

## AN ANALYSIS OF THE STATISTICS OF THE *HUBBLE SPACE TELESCOPE* KUIPER BELT OBJECT SEARCH

MICHAEL E. BROWN AND SHRINIVAS R. KULKARNI

Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA 91125; mbrown@gps.caltech.edu, srk@astro.caltech.edu

AND

TIMOTHY J. LIGGETT

Department of Physics, Harvey Mudd College, Claremont, CA 91711

Received 1997 June 26; accepted 1997 September 4; published 1997 October 22

### ABSTRACT

We calculate statistical limits to the detection of Kuiper belt objects in the *Hubble Space Telescope* (*HST*) data of Cochran et al., in which they report the discovery of a population of Halley-sized objects in Pluto-like orbits. Detection of a population of faint objects in these data is limited by the number of false objects that appear owing only to random noise; the number of real objects must exceed the uncertainty in the number of these false objects for the population to be observable. We determine the number of false objects expected owing to random noise in the data of Cochran et al. by measuring the pixel-to-pixel noise level in the raw *HST* data and propagating this noise through the detection method employed by Cochran et al. We find that the uncertainty in the number of false objects exceeds by 2 orders of magnitude the reported number of objects detected by Cochran et al. The detection of such a population of Halley-sized Kuiper belt objects with these data is therefore not possible.

*Subject headings:* methods: statistical — comets: general

### 1. INTRODUCTION

The existence of a belt of remnant planetesimals beyond the orbit of Neptune has been suspected for almost 50 years (Edgeworth 1949; Kuiper 1951). This Kuiper belt has more recently been proposed as the origin of the Jupiter family comets (Fernandez 1980), a suggestion that has been supported by numerical integrations of Duncan, Quinn, & Tremaine (1988). As of this writing, 53 large ( $\geq 100$  km diameter) objects have been discovered between 36 and 49 AU by ground-based observers (see, for example, Jewitt & Luu 1995), and the existence of the belt is well established. Smaller objects, however, such as the 1–10 km progenitors of typical short-period comets, cannot be seen in the ground-based searches, and the direct connection between the short-period comets and the Kuiper belt remains uncertain.

Recently, Cochran et al. (1995; hereafter CLSD) reported the statistical detection of  $\sim 10$  km Halley-sized objects in the Kuiper belt from deep *Hubble Space Telescope* (*HST*) images. Their analysis suggests a density of  $\sim 2.5 \times 10^4$  objects per square degree brighter than  $V = 28.6$ . This density of Kuiper belt objects (KBOs) is higher than might be expected: extrapolation from detections of brighter objects predicts a density 2 orders of magnitude lower. The reported detection appears to be in agreement, however, with the number of objects required to account for the population of Jupiter family comets (Weismann & Levison 1996).

We have examined the data of CLSD and have calculated statistical limits to the detection of KBOs. We find that, contrary to the analysis of CLSD, the reported population of Halley-sized objects could not have been detected in these data. In the following sections, we present a brief synopsis of the *HST* data used for the KBO search, our analysis of the search method of CLSD, and a discussion of the discrepancies between our calculated limits and the reported detections of CLSD.

### 2. THE *HST* OBSERVATIONS AND SEARCH METHOD

The data of CLSD consist of 34 WFPC2 F606W (a 1600 Å wide filter center at 5840 Å) 500–600 s exposures (for a total

integration time of 18,000 s) obtained with the telescope pointing at a fixed location in the ecliptic approximately in the direction of the Earth's motion. This pointing geometry is ideal for searching for KBOs given the small field of view of WFPC2: almost no parallax motion appears, and nearly all observed motion is due to the small orbital motion of the objects. At typical KBO distances, orbital motions project to approximately  $1'' \text{ hr}^{-1}$ , which corresponds to about  $0''.13$  in a 500 s exposure. Over the 30 hr elapsed time from the first to the last exposure, a KBO crossing the field will move approximately one-third of the way across a single WFPC2 CCD.

No KBOs appeared in the individual images, so the images had to be combined in such a manner as to enhance the visibility of any KBOs present. To accomplish this enhancement, CLSD chose a potential orbit, consisting of an orbital speed and inclination, and shifted the 34 images so that any object moving along the chosen orbit appeared stationary between images. The 34 shifted images were then combined in a median sum to make a higher signal-to-noise ratio image of any KBOs in the field with the chosen orbit. CLSD found no bright KBOs at this step. A search for faint objects in the median image was then performed by summing values of the 5 contiguous pixels in a “+” pattern centered on each pixel and searching for sums brighter than a given magnitude limit. Additional filtering was performed in an attempt to remove the effects of cosmic rays. The speed and direction of KBO motion differ depending on the orbit of the object, so different orbits were considered separately. For a complete search, this procedure is then repeated for all potential KBO orbits.

In their limited preliminary analysis, CLSD considered only 154 positively inclined orbits in the 3:2 mean motion resonance with Neptune; a total of 53 candidate KBOs brighter than  $V = 28.6$  was identified. These candidate objects include both potentially real objects and objects that appear real owing only to the fortuitous combination of random noise (“false objects”). To determine the number of these false objects, the entire search procedure was repeated using retrograde orbits, which no real objects are expected to have. A total of 27 objects was found

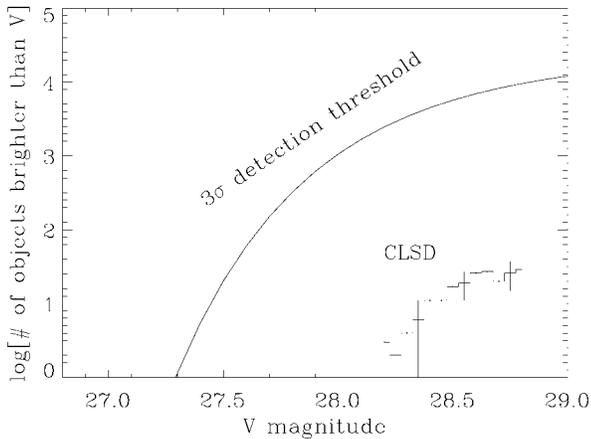


Fig. 1.—Comparison of the number of KBOs reported by CLSD and the statistical detection threshold determined by the uncertainty in the number of objects expected to be detected owing to random noise in the data (“false objects”) as a function of limiting magnitude. A statistical detection of objects can be made only if the number of real objects exceeds the uncertainty in the false object rate; the solid line therefore gives a lower limit to the number of objects that could possibly be detected in the data. The reported detection of CLSD is more than 2 orders of magnitude below this lower limit.

in 154 retrograde orbits. The excess of objects in prograde orbits over those in the false object retrograde orbits was interpreted as a statistical detection of  $26 \pm 9$  real KBOs brighter than  $V = 28.6$ .

### 3. ANALYSIS OF THE CLSD METHOD

We now calculate detection limits for the data of CLSD. Here detection limits refer not to the ability to detect single objects (most objects under consideration here are multi- $\sigma$  events in the data) but rather to the ability to determine statistically that the number of objects brighter than a given magnitude limit exceeds the expected number of false objects. The signal-to-noise ratio of a detection in this case is the number of real objects divided by the uncertainty in the number of false background objects, with the uncertainty in the false object background being the square root of the number of false detections. Using these criteria, we can calculate the minimum number of real objects that must be present for a statistically significant detection of a population of KBOs. For example, if, at a given magnitude limit, 100 false object detections are made, a minimum of  $3 \times \sqrt{100} = 30$  real KBOs brighter than this magnitude limit must be present to be detected at a  $3 \sigma$  confidence level above the background. If fewer real objects are present, they will not be distinguishable from the uncertainty in the number of false objects.

Using measurements of the pixel-to-pixel noise level in the raw *HST* data of CLSD (extracted from the *HST* data archive) and propagating this noise through the detection method employed by CLSD, we determine the number of false objects expected in the data and, therefore, the lower limit to the number of real objects that are detectable at a given magnitude. To determine the pixel-to-pixel noise level in the 34 individual images, we first construct a median image from the 34 images and subtracted this image from each of the individual images to remove any real stationary objects. The noise level is then determined in a series of small boxes in various locations around the image, taking care to exclude both cosmic rays and any regions near bright objects in the medianed image (we exclude the regions near bright objects because, though the objects themselves have been removed, the nearby regions will

still show higher than average background noise). We find that the average  $1 \sigma$  pixel-to-pixel noise level is 1.67 counts for 500 s exposures and 1.80 counts for 600 s exposures. These noise levels agree well with the values of 1.63 and 1.78 counts expected for the noise due to readout, dark current, and sky background for WFPC2 exposures of these lengths with the F606W filter in the direction pointed (Biretta et al. 1996). Adding the noise in quadrature, the  $1 \sigma$  pixel-to-pixel noise is  $12.5 \text{ pixel}^{-1}$  in a median sum of the 34 images [the noise in the median is  $(\pi/2)^{1/2}$  greater than the noise in a mean], or 27.9 counts in the sum of a 5 pixel “+” pattern as used by CLSD.

We must now consider how much of the light from a typical KBO will appear inside the 5 pixel “+” pattern. For a point source perfectly centered within the central pixel, only 64% of the light will appear inside the 5 pixel “+” pattern (Biretta et al. 1996). For a sum of 34 images shifted to a common center, even less light will appear inside the 5 pixel “+” pattern. Furthermore, mismatches between the true orbital parameters of the KBO and those searched will also smear the object. To determine the effects of these problems, we have performed a simulation in which we embed objects of known magnitude and orbital parameters into 34 artificial images and shift the images using matched and slightly mismatched test orbits. In CLSD, parameter space is explored with a grid spacing of  $1^\circ$  in inclination and  $\sim 0.043 \text{ hr}^{-1}$  in orbital speed; thus, the test orbits could be off by as much as  $0.5^\circ$  in inclination and  $0.021 \text{ hr}^{-1}$ . For a perfectly matched orbit, we find that 52% of the light falls within the 5 pixel “+” pattern. For the maximum mismatch, the fraction is only 22%. Accurate incorporation of this effect into the statistical limits we are calculating is difficult (and the effect is ignored by CLSD); instead, we will generously assume that all objects detected by CLSD have perfectly matched orbits and that 52% of the light falls within the 5 pixel “+” pattern. This assumption causes us to overestimate the ability to detect faint moving objects; thus, our calculation will be only a lower limit to the number of objects required for a statistically significant detection.

With the noise level calculated above, we can now calculate the expected number of false objects brighter than a given limiting magnitude. The number of counts,  $n$ , produced by an object of magnitude  $m$  in an exposure of length  $t$  using filter F606W of WFPC2 is (Leitherer et al. 1995)

$$n = t \times 10^{(22.933 - m)/2.5}. \quad (1)$$

An object of magnitude 28.6 (the limiting magnitude of CLSD) produces 97 total counts in 18,000 s (the combined exposure time of the 34 CLSD images). Of these counts, 52%, or 50 counts, appear inside a 5 pixel “+” pattern. This count level is 1.79 times the  $1 \sigma$  background level of 27.9 counts, calculated above, so at this magnitude limit, a false object will appear in  $1 - f(1.79) = 3.67\%$  of the 5 pixel “+” patterns examined, where  $f(x)$  is the Gaussian distribution function. For a typical orbit, CLSD examine  $\sim 10^6$  such 5 pixel “+” patterns (the number is determined by the size of the WFPC2 field and the distance that KBOs move in the 30 hr between the first and last image), so in each orbit  $3.7 \times 10^4$  false objects should appear. A  $3 \sigma$  detection of real objects over this background therefore requires the presence of at least  $3 \times (3.7 \times 10^4)^{1/2} = 574$  real objects per orbit considered, or  $574 \times \sqrt{154} = 7134$  objects in the 154 orbits of CLSD. Any smaller number of real objects, such as the 27 objects reported by CLSD, cannot be detected above the false object uncertainty in these data.

Figure 1 shows the minimum number of objects detectable

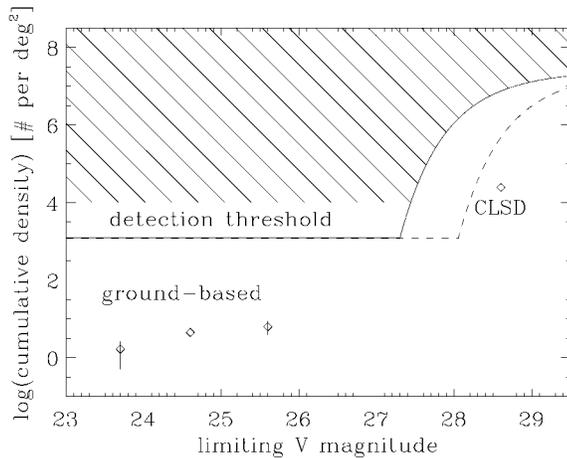


FIG. 2.—Lower limits to the detection of Kuiper belt objects in the data of CLSD. The solid line shows the  $3\sigma$  detection lower limit; only above this line can objects be detected by the statistical methods of CLSD. Below this line the uncertainty in the number of false objects is greater than the number of real objects, so no detection can be made. (For objects brighter than  $V = 27.4$ , no false objects are expected, so the detection limit is simply set by the requirement that one object appear within the small field of view of the WFPC2.) For comparison, the reported detection of CLSD and the number of objects detected from ground-based surveys (Jewitt & Luu 1995) are also shown. The dashed line shows the approximate detection threshold that would be expected for observations with a total integration time 4 times larger than in the current data, such as those approved for *HST* cycle 7.

above the false object uncertainty as a function of limiting magnitude, calculated using the above method. At all magnitude limits, the uncertainty in the false object background level greatly exceeds the number of objects reported by CLSD. In Figure 2 we convert the calculated detection limits into number density limits and again compare these to CLSD.

Note that in these last two sections, we have ignored several of the CLSD data processing steps, in particular the subtraction of a smooth background from the data and the cosmic-ray filtering (the CLSD six-group median, discussed in more detail below). Our analysis assumes as initial conditions uniform Gaussian noise and a perfectly flat background; the processing steps of CLSD are an attempt to make the real data approach this ideal. Because we are ignoring the effects of cosmic rays and nonuniform backgrounds, the noise level that we calculated for the data is actually a lower limit to the noise in the real data. The expected number of false objects and the detection limits are therefore also lower limits based on the assumption of ideal noise and perfect analysis.

#### 4. DISCUSSION

Objects of the faintness of those reportedly detected by CLSD are detectable above the expected false object background only if they are more than 2 orders of magnitude more abundant than the number reported by CLSD. The discrepancy between the limits calculated here and the reported detections of CLSD apparently arises from an underestimate of the false object rate by CLSD. We have used the noise characteristics of the data to calculate an expected false object rate of  $3.7 \times 10^4$  objects with  $V < 28.6$  per orbit; CLSD estimated a false object rate of 0.17 per orbit by their search for objects in retrograde orbits. These two methods should be equivalent, yet the statistical lower limit to the expected false object background is more than 5 orders of magnitude higher than the CLSD apparent measured background.

We suggest two effects that likely contribute to the CLSD

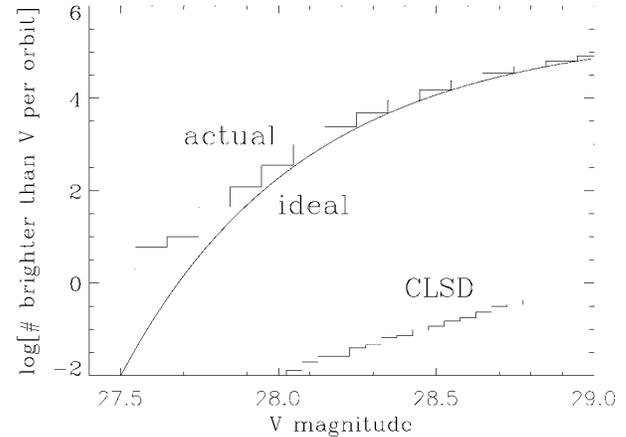


FIG. 3.—Number of false objects found in the CLSD data in a single orbit using an optimized technique that avoids the smoothing and overfiltering of the CLSD method. The number of objects found agrees well with the number calculated from ideal statistical expectations of the data. In contrast, the number of false objects reported by CLSD is more than 5 orders of magnitude lower than the actual and expected lower limit.

underestimate of the background rate. The first effect is the inadvertent smoothing of the images during processing. Small amounts of smoothing can produce enormous changes in the number of background objects this close to the noise level, and images are smoothed every time they are shifted by fractional pixel amount and the new pixel values are interpolated from the old. Shifting 34 frames by random fractional pixel amounts smooths the summed data by about 34%, which results in an apparent 38% reduction in the number of false noise objects.

Another method by which CLSD likely artificially reduce the number of false objects is by overfiltering the data. CLSD attempt to develop a filtering method for “determining which of the resultant ‘objects’ were real and which were just coincidental alignments of noise.” Unfortunately, no such filter exists: there is no statistical difference between the distribution of pixel values that make a real object and the distribution of pixel values that make a false object with the same mean (other than the small extra noise added by the Poisson statistics of the real object). This situation is a consequence of the fact that, for a Gaussian-distributed set of numbers, the mean and variance of the distribution are statistically independent (Frodeson, Skjeggstad, & Tofte 1979), so random noise and real objects cannot be distinguished by differences in variance about the mean. The CLSD multistage filtering does indeed reduce the apparent number of false objects, but only at the expense of also removing the same number of real objects, since the two are indistinguishable. Depending on the filtering method used, an arbitrarily small number of false objects can be made to appear, but this filtering does not improve the ability to detect real objects above the false object background.

To demonstrate empirically the large number of expected false objects, we have taken the CLSD data and have processed a single retrograde orbit to determine the number of false objects detected as a function of limiting magnitude. To avoid the problems discussed above, we modify the CLSD method somewhat. First, shifting of the images is performed only once, in order to take into account the total effects of telescope dithering, WFPC2 geometric distortion, and orbital motion. In addition, the shift performed is a nearest neighbor, rather than a linear, interpolation. Such a shift preserves the noise characteristics of the original data while slightly broadening the PSF of any real objects. Second, we take care not to overfilter the

data when searching for objects. Instead of the two-step six-group–median scheme employed by CLSD, we perform a single clipped sum to remove cosmic rays: the pixel values in each of 34 stacked 5 pixel “+” patterns are sorted, the highest 30 values are discarded, and sum is taken of the remaining 140 pixel values. Based on statistics of the numbers of cosmic rays in the data, no cosmic-ray events should remain in the clipped sum. This filter comes close to reproducing the ideal cosmic-ray–free method of simply summing the pixel values. Figure 3 shows the number of false objects found brighter than a given magnitude for a single retrograde orbit using this method and compares it to the expected number of false objects assuming perfect Gaussian statistics as calculated using the method above. The actual number of false objects found slightly exceeds the calculated number, as is expected owing to the inherent non-Gaussian nature of effects such as cosmic rays, but

the otherwise close match between the expected and actual numbers demonstrates in practice as well as in theory the impossibility of the detection reported by CLSD.

Regardless of the cause of the underestimate of the number of false background objects, which remains unclear, these simple statistical calculations show that the small population of Halley-sized bodies in the Kuiper belt reported by CLSD cannot be detected in these data. Statistical detection limits for these data have been estimated, and a complete reanalysis of the data will be necessary to determine if any objects appear above these higher limits.

This research was supported by grant HF-01056-94A from the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Incorporated, under NASA contract NAS5-26555.

#### REFERENCES

- Biretta, J. A., et al. 1996, WFPC2 Instrument Handbook, Version 4.0 (Baltimore: STScI)
- Cochran, A. L., Levison, H. F., Stern, A. S., & Duncan, M. J. 1995, *ApJ*, 455, 342
- Duncan, M., Quinn, T., & Tremaine, S. 1988, *ApJ*, 328, L69
- Edgeworth, K. E. 1949, *MNRAS*, 109, 600
- Fernandez, J. A. 1980, *MNRAS*, 192, 481
- Frodesen, A. G., Skjeggstad, O., & Tofte, H. 1979, *Probability and Statistics in Particle Physics* (Bergen: Universitetsforlaget)
- Jewitt, D., & Luu, J. 1995, *AJ*, 109, 1867
- Kuiper, G. P. 1951, in *Astrophysics: A Topical Symposium*, ed. J. A. Hynek (New York: McGraw-Hill), 357
- Leitherer, C., et al. 1995, *HST Data Handbook*, Version 2 (Baltimore: STScI)
- Weismann, P., & Levison, H. 1996, *Pluto-Charon* (Tucson: Univ. of Arizona Press), in press