shadow where it crosses the telescope) is given by

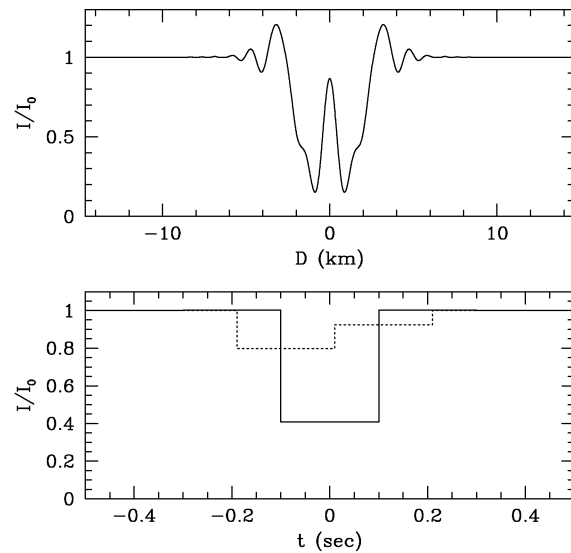
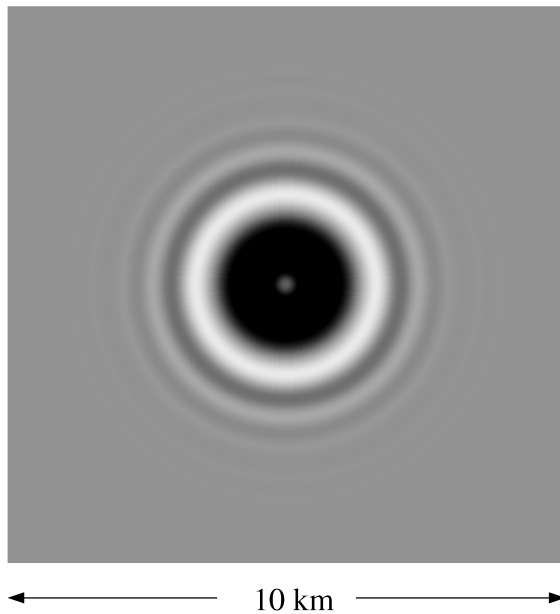


FIG. 1.—*Left panel*: diffraction shadow projected onto the surface of the Earth from a 3 km diameter KBO at 43 AU. *Right panel, top*: perfectly sampled light curve assuming zero impact parameter. *Right panel, bottom*: same light curve as sampled by the TAOS system at 5 Hz. *Solid curve* has measurements centered on event, *dotted line* shows light curve where sampling is out of phase with event.

$$W = \sqrt{H^2 - b^2},$$

where b is the impact parameter, and H is the event cross section, which we define as the diameter of the first Airy ring of the diffraction shadow, and which can be approximated by (Nihei et al. 2007)

$$H \approx [(2\sqrt{3}F)^{\frac{2}{3}} + D_*^{\frac{2}{3}}] + \theta_* \Delta, \quad (2)$$

where θ_* is the angular size of the occulted star.

For the very small objects ($D \lesssim 1$ km) targeted by this survey and stars with small angular diameters (the vast majority of stars covered by this survey), the minimum event cross section is set by the Fresnel scale:

$$H_{\min} \approx 2\sqrt{3}F. \quad (3)$$

At 43 AU, $H_{\min} \approx 5$ km. At opposition ($\phi = 0$), $v_{\text{rel}} \approx 25 \text{ km s}^{-1}$, and with $b = 0$ the resulting event duration is 200 ms, with the duration getting smaller as b is increased.

This is illustrated in the right panels of Figure 1. The top panel shows a slice through the simulated diffraction shadow, assuming the KBO crosses the line of sight to the star ($b = 0$). Note that the event width, given by the distance between the two top peaks, is about 5 km. (In this case, the event width is dominated by the Fresnel scale, so the approximation given in eq. [3] applies.) The bottom panel shows this event as it would be measured by the TAOS system at 5 Hz. The solid line shows the light curve which would be measured if the sampling was in phase with the event, that is, the measurement at $t = 0$ is centered on the epoch when the KBO is centered on the line of sight to the target star. The dotted line shows the same event with the sampling out of phase with the event.

Typical occultation events for small KBOs at opposition will thus manifest themselves in the TAOS data as a reduction in flux on one or two consecutive photometric measurements of a star with an otherwise flat light curve. However, when observing away from opposition, the relative velocity decreases, as indicated by equation (1). Furthermore, TAOS is also sensitive to objects more distant than the Kuiper Belt. The discovery of Sedna (Brown et al. 2004) indicates the possibility of a large, heretofore unknown population of objects at distances of 100 to 1000 AU (see Wang et al. 2009 and references therein). Such events will also be of a longer duration due to the increased angular size of the Fresnel scale, as indicated by equation (2). Figure 2 shows a simulated light curve with an occultation by a 5 km object at 500 AU, observed at 70° from opposition. The width of the event is about 22 km, and with a relative velocity of about 9 km s^{-1} , the event duration is about 2.5 s, corresponding to a total of 13 measurements at our cadence of 5 Hz. (Once again, the approximation for the event width given in eq. [3] applies.)

The goals of the TAOS statistical analysis described in this article are to find as many such events as possible, minimize the

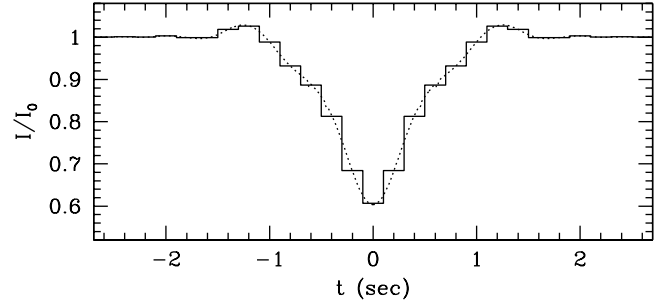


FIG. 2.—Simulated light curve with an occultation by a 5 km object at 500 AU, measured at 70° from opposition. Diffraction features are smoothed out due to finite angular size of the occulted star. *Dotted line* is the infinitesimally sampled light curve, and the *solid line* indicates the light curve as would be measured with 5 Hz sampling. See Nihei et al. (2007) for a discussion of occultation events from objects at such distances.

false-positive rate, and provide a method to robustly estimate the statistical significance of any candidate event. The statistical technique should be sensitive to both the one- and two-point events shown in Figure 1 and the longer duration events such as that shown in Figure 2. In the following sections the application of rank statistics to meet these goals will be described. The discussion will begin with a focus on single-point events, and the extension of the statistical technique to multipoint observations will be presented in § 3.4.

3. RANK STATISTICS

The idea of rank statistics is quite simple, and is best introduced with a single series. Take a time series of flux measurements f_1, \dots, f_{N_p} from one telescope, where N_p is the number of points in the time series. Replace each flux measurement f_j with its rank r_j . That is, the lowest f_j will be assigned rank 1 and the highest assigned rank N_p . If we use a total of T telescopes, we replace the time series for each telescope with its rank within the light curve from that telescope, giving a set of T rank time series r_{ij} . Thus for each time point t_j , we have a *rank tuple*

$$(r_{1j}, \dots, r_{Tj}).$$

If, for each telescope i , the rank r_{ij} follows a uniform distribution on $\{1, \dots, N_p\}$ at each time point t_j , and if the light curves f_{ij} are independent between the different telescopes, then each rank tuple combination is equally likely at each time point, and we can calculate exact probability distribution of these rank tuples. The calculation of the probability distribution of the raw data is impossible to perform on the original time series measurements, since the underlying distributions of the flux measurements f_{ij} in each light curve are unknown. That is, by working with the ranks, we replace something unknown, the distribution of the data, with something known, the distribution of the ranks.

A time series is *stationary* if the distribution of any finite subset of the series is invariant under time shift. A stationary time series f_j is *ergodic in mean* if, for any function G with an expected value $\mathbb{E}(G(f)) < \infty$, we have the following convergence with probability of 1 (law of large numbers):

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n G(f_i) \rightarrow \mathbb{E}(G(f)).$$

It can be shown that if a time series f_j of length N is ergodic in mean, then the distribution of ranks r_j/N will converge to the uniform distribution for any sequence $1 \leq j \leq N$, and it is well known that ergodicity in mean can be assured under very weak assumptions on temporal dependence within a light curve. This proof is beyond the scope of this article, but it is published in Coehlo (2010).

Therefore, if the data f_j are stationary and ergodic in mean, this implies that at each time point t_j , r_j will be uniform on $\{1, \dots, N_p\}$. In addition, if the light curves from different telescopes are independent, then the rank tuples (r_{1j}, \dots, r_{Tj}) will be uniform on

$$\{1, 2, \dots, N_p\}^T,$$

and we can calculate exact probability distributions of the rank tuples. However, most of the light-curve sets exhibit slowly varying trends that are highly correlated between the different telescopes, so our light curves are neither uncorrelated nor stationary. We have developed a filtering algorithm to remove these trends, and in most cases the resulting individual light curves can be plausibly modeled as stationary and ergodic in mean. In most cases the correlations between the light curves from different telescopes are also removed by the application of the filter. This filtering algorithm will be described in § 4.

Figure 3 introduces the *rank-rank diagram*, which is a scatter-plot of the ranks on two telescopes. Similar plots will be used throughout the remainder of the article to illustrate various statistical tests. Note that each rank must occur once and only once in each time series. Thus there must be exactly one point in each row and in each column. The ranks within a single light curve are thus not independently distributed. However, if the conditions on ergodicity in mean and no dependence between telescopes are met, then the rank pairs will be uniformly distributed throughout the diagram.

3.1. The Rank Product Test Statistic

As can be seen in Figure 1, events consistent with occultations by KBOs will appear as one or two consecutive flux drops in all four telescopes. Our test is thus designed to find those rank tuples where all of the ranks are small, corresponding to a region toward the lower left corner of the rank-rank diagram shown in Figure 3 (expanded to T dimensions, where T is the number of

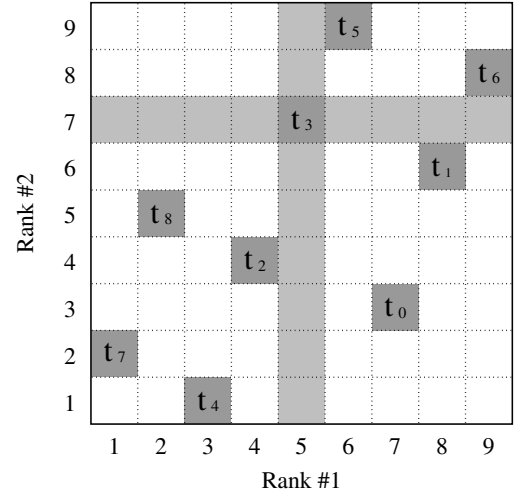


FIG. 3.—Schematic of a rank-rank diagram, with $N_p = 9$. Axes are ranks of photometric intensity for individual data points on two different telescopes. A single two-telescope photometric measurement will correspond to a rank doublet on this plot. These are marked with the *dark squares* and labeled with the time at which they were measured. For example, note the highlighted rank pair at (5,7), measured at time t_3 . Note that each rank value occurs once and only once in the time series for each telescope.

telescopes). The assumptions on the rank statistics and conditions placed on the raw data allow us to calculate the significance level α of various test statistics corresponding to this region.

The statistical analysis is designed to use each rank tuple (r_{1j}, \dots, r_{Tj}) to perform a hypothesis test that there is an event at time t_j . Each measurement r_{ij} can be used as a test statistic for the null hypothesis of no occultation event versus the alternative that there is an occultation, yielding a p -value given by

$$\mathbb{P}(R \leq r_{ij}) = \frac{r_{ij}}{N_p}.$$

The goal is to use the tuple of T p -values at time t_j to calculate a single test of significance. Fisher proposed that the product of the p -values be used as a test statistic for this general problem (Fisher 1958; Mosteller & Fisher 1948). Given the product of ranks at time t_j over all telescopes T

$$y_j = \prod_{i=1}^T r_{ij},$$

we define our rank product statistic as

$$z_j = -\ln\left(\frac{y_j}{N_p^T}\right). \quad (4)$$

Event detection based on the rank product statistic was described in Lehner et al. (2006) and Zhang et al. (2008). In the

description presented in Lehner et al. (2006), we made the assumption that the distribution of p -values r_{ij}/N_p was uniform on the *continuous* interval $[0, 1]$. In this case, it can be shown that the rank product distribution has a distribution of the form

$$p(z_j) = \frac{1}{\Gamma(T)} z_j^{T-1} e^{-z_j}, \quad (5)$$

which is simply the Γ distribution. However, the distribution of ranks is in fact uniform on the *discrete* set $\{1/N_p, 2/N_p, \dots, 1\}$, and we found that the assumption of a continuous distribution leads to substantial errors for large values of z . The true distribution of the rank product can be calculated using the function $K(n; T, N_p)$, which we define as the number of ways to get a product of n by multiplying T integers (number of telescopes) between 1 and N_p (number of points in the light curves). This function can be calculated numerically, and we have developed a simple algorithm to calculate $K(n; T, N_p)$ when $n \leq N_p$ (see Appendix). Some values of this function for $T = 4$ telescopes are shown in Table 1. Note that this function is independent of N_p if $n \leq N_p$.

Rather than using the function K to calculate the probability density of the rank product statistic z , it is simpler to calculate the distribution as a function of the rank product y as

$$p(y) = \frac{1}{N_p^T} K(y; T, N_p).$$

We thus calculate the significance, or p -value, of any candidate event as

$$\mathbb{P}(Y \leq y) = \frac{1}{N_p^T} \sum_{i=1}^y K(i; T, N_p). \quad (6)$$

However, for clarity we continue to display results in terms of the rank product statistic z because candidate events are more easily distinguished on the tail of the distribution (see Fig. 4). Given the relation between z and y , it clearly follows that

$$\mathbb{P}(Z \geq z(y)) = P(Y \leq y).$$

Note that the results published in Zhang et al. (2008) and Bianco et al. (2010) use the correct probability distribution based on the discrete rank distribution.

The efficacy of the rank product method is shown in Figure 4 (right panel). A three-telescope (top) and a four-telescope (bottom) data run were simulated. On the three-telescope run, an event was added with a rank triplet $\{10, 10, 10\}$, and on the four-telescope run, an event was added with ranks $\{10, 10, 10, 10\}$. The four-telescope event has a p -value of 3.7×10^{-12} under the null hypothesis, while the three-telescope event has a p -value of 1.5×10^{-9} . This simple example illustrates the value of using multiple telescopes, in that the absence of the fourth telescope decreases the significance of the event by more than 2000, while keeping the false-positive rate fixed.

The rank product test statistic is based on subsets of the rank tuples where events would plausibly be expected to be found. However, in general, the subset of rank tuples that provides the most sensitive detection is composed of those tuples which are most likely in the event of an occultation. We could imagine identifying this subset by running an enormous simulation of occultations which produced a probability for each of the T^{N_p} tuples. The rejection region for the test would then be composed of the quadruplets with largest probabilities, the number

TABLE 1
RANK QUADRUPLTS USED TO CALCULATE $K(z; T, N_p)$ FOR $T = 4$ AND $z \leq N_p$

$K(1) = 1$	$K(2) = 4$	$K(3) = 4$	$K(4) = 10$	$K(5) = 4$	$K(6) = 16$
1111	1112	1113	1114	1115	1116
	1121	1131	1141	1151	1161
	1211	1311	1411	1511	1611
	2111	3111	4111	5111	6111
			1122		1123
			1212		1132
			2112		1213
			1221		1312
			2121		2113
			2211		3112
					1231
					1321
					2131
					3121
					2311
					3211

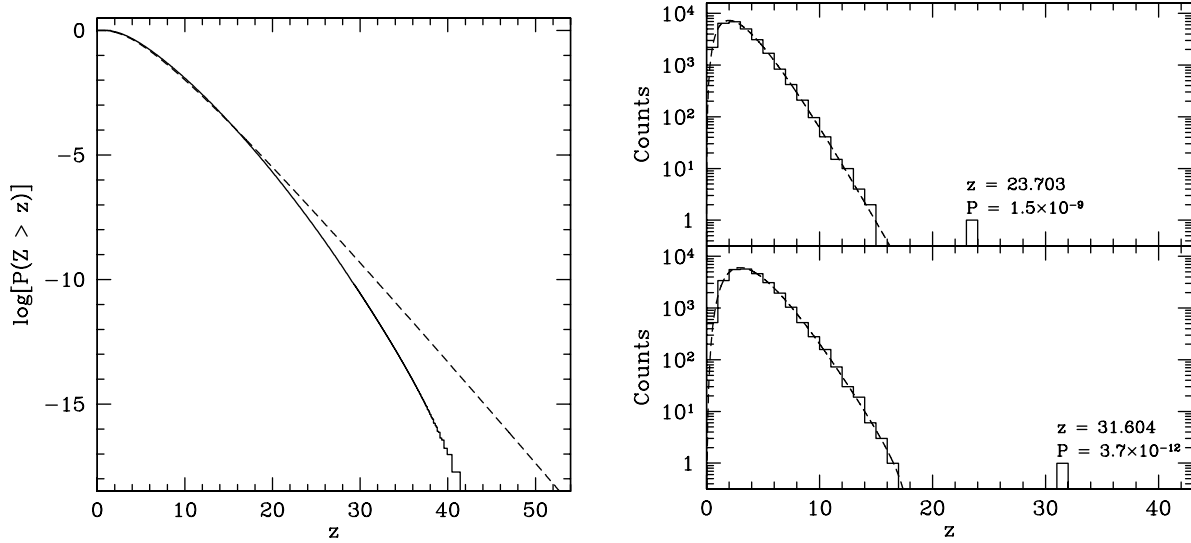


FIG. 4.—*Left panel:* event significance as a function of z (solid line), assuming $N_p = 27,000$ and $T = 4$. Also shown for comparison is the event significance calculated using the Γ distribution approximation (dashed line). *Right panels:* the results of two simulations illustrating the power of the rank product method for event selection. On the *top panel*, the histogram shows the parameter z for $T = 3$ and $N_p = 27,000$, and the dashed line shows the true distribution given the null hypothesis. The rank triplet on the tail has ranks $\{10, 10, 10\}$. The *bottom plot* is the same, but with $T = 4$, and the outlier arises from a rank quadruplet of $\{10, 10, 10, 10\}$.

being determined by the desired false-positive rate. Note that the rejection region might not be symmetric in the telescopes (invariant to the telescope labels), which might be desirable if light curves from some telescopes had much better signal to noise ratios than from others. We have not carried out such a simulation, but a modest simulation indicates that the rejection region determined by the rank product statistic is sufficient for the purpose of event detection.

3.2. False-Positive Rate

The methodology we employ is to search for an event at every time point in every light-curve set that the survey has collected. Hence, the total number of hypotheses tested is

$$N_{\text{hyp}} = \sum_l N_p(l),$$

where the sum is over all light-curve sets l in the data set.

If we set a significance threshold of α to declare an event at time point t_j in light-curve set l , and we use the same significance threshold at all times in all light curves, then the expected number of declared events due to chance would be

$$\alpha \times N_{\text{hyp}} = \alpha \times \sum_l N_p(l).$$

Therefore, to control false positives we must make α very small. For the results published in Zhang et al. (2008), the data set (after diagnostic cuts, see § 4.3 for details) comprised a total

of 2.3×10^9 tuples, and the threshold used was $\alpha = 10^{-10}$, which gives a predicted 0.23 false-positive events. To keep the false-positive rate low for the larger data set (9.0×10^9 tuples) used in Bianco et al. (2010), we used $\alpha = 3 \times 10^{-11}$, corresponding to an expected number of 0.27 false positives.

In all likelihood there will be at most one occultation in a light-curve set, and if an occultation occurred over consecutive time points it would only be counted once. Hence, one could consider performing a hypothesis test over the entire light-curve set rather than at each time point by looking at the minimum of the rank product over all time points in the light-curve set:

$$\beta = \min_j \prod_i r_{ij}.$$

If we test based on β at level α' , then the expected number of false positives is

$$\sum_l \alpha' = \alpha' \times \text{number of light-curve sets}.$$

The distribution of β can be evaluated numerically, but it is much easier to work with the rank product at every time point. Given a constraint on the false-positive rate, and given the lengths of the series of interest, and the part of the distribution we are interested in (the tail), it has been found (Coehlo 2010) that there is little difference if we work with β or with the rank product at all time points; the same events will be detected.

3.3. Power of Rank Product Test

The primary advantage of the rank product test is that one can calculate the exact *level* of the test, that is, the probability of measuring a particular rank product under the null hypothesis (no event is present). However, the replacement of the original data by the ranks leads to a loss of *power* of the test, that is, the probability of detection under the alternative hypothesis (an event *is* present). This is due to the loss of information when mapping the photometric measurements onto the rank space.

It is impossible to know how much power is lost by using the rank product test since the underlying distributions of the photometric measurements are not known. However, we can calculate the loss of power when the data are Gaussian by comparing the rank product with the power of the likelihood ratio test (after co-adding the light curves in a light-curve set), which is optimal under the Gaussian assumption. We have performed such a simulation, using $T = 3$ telescopes and light curves of length $N_p = 27,000$. For the simulation, we generated a total of 10^5 light-curve sets. Each light curve is generated with a mean value μ and a noise level σ . To each light-curve set we added a single-point occultation event of depth

$$\delta = a \times \mu,$$

where a is a free parameter. We then found how many events were recovered by each test, using a threshold level of $\alpha = 10^{-10}$. It can be shown (Coehlo 2010) that the power of each statistical test is a function of

$$c = \frac{\delta}{\sigma} = a \times S/N,$$

and is otherwise independent of μ and σ . We thus repeated the test for several values of c , and the results are shown in Figure 5. The top panel of Figure 5 shows the power of the Gaussian and rank product tests as a function of the event depth c , and the bottom panel shows the ratio of the powers of the rank product and Gaussian tests. The power of both tests is very low for occultation event depths $c < 3$. For values of $c \approx 3$, the power of the rank product test is about 70% of the power of the Gaussian test, and for larger values the ratio rapidly approaches one. Also note that for $c > 5$, nearly every event is detected by both tests. Since many occultations produce values of c outside the $3 < c < 5$, and the power reduction is modest inside $3 < c < 5$, we conclude that we do not lose much power in the Gaussian case.

3.4. Detection of Multipoint Occultation Events

If we had an occultation from a large object in the Kuiper Belt, it would cause a substantial flux drop for several consecutive time points, resulting in several values of the rank product that pass the significance threshold. On the other hand, if the object were at 200 AU, it might cause a modest flux drop

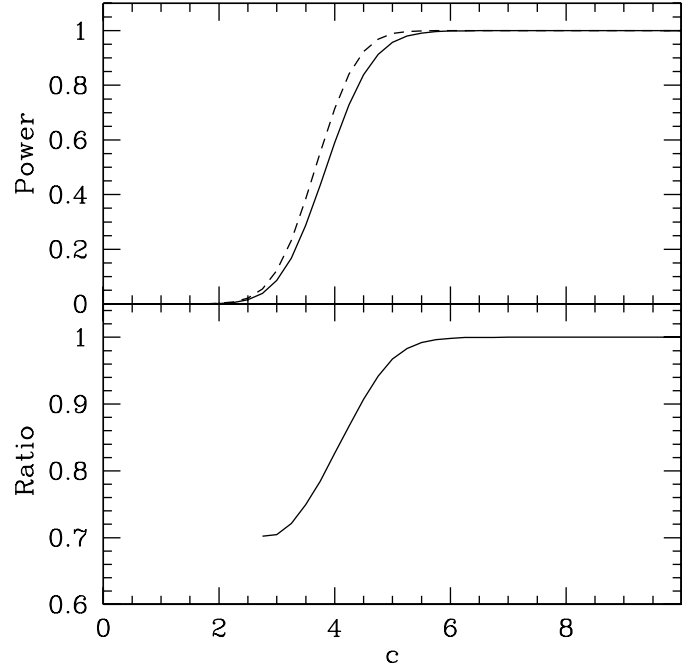


FIG. 5.—*Top panel:* power of the rank product test (solid line) and Gaussian test (dashed line) as a function of occultation depth c . *Bottom panel:* ratio of the power of the rank product test and the power of the Gaussian test as a function of occultation depth c .

for several time points, none of them big enough to pass the threshold. In the latter case, it is useful to consider functions of the data that look at neighboring time points for detection.

Let our original time series be f_1, \dots, f_N , and suppose we form a new series by

$$a_j = b(f_{j-k}, \dots, f_j, \dots, f_{j+k}).$$

For example, b could be a moving average:

$$b(f_{j-k}, \dots, f_j, \dots, f_{j+k}) = \frac{1}{2k+1} (f_{j-k} + \dots + f_{j+k}).$$

The series a_j might show a larger response at the center of the modest signal than f_j , leading to better detection efficiency. Another possibility for b is to take the inner product with some signal. For example, a series of event templates could be used as the function b to search for occultation events from objects of specific sizes and distances, which is what was done by Schlichting et al. (2009), Wang et al. (2010), and Bickerton et al. (2008).

Such manipulation of the data will introduce significant autocorrelation into the light curves. However, if the series f_j is stationary and ergodic in mean, and if k is small relative to the length of the light curves, then it follows that a_j will also be stationary and ergodic in mean, so the rank product distribution will still be satisfied. This is because the autocorrelation

structure is expected to be the same throughout the light curve. See Coehlo (2010) for a detailed discussion.

4. LIGHT CURVE FILTERING

As discussed in § 3, the tests based on rank statistics are valid only if the light curves from each telescope are stationary, ergodic in mean, and independent of those from other telescopes. However, in the actual data, significant correlations and nonstationarity are evident, as can be seen in the top left panel of Figure 6. Trends like those evident in the light curves in Figure 6 can arise due to changing air mass and atmospheric transparency throughout the duration of a run. The bottom left panel of Figure 6 shows the rank-rank diagram corresponding to this light-curve set (telescopes A and B are shown). Under the assumption of independence, the points should be distributed uniformly across the diagram; clearly this is not the case.

To solve this problem, we apply a *mean filter* to the light curves in order to remove the slowly varying trends. Each photometric measurement f_j in a light curve is replaced with

$$g_j = f_j - \bar{f}_j,$$

where \bar{f}_j is defined as a 3σ -clipped (Bertin & Arnouts 1996; Da Costa 1992) mean taken over a window of size W_μ which is centered on the point f_j .

After application of the mean filter, we found many light curves that exhibit fluctuations in variance over time. In periods of higher variance, more extreme high or low rank values are more likely, and our assumption on the uniform distribution of ranks throughout a light curve is invalid. We thus correct for changes in the variance by applying a *variance filter*, where we replace every point g_j with

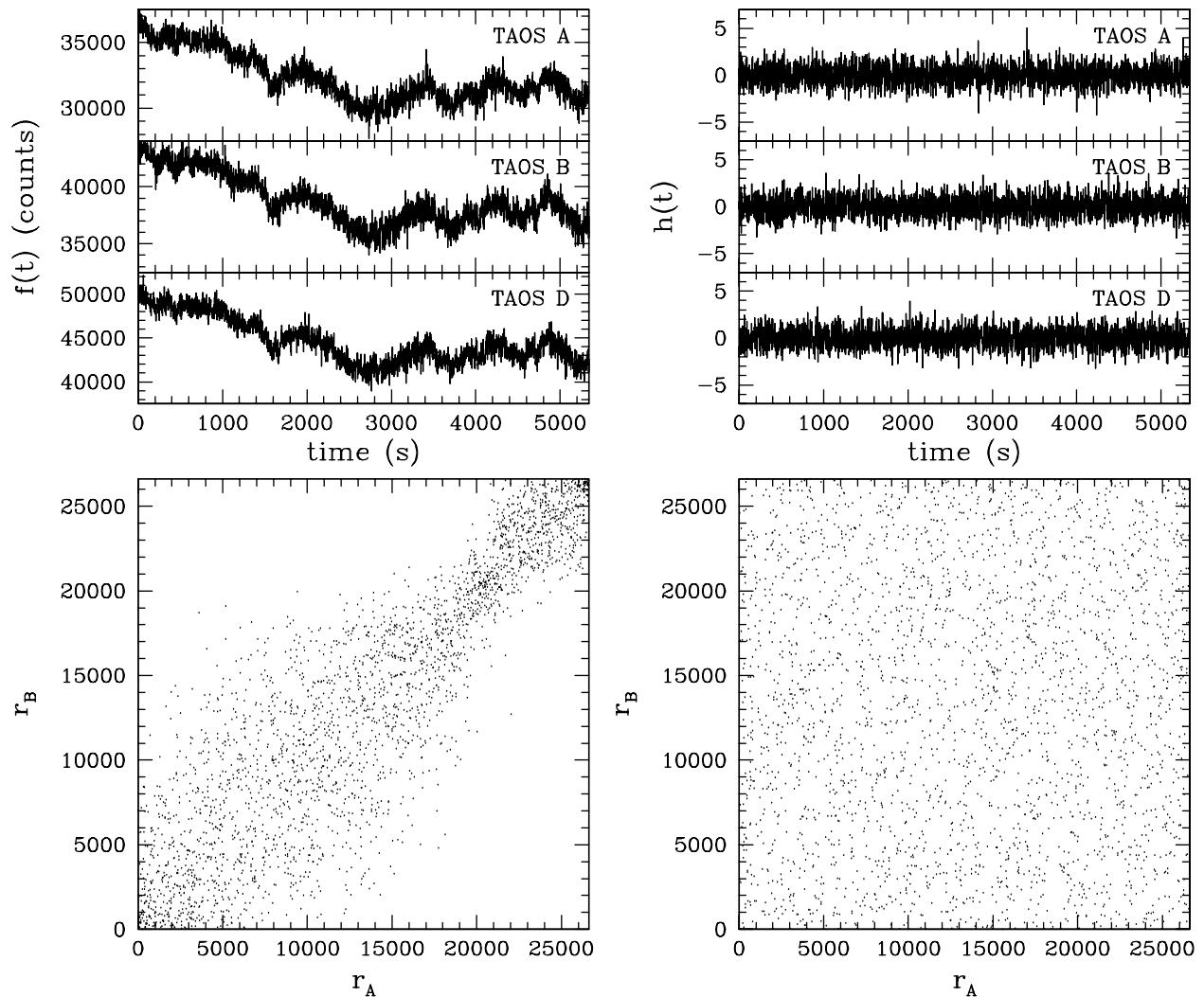


FIG. 6.—*Top left:* an unfiltered three-telescope light-curve set for a single star. Note the correlated variations in the light curves. *Bottom left:* rank-rank diagram for telescopes A and B. *Top right:* same light-curve set after filtering. *Bottom right:* rank-rank diagram after filtering. All four plots are reproduced from Zhang et al. (2008).

$$h_j = \frac{g_j}{\sigma_j},$$

where the standard deviation σ_j is calculated over a window of size W_σ centered on the point g_j , and 3σ clipping is applied here as well.

We want to choose the window sizes to be small enough to accurately correct for high frequency trends, but we also want them large enough to enable accurate determination of \bar{f} and σ . After testing various window sizes, we found that $W_\mu = 33$ and $W_\sigma = 151$ work best (the variance fluctuates much more slowly than the mean, hence the larger window size).

We note that much work has been done in the past on removing such trends from light curves, most of which involves removing correlated trends in light curves from different stars in the same series of images (for example, see Kovács et al. 2005; Tamuz et al. 2005; Bianco et al. 2009). The simpler approach we have adopted works well enough for our purposes, but we are considering adopting similar techniques for future analysis.

The top right panel of Figure 6 shows the same light-curve set shown in the left panel, after filtering. The trends in the mean and variance have clearly been removed. The rank-rank diagram of the filtered light-curve set is shown in the bottom right panel of Figure 6. No dependence is evident in this diagram.

Figure 7 shows autocorrelation functions (ACFs) after application of the mean and variance filters. Three of the panels show ACFs of light curves in the TAOS data after filtering (such as those shown in the top right panel of Fig. 6), and one of them shows the ACF of a synthetic light curve of white noise, after the same filters have been applied. The autocorrelation is insignificant in all cases after a time lag of 6.6 s, which corresponds to

the window size W_μ of the mean filter. The ACF of the simulated light curve demonstrates that the observed features in the ACFs are consequences of the mean filter. The small feature in Figure 7a evident at time lag of 30 s is likely due to the variance filter which has a window size of $W_\sigma = 151$ points. The short timescale (relative to the length of the light curve) of the significant autocorrelation features is consistent with our modeling of the filtered light curves as stationary and ergodic in mean, as dependence over longer timescales would invalidate our assumptions that all possible ranks are equally likely at each time point (Coehlo 2010).

While the filter appears to work well on the example light-curve set shown in Figure 6, we still need to quantify how well it actually works. This is important because some data runs may exhibit variations that are not adequately corrected for by the filters we apply. In particular, data runs with extremely rapid fluctuations in the mean (due to fast-moving cirrus clouds or other phenomena) will not be removed if the event width is small when compared with W_μ . This is illustrated in Figure 8, which shows a light-curve set taken during a night with periods of fast-moving cirrus clouds. Significant correlations are evident in the filtered light curves. The corresponding rank-rank diagram of telescopes TAOS A and TAOS B is shown in Figure 9. Significant overdense regions are evident in the lower left and upper right corners of this diagram. In order for the application of the rank statistics to be valid, such data need to be flagged and cut from the data set before the application of the rank product test.

We have thus developed two diagnostic tests to be applied to each light-curve set to assess the quality of the data after the application of the filters. We have found that phenomena inducing correlations in the light-curve sets tend to affect the entire

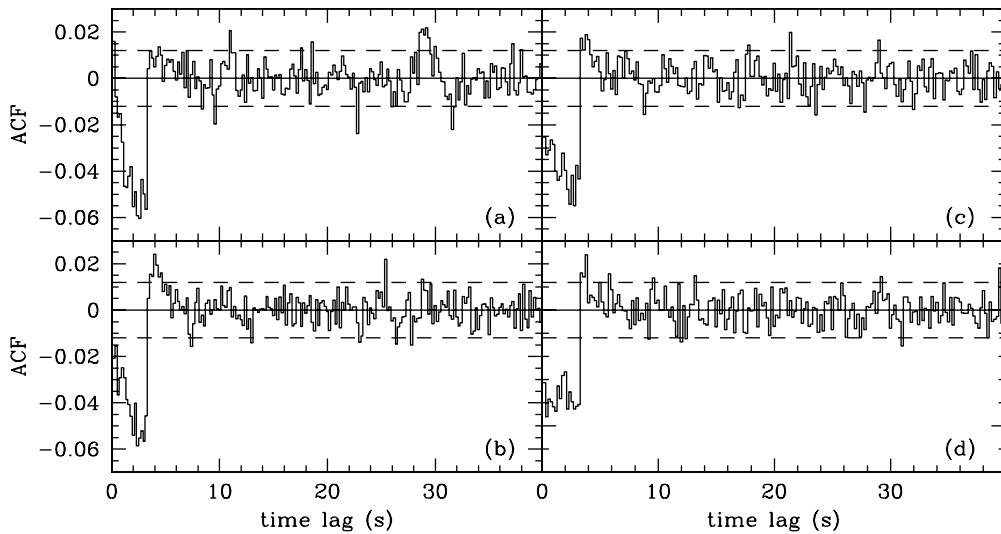


FIG. 7.—Autocorrelation plot from four light curves. Panels (a), (b), and (c): autocorrelation plots from three filtered TAOS light curves. Panel (d): autocorrelation plot from a synthetic white-noise light curve, after application of the mean and variance filters. Dashed lines are the 95% confidence level limits for what would be expected for randomly distributed light curves.

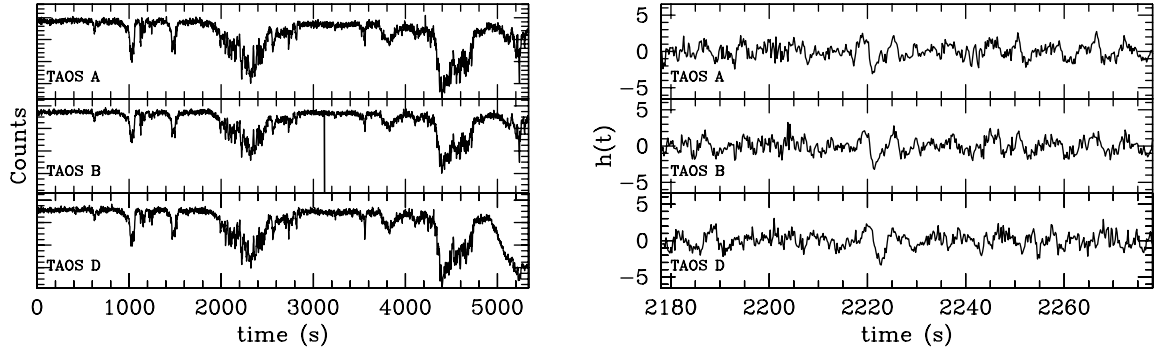


FIG. 8.—*Left panel*: An unfiltered light-curve set on a night with periods of cirrus cloud cover (during the periods of significant flux drops). *Right panel*: the same light-curve set after filtering, zoomed in to a period of cloud cover. Significant correlations are evident. Note that the correlation is stronger between telescopes TAOS A and TAOS B, which are close together (6 m separation). TAOS D, which is about 100 m away, has similar features but they are offset in time.

data run. Therefore, these diagnostic tests (described in the following subsections) are applied to entire data runs rather than individual light-curve sets. Data runs failing these tests are not considered for further analysis. The tests, described in the following subsections, were used in Zhang et al. (2008), Wang et al. (2009) and Bianco et al. (2010). An improved version of these tests has been developed for use in future analysis runs, and these new tests will be described in § 4.4.

4.1. Pearson's χ^2 Statistic

A simple test to determine if the light curves in a light-curve set are dependent is to divide the multi-telescope rank space into a grid and count the number of rank tuples in each grid element.

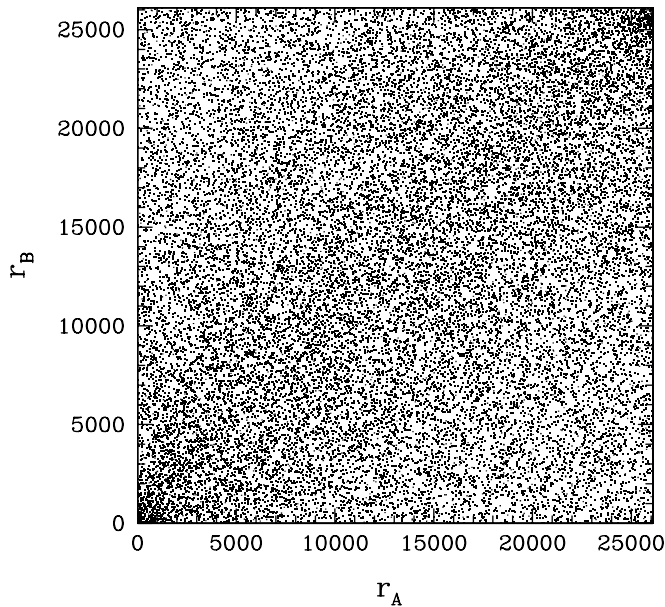


FIG. 9.—Rank-rank diagram (telescopes TAOS A and TAOS B) of the light-curve set shown in Fig. 8. The rank pairs are not uniformly distributed, as there are denser than average regions in the lower left and upper right corners.

This is illustrated in the left panel of Figure 10, in the case of two telescopes where the rank-rank diagram is divided into a $N_g \times N_g$ grid, where $N_g = 3$. With $N_p = 9$ and 9 grid elements, the expected number of rank pairs in each grid element is 1.

One can then perform a Pearson's χ^2 test on the number of rank pairs in each grid element by calculating

$$\chi^2 = \sum_{i=1}^{N_g^T} \frac{(O_i - \mathbb{E}_i)^2}{\mathbb{E}_i},$$

where O_i is the observed number of rank tuples in grid element i , and

$$\mathbb{E}_i = \frac{N_p}{N_g^T}$$

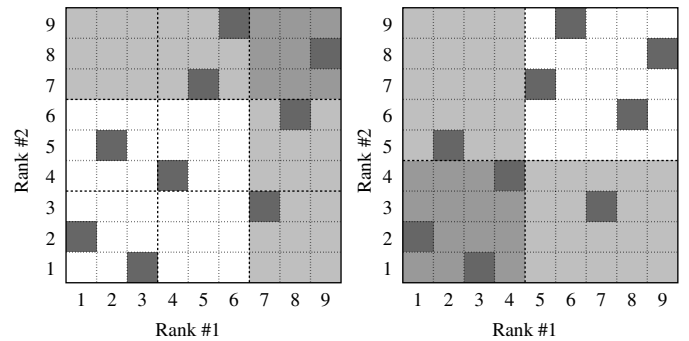


FIG. 10.—*Left panel*: rank-rank diagram illustrating the Pearson's χ^2 test. The rank-rank diagram is divided into a $N_g \times N_g$ grid ($N_g = 3$ in this case), and the number of rank pairs in each box is tabulated and compared with the expected uniform distribution. Counts in the gray elements are not free parameters. *Right panel*: rank-rank diagram illustrating the hypergeometric test. The test counts the number of objects in the lower left corner (dark shaded region) of the rank-rank diagram, in this case with a box size of $R = 4$. Note that there are four rank doublets with $r_{1j} \leq 4$ and four rank doublets with $r_{2j} \leq 4$ (light shaded regions) since each rank must occur exactly once in a light-curve set for each telescope. In this case there are three rank doublets where $r_{ij} \leq 4$ for both telescopes.

is the expected number of rank tuples in grid element i . (Note that \mathbb{E}_i may vary slightly among grid elements if N_p is not an exact multiple of N_g .)

For a given data run, we expect the distribution of the Pearson's χ^2 statistic to follow the χ^2 distribution, given by

$$p(u_c, \nu) = \frac{1}{2^{\nu/2} \Gamma(\nu/2)} u_c^{(\nu/2)-1} e^{-u_c/2}, \quad (7)$$

where $u_c = \chi^2$ and ν is the number degrees of freedom. The derivation of ν can be illustrated by Figure 10. It is important to note that every rank must appear once and only once in the time series for each telescope, and this constrains the value of ν . The degrees of freedom is the number of cells in the grid minus the number of independent constraints. First, the cell counts must sum to N_p , giving one constraint. Secondly, the counts in each grid row and each grid column must sum to three, giving two constraints on the three rows and on the three columns which are independent of each other and of the first constraint. Thus the total degrees of freedom are $9 - 2 - 2 - 1 = 4$. To illustrate, note that in the left grid column, the ranks 1, 2, and 3 must appear in telescope 1. Therefore, for telescope 2, since there are two doublets in the bottom left grid element and one doublet in the middle left grid element, there must be zero doublets in the top left element. The gray grid elements in the rank-rank diagram are thus not free parameters. For an arbitrary number of telescopes T , the number of degrees of freedom can be shown to be equal to

$$\nu = N_g^T - T(N_g - 1) - 1.$$

4.2. The Hypergeometric Test

While the Pearson's χ^2 test validates that the rank tuples are spread uniformly over $\{1 \dots N_p\}^T$, it is also useful to demonstrate that there is no bias toward rank quadruplets with all ranks relatively low, since these are the target events in the survey. Given a rank limit R , we define the variable u_h as the number of rank quadruplets with $r_{ij} \leq R$ for all telescopes i . This is illustrated in the case of two telescopes in the right panel of Figure 10, where we choose $R = 4$. In this figure, $u_h = 3$ is the number of rank doublets in the shaded lower left corner of the rank-rank diagram. Note that with $R = 4$ there are exactly four rank doublets with both $r_{1j} \leq R$ and $r_{2j} \leq R$. The probability distribution of the number of rank doublets with *both* ranks $r_{ij} \leq R$ is given by the hypergeometric distribution

$$\mathbb{P}(U = u_h) = \frac{\binom{R}{u_h} \binom{N_p - R}{R - u_h}}{\binom{N_p}{R}}, \quad (8)$$

where $u_h \leq R$ (if $u_h > R$ then $\mathbb{P} = 0$).

To expand this calculation to more than two telescopes, we use the *law of total probability* to calculate

$$\mathbb{P}(U_{i+1} = u_h) = \sum_{l=u_h}^R [\mathbb{P}(U_{i+1} = u_h | U_i = l) \times \mathbb{P}(U_i = l)], \quad (9)$$

where u_h is the number of measurements with $r \leq R$ on all telescopes 1 to $i + 1$. The conditional probability is defined as

$$\mathbb{P}(U_{i+1} = u_h | U_i = l) = \frac{\binom{R}{u_h} \binom{N_p - R}{l - u_h}}{\binom{N_p}{l}}, \quad (10)$$

given that each rank must occur exactly once for each telescope,

$$\mathbb{P}(U_1 = l) = \delta_{lR},$$

and one can thus expand equation (9) to include an arbitrary number of telescopes.

4.3. Application of Diagnostic Statistics

To date, the TAOS project has only analyzed data sets with light curves from three telescopes (Bianco et al. 2010; Wang et al. 2009; Zhang et al. 2008). Therefore we only describe the application of the diagnostic tests to three-telescope data. The development of these tests was a work in progress when the results presented in Bianco et al. (2010); Wang et al. (2009); Zhang et al. (2008) were calculated, and an improved method is described in § 4.4. However, we now present the original methods used to apply the diagnostic tests to the data to illustrate what was done to derive our previously published results. We note that for any analysis of TAOS data we will perform in the future, we will use the improved methods described in § 4.4.

For each data run, we apply both the Pearson's χ^2 statistic u_c and the hypergeometric test statistic u_h to each light-curve set. For the Pearson's χ^2 test, we use a grid size of $N_g = 5$, which corresponds to a total of $\nu = 112$ degrees of freedom. For the hypergeometric test, we set $R = N_p/5$, rounding to the nearest integer. (A typical 90 minute data run will have $N_p = 27,000$, but many runs are truncated due to bad weather.) Due to the fact that any correlations in the data may not show up in light curves with low S/N values, we perform the diagnostic tests only on those light-curve sets with $S/N \geq 10$. Details of the algorithm used to calculate S/N values of our light curves are given in Zhang et al. (2009). To summarize, we first calculate a 5σ -clipped rolling mean similar to that calculated in the mean filter, and then average the value of the rolling mean to get the signal. We then subtract the rolling mean from the raw light curve, and calculate a 5σ -clipped standard deviation of the new light curve, which we use as an estimate of the noise.

Even in the case of completely independent light curves, random chance will give rise to a number individual light-curve sets with aberrant values of the test statistic. We therefore look at the ensemble of test statistics for each data run, and require a match to the theoretical distributions. A set of examples is shown in Figure 11. The top panels show histograms of u_c and u_h for all of the light-curve sets in a data run with no evident dependence among the light curves, and the bottom panels show the same data for the pathological data run that contains the light-curve sets shown in Figures 8 and 9. Clearly, the data shown in the top panels match the theoretical distribution quite well, while the data in the bottom panels do not.

To determine which data runs are to be rejected, we test the goodness of fit of the distribution of test statistics over the light-curve sets in a data run to their theoretical distributions. To set a threshold, for each data run we calculate the quantity D_{\max} , which is defined as the absolute value of the maximum difference between the cumulative distribution of measured test statistics and the theoretical cumulative probability distribution. This is analogous to the Kolmogorov-Smirnov test (see Press et al. 1994, and references therein). For each data run, we calculate D_{\max} for both the Pearson's χ^2 test and the hypergeometric test. A scatter plot of these values is shown in Figure 12.

Data runs that fail either of the two tests are removed from the occultation event search. For data runs exhibiting widespread dependence between the light curves from different telescopes, we expect the measured distributions to differ significantly from the theoretical distributions, giving rise to large values of D_{\max} . Visual inspection of several data runs indicated that setting a cut on $D_{\max} > 0.2$ allowed us to reject

nearly all of the runs where the light curves exhibit clear, widespread dependence.

4.4. Improved Application of Diagnostic Tests

While the diagnostic tests described in the previous section are sufficient to remove nearly all of the data runs with significant dependence between light curves from different telescopes, they suffer from some limitations which motivated us to improve the techniques. First, the threshold D_{\max} was chosen somewhat arbitrarily after visual inspection of many data sets, since there is no way to determine empirically what the optimum threshold actually is. Second, in order for the measured statistical distributions to match the theoretical χ^2 and hypergeometric distributions, the original time series data in the light curves are required to be independent and identically distributed (i.i.d.), which is a stricter requirement than stationary and ergodic in mean. If some autocorrelation structure were present in the light curves, the light curves could still be stationary and independent from each other, however, the test statistics would not be expected to match the theoretical distributions. Finally, and most importantly, we would like to apply the same tests to light curves when searching for multipoint events. As discussed in § 3.4, for such event searches we would take a moving average of the light-curve data, or perhaps take the inner product of the light curve with some event template. Such filtering will induce autocorrelations into the light curves, and increase the S/N values as well. If we increase the S/N enough, some insignificant correlations between the light curves might in fact become significant in the filtered data. So it would be useful to apply the diagnostic tests to the data runs after the application of these

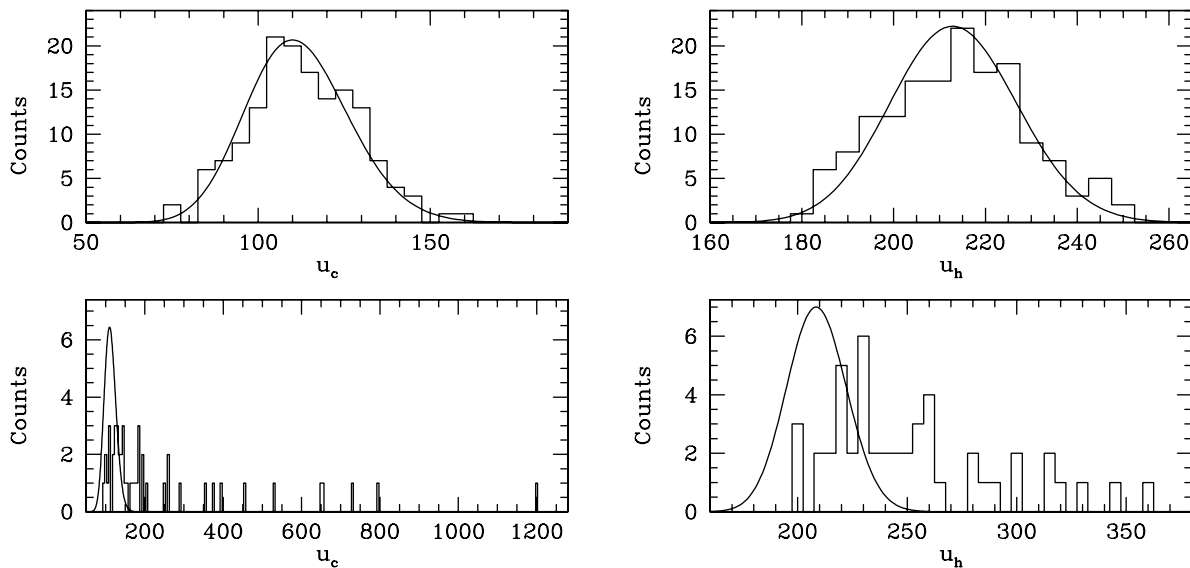


FIG. 11.—Results of diagnostic tests. *Histograms* indicate actual data, and *solid lines* indicate theoretical distributions. *Top panels*: Pearson's χ^2 test (*left*) and hypergeometric test (*right*) for a data run with no evident dependence between telescopes. *Bottom panels*: same as top panels, but for a data run with significant dependence between light curves. The light-curve set shown in Figs. 8 and 9 comes from this data run.

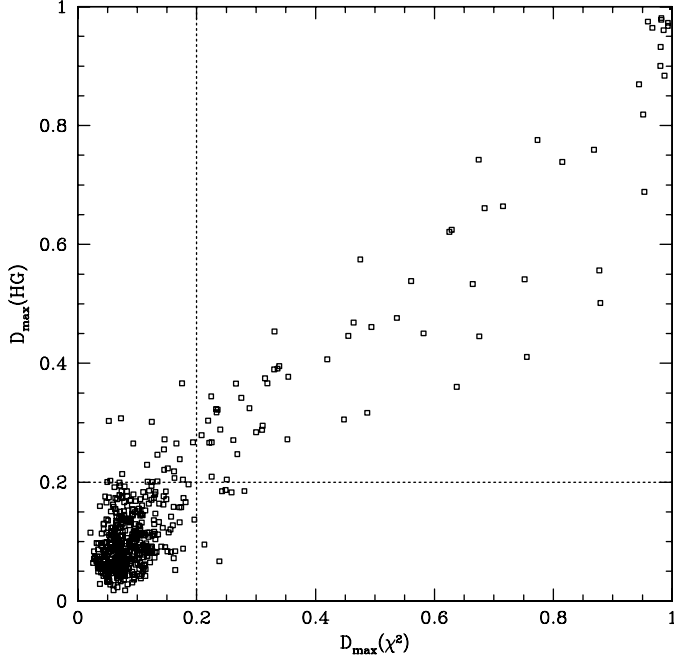


FIG. 12.—Scatter plot showing D_{\max} values for the Pearson's χ^2 test (x -axis) and hypergeometric test (y -axis). Each point corresponds to a single data run. For each test statistic, we reject data runs with $D_{\max} > 0.2$ (dotted lines).

filters. However, the introduction of significant autocorrelations into the light-curve data will more or less guarantee that all of the data runs will fail the diagnostic tests since the light curves will not be i.i.d.

We have thus developed a new technique based on the Blockwise Bootstrap (Künsch 1989) method (hereinafter BBS), which uses both the Pearson's χ^2 statistic u_c and the hypergeometric statistic u_h , described earlier, but requires no assumptions about the theoretical distributions for either statistic. The BBS test is implemented as follows. First, for a given light-curve set, we calculate both u_c and u_h . Then we divide each light curve in the light-curve set into 100 subsets, or *blocks*, of data. We then permute the blocks randomly, with each light curve in the light-curve set undergoing a different random permutation, and recalculate the diagnostic test statistics. We repeat this step a total of 99 times, and we are thus left with 100 statistical measurements for each of the diagnostic tests.

This is illustrated schematically in Figure 13. The top panel shows the original, unpermuted light curve, divided into five blocks. The blocks are labeled 1 through 5 for clarity. The bottom four panels show the same light curve with the five blocks randomly permuted. Note that the data within each block remain unchanged.

For each diagnostic test, we have now calculated 100 different values, one for the original light-curve set and 99 for the randomly permuted light-curve sets. If we permute the blocks we still preserve the stationary structure as long as the block size is large in comparison to the time scale of any autocorrelation.

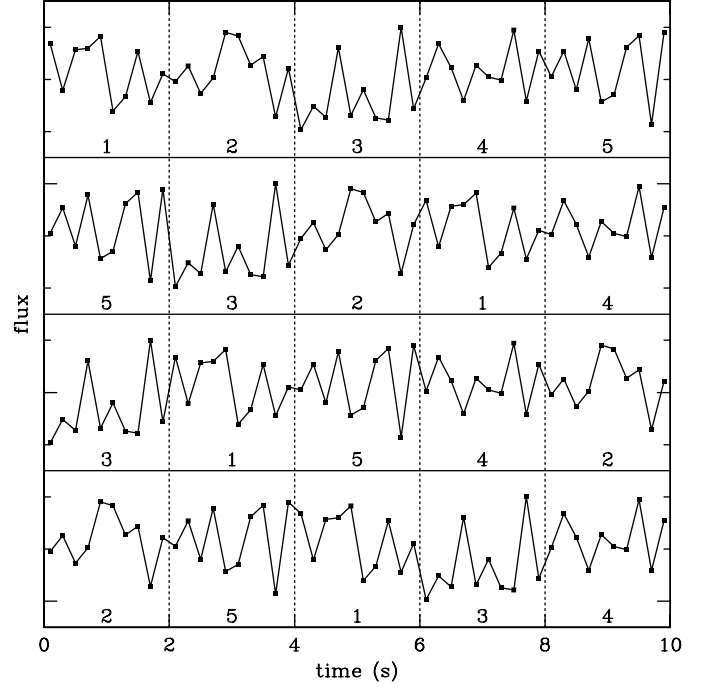


FIG. 13.—Illustration of the BBS test. *Top panel*: original light curve, divided into five blocks of data (dotted lines). *Bottom panels*: four light curves with the blocks permuted randomly. Blocks are labeled 1 through 5 for reference.

So if the light curves are independent, our 100 values are like 100 independent draws from the same distribution. Thus, if we rank each of the series of 100 test statistics from 1 to 100 (where a rank of one corresponds to the largest value of u_c or u_h , which would be the worst match to the expected distribution), the ranks should be uniform and we can calculate associated p -values as

$$\mathbb{P}(V < v_c) = \frac{v_c}{100} \quad (11)$$

and

$$\mathbb{P}(V < v_h) = \frac{v_h}{100}, \quad (12)$$

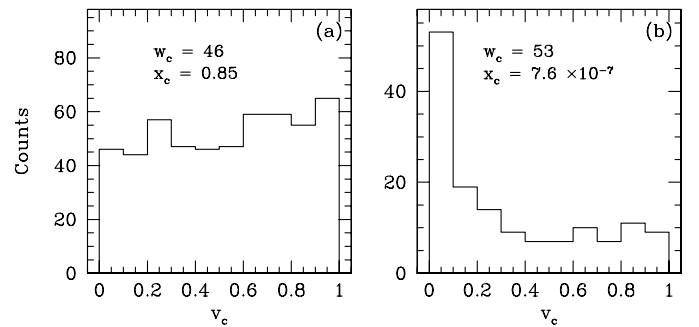


FIG. 14.—(a): distribution of p -values v_c from the Pearson's χ^2 test from a good data run. (b): same plot for a rejected data run.

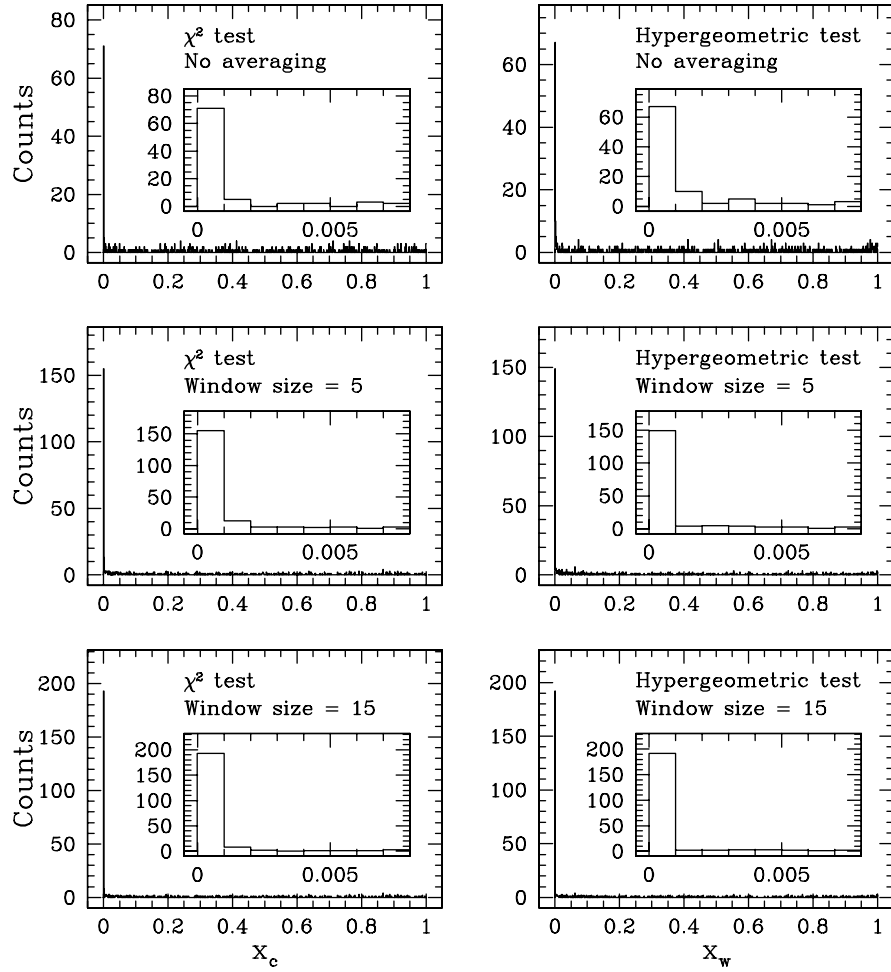


FIG. 15.—(a): Distribution of p -values from all data runs in the data set described in Bianco et al. (2010) for both the χ^2 and hypergeometric tests, with the original data and data after the application of moving average filters with window sizes of 5 and 15. *Inset plots* are zoomed in to the lowest p -values. The excess counts in the lowest bins are the runs exhibiting significant correlations between the telescopes.

where v_c and v_h correspond to the ranks of the test statistics u_c and u_h from the unpermuted (original) light-curve sets.

The BBS test is then performed on every light-curve set in a data run with $S/N > 10$. In the case of a data run that does not exhibit any strong dependence between the telescopes, the p -values v_c and v_h should be uniformly distributed on $\{0.01, \dots, 1\}$. However, in the case where there is significant dependence between the telescopes, we would expect the distributions of v to be clustered at small values, since any correlation between the telescopes would disappear when the blocks are randomly permuted. This is illustrated in Figure 14, which shows a histogram of the values v_c from each light-curve set in two different data runs. The histogram in Figure 14a shows the results from a data run with little dependence between the telescopes, while Figure 14b shows a data run with strong dependence.

In order to quantify the amount of dependence between the telescopes in a data run, we define two new test statistics, w_c and

w_h , which are defined as the number of light-curve sets in a data run with $v_c < v_t$ and $v_h < v_t$ respectively, where we choose $v_t = 0.1$ ¹⁵ (This corresponds to the lowest bin in the histograms shown in Fig. 14). In the case of independence between telescopes, the distributions of w_c and w_h follow the binomial distribution of the form

$$p(w_c) = \binom{L}{w_c} v_t^{w_c} (1 - v_t)^{L - w_c},$$

$$p(w_h) = \binom{L}{w_h} v_t^{w_h} (1 - v_t)^{L - w_h},$$

where L is the number of light-curve sets with $S/N > 10$ in the data run that are used to calculate the test statistics u_c and u_h .

¹⁵Tests have shown that as long as v_t is relatively small, the exact value chosen for v_t has no significant effect on the final results.

Using these distributions, we can calculate two test statistics for the entire data run, which we define as

$$x_c(w_c) = \mathbb{P}(W > w_c), \quad x_h(w_h) = \mathbb{P}(W > w_h).$$

For the data run in Figure 14a, we thus calculate $x_c = 0.85$, while for the data run in Figure 14b, we have $x_c = 7.6 \times 10^{-7}$.

We can now reject a data run for significant dependence by setting thresholds on x_c and x_h . In the absence of any significant dependence, the values of x_c and x_h should be distributed uniformly on the interval $[0, 1]$. Plots of the distributions of x_c and x_h statistics are shown in Figure 15. In order to illustrate the application of the BBS test to multipoint occultation searches, we also show the distributions after taking moving averages on the light curves with bin sizes of 5 and 15. The histograms shown have a bin size of 0.001, and with a total of 524 data runs we expect a value of about 0.5 for each bin. However, while the distributions appear to be uniform over most of the ranges in x values, note the large number of counts in the lowest bins. These are the light-curve sets that show dependence between telescopes. Note that some of the histograms show a slight excess in the second bin as well. By rejecting all data runs that appear in the first two bins ($x_c \leq 0.002$ and $x_h \leq 0.002$), we are clearly rejecting nearly all of the data runs exhibiting widespread dependence between the telescopes. Note that we only expect a total of one data run in the first two bins from random chance.

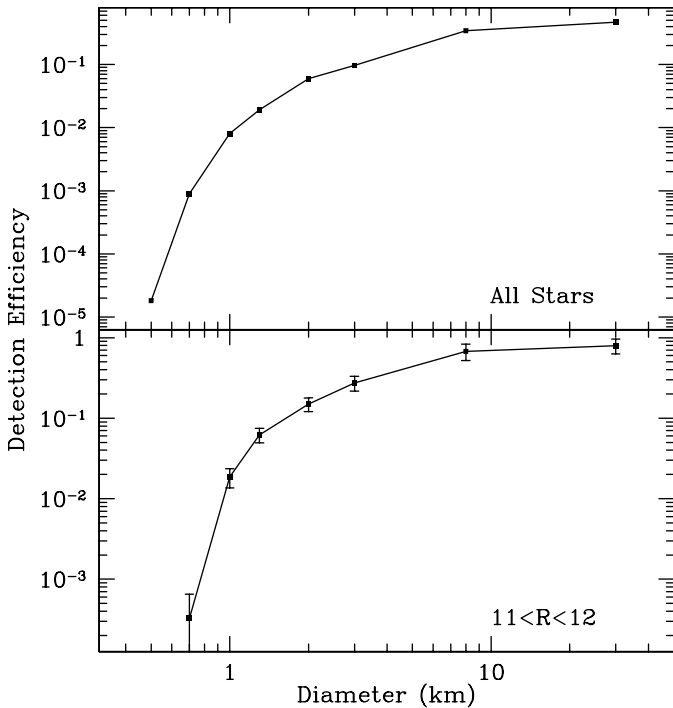


FIG. 16.—*Top panel:* detection efficiency vs. diameter for results published in Bianco et al. (2010) for all stars in the data set. *Bottom panel:* same as top panel, but only for stars with magnitudes $11 < R < 12$.

Note that our thresholds on x_c and x_h of 0.002 are based on the fact that the distributions of these values are uniform above these thresholds as expected, and there are spikes where we expect to find the worst data runs. These are reasonable choices of thresholds, given the clear changes in the distributions of x values from each of the tests. However, we note that there is no underlying theory which would allow one to set an optimal threshold. Our choice of threshold is thus context and data driven and not motivated by optimality conditions. The thresholds on these values will thus be determined on a case by case basis for future analyses.

Finally, note that in Figure 15, the larger the bin size on the moving average, the more data runs that are rejected. This is because of low level correlations that are insignificant in the unbinned data, but become significant in the binned data due to the increased S/N of the binned light curves.

The BBS test is clearly a superior method to the simple comparison of the test statistics to their theoretical distributions. It is very clear where the thresholds on x_c and x_h should be, the test will not reject data runs where the light curves are stationary but not i.i.d., and the tests are capable of robustly rejecting data runs when performing searches for multipoint occultation events.

5. CONCLUSION

We have developed a technique to search for extremely rare coincident events in voluminous multivariate (multi-telescope) time series data. Using rank statistics, this technique enables robust determination of event significance and false-positive rate, *independent of the underlying noise distributions in the time series data*. Furthermore, we have developed a method to test for widespread dependence between light curves in a data run, which allows us to reject runs with inherent characteristics that could possibly give rise to a larger false-positive rates. We note that while the method described in this article is sufficient for the calculation of the rate of false-positive events that arise due to random statistical chance, it is not capable of estimating the background event rate due to systematic errors in the TAOS photometry (Zhang et al. 2009). For example, tracking errors or moving objects in the images could give rise to false detections in the data set. A description of how such background events are handled is described in Bianco et al. (2010).

This method has been used to search for rare occultation events by KBOs in over 500 data runs comprising a total of nearly 370,000 light-curve sets (Bianco et al. 2010). To calculate the upper limit on the size distribution, we performed a simulation in which simulated events were added to the light-curve data. A search algorithm based on the statistical algorithm described in this article was then used to measure our detection efficiency. The results are shown in Figure 16. The top panel shows our detection efficiency as a function of object diameter averaged over all of the light-curve sets in the data set. In this plot, all of the detection efficiency at the small end comes from the brightest stars with the highest S/N. At the larger end, the

efficiency reaches a maximum of about 47%. This is due to the inclusion of many faint stars in the data set with such low S/N values that any event is virtually undetectable. (Note that the inclusion of these stars in the data set has no bearing on the final upper limit). In order to better illustrate the effectiveness of the search algorithm, we plot the detection efficiency as a function of diameter only for those stars with magnitudes $11 < R < 12$ in the bottom panel of Figure 16. (We thus exclude the very bright stars with the highest S/N, as well as the faint stars with which no events are detectable). The detection efficiency is very high for objects with diameters $D \gtrsim 3$ km, and drops rapidly when $D < 1$ km.

Finally, we note that the rank product test is useful only for the rejection of the null hypothesis that no occultation event is present, and to estimate the false-positive rate. In order to definitively show that outer solar system objects have been detected, a number of events would need to be detected in order to correlate the surface density with ecliptic latitude. Any physical parameters of an occulting object (such as size and distance) can be estimated for high S/N events from the shape of the light curves (Nihei et al. 2007). However, this would be difficult to do with the TAOS data, primarily due to the relatively slow cadence of the TAOS observations. (Some limited size and distance information could be determined for occultations by objects well beyond 100 AU [Wang et al. 2009]). Surveys with a higher readout rate would in fact be able to make reasonable estimates of these physical parameters for

high S/N events (Schlichting et al. 2009; Bickerton et al. 2009, 2008; Roques et al. 2006, 2003; Chang et al. 2007), and a next generation multi-telescope occultation survey (TAOS II, currently in the early development stage), which will operate with a readout cadence of 20 Hz, will also be able to do so after using the rank product method to identify candidate events.

This work was supported in part by the thematic research program AS-88-TP-A02 at Academia Sinica. N. K. C.'s work was supported in part by the National Science Foundation under grant DMS-0636667. Work at the Harvard College Observatory was supported in part by the National Science Foundation under grant AST-0501681 and by NASA under grant NNG04G113G. The work at National Central University was supported by the grant NSC 96-2112-M-008-024-MY3. J. A. R.'s work was supported in part by the National Science Foundation under grant AST-00507254. S. L. M.'s work was performed under the auspices of the US Department of Energy by Lawrence Livermore National Laboratory in part under Contract W-7405-Eng-48 and by Stanford Linear Accelerator Center under Contract DE-AC02-76SF00515. Y. I. B. acknowledges the support of National Research Foundation of Korea through Grant 2009-0075376. K. H. C.'s work was performed under the auspices of the US Department of Energy by Lawrence Livermore National Laboratory in part under Contract W-7405-Eng-48 and in part under Contract DE-AC52-07NA27344.

APPENDIX

EVALUATION OF THE K FUNCTION

Here we present an algorithm to evaluate $K(x; k, n)$, the number of ways to get a product of x by multiplying k integers between 1 and n , which is applicable when $x \leq n$.

Note that $K(1; k, n) = 1$. For $x > 1$, consider the prime decomposition of x where the p s are unique primes and d is their degree so that:

$$x = p_1^{d_1} \times p_2^{d_2} \times \dots \times p_m^{d_m},$$

where m is the total number of prime factors of x .

We claim that

$$K(x; k, n) = \prod_{i=1}^m \binom{d_i + k - 1}{k - 1}.$$

A1. Proof

Suppose $A_1 \times A_2 \times \dots \times A_k = x$ and take prime decompositions of each number

$$\begin{array}{lll} A_1 & = & p_1^{d_{1,1}} \times p_2^{d_{1,2}} \times \dots \times p_m^{d_{1,m}} \\ A_2 & = & p_1^{d_{2,1}} \times p_2^{d_{2,2}} \times \dots \times p_m^{d_{2,m}} \\ \vdots & & \vdots \\ A_k & = & p_1^{d_{k,1}} \times p_2^{d_{k,2}} \times \dots \times p_m^{d_{k,m}}. \end{array}$$

Note that

$$\sum_{i=1}^k d_{i,j} = d_j \quad \forall j.$$

Hence,

$$K(x; k, n) = \prod_{i=1}^m S(d_i; k),$$

where $S(d; k)$ is the number of ways to get a sum of d by adding k integers where $0 \leq k \leq d$. The calculation of the function S is best illustrated with an example. Consider the case of $d = 10$ and $k = 4$. If we illustrate the sum $d = 10$ with 10 dots in the top row of Figure 17, the function S is simply the number of ways to divide the dots into 4 groups (using the bars shown).

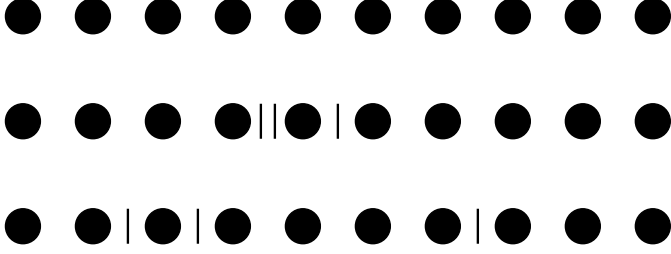


FIG. 17.—Schematic illustrating the calculation of the function $S(d; k)$. In this case, $d = 10$, as indicated by the top row of 10 dots. The bars in the bottom two rows indicate possible ways to split the 10 dots into four addends. The second row indicates the tuple $(4, 0, 1, 5)$, and the third row indicates the tuple of $(2, 1, 4, 3)$.

For example, the second row of Figure 17 corresponds to a tuple of $(4, 0, 1, 5)$, while the third row corresponds to a tuple of $(2, 1, 4, 3)$. So the number of possible 4-tuples is the number of ways to choose 3 bar locations in a total $10 + 3 = 13$ possibilities. This gives

$$S(d; k) = \binom{d+k-1}{k-1}.$$

For example, to calculate $K(6; 4, n)$ where $n \geq 6$, we note that $6 = 2 \times 3$ is the product of two primes to the first power, and in the four-telescope case it is equal to

$$K(6; 4, n) = \binom{1+4-1}{4-1}^2 = \binom{4}{3}^2 = 4^2 = 16,$$

in agreement with Table 1. Note that this formulation is only valid if $x \leq n$. For example, $K(6; 4, 5) = 12$, since any of the rank tuples with a rank value of 6 would be impossible in a light-curve set containing only 5 points.

As a second example, for the case of $360 = 2^3 \times 3^2 \times 5$, we have three primes with degrees 3, 2, and 1. We thus have (for $n \geq 360$)

$$\begin{aligned} K(360; 4, n) &= \binom{3+4-1}{4-1} \binom{2+4-1}{4-1} \binom{1+4-1}{4-1} \\ &= \binom{6}{3} \binom{5}{3} \binom{4}{3} = 20 \times 10 \times 4 = 800. \end{aligned}$$

REFERENCES

- Bertin, E., & Arnouts, S. 1996, *A&AS*, 117, 393
 Bianco, F. B., et al. 2009, *AJ*, 138, 568
 ———. 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. E., Trujillo, C., & Rabinowitz, D. 2004, *ApJ*, 617, 645
 Chang, H.-K., et al. 2007, *MNRAS*, 378, 1287
 Coehlo, N. K. 2010, Ph.D thesis, Univ. California, Berkeley
 Cooray, A. 2003, *ApJ*, 589, L97
 Cooray, A., & Farmer, A. J. 2003, *ApJ*, 587, L125
 Da Costa, G. S. 1992, in *ASP Conf. Ser. 23, Astronomical CCD Observing and Reduction Techniques*, ed. S. B., Howell, 90
 Fisher, R. A. 1958, *Statistical Methods for Research Workers* (New York: Hafner)
 Kovács, G., Bakos, G., & Noyes, R. W. 2005, *MNRAS*, 356, 557
 Künsch, H. R. 1989, *Ann. Stat.*, 17, 1217
 Lehner, M. J., et al. 2006, *Astron. Nachr.*, 327, 814
 ———. 2009, *PASP*, 121, 138
 Mosteller, F. J., & Fisher, R. J. 1948, *The American Statistician*, 2, 30
 Nihei, T. C., et al. 2007, *AJ*, 134, 1596
 Press, W. H., et al. 1994, *Numerical Recipes in C* (New York: Cambridge Univ. Press)
 Roques, F., et al. 2003, *ApJ*, 594, L63
 ———. 2006, *AJ*, 132, 819
 Schlichting, H. E., et al. 2009, *Nature*, 462, 895
 Tamuz, O., Mazeh, T., & Zucker, S. 2005, *MNRAS*, 356, 1466
 ———. 2009, *AJ*, 138, 1893
 ———. 2010, *AJ*, 139, 2003
 Zhang, Z.-W., et al. 2009, *PASP*, 121, 1429
 Zhang, Z.-W., et al. 2008, *ApJ*, 685, L157