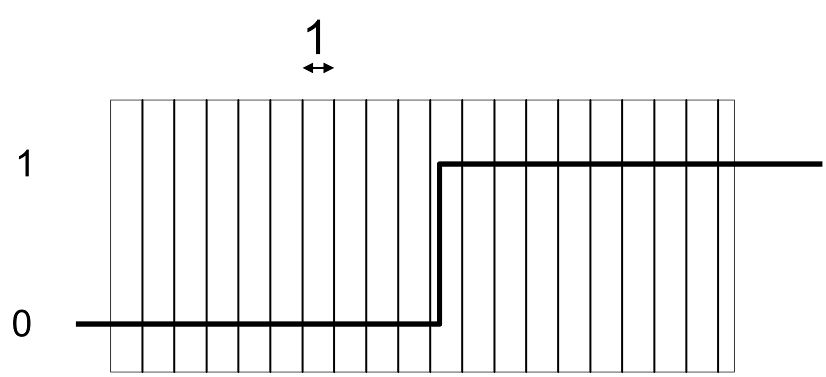
A Study of Occultation Timing Accuracy – Revision 2

Frank Freestar8n

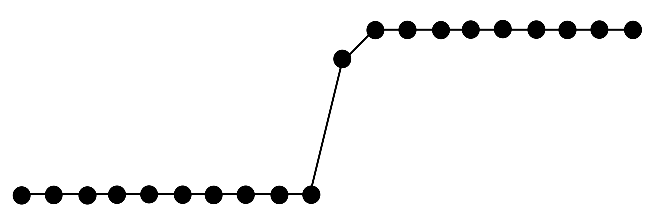
Abstract

There are many ways to time an occultation, but the general principle is to record the light intensity of a star at regular intervals as it disappears and reappears. The resulting accuracy of the times depends on the noise in the measurement, the time between measurements, and the nature of the intensity change, e.g. the presence of diffraction and the possibility of multiple steps due to the star being double. In this study I examine the specific case of a single abrupt transition in the presence of noise, and I use simulation methods to determine the time of the event and its associated error. I show that for small amounts of noise the timing error is proportional to the noise and can be much less than the interval between measurements, whereas at the high end it becomes quadratic with noise and can be much greater than the time interval.

In this work I focus exclusively on the timing accuracy of a single abrupt transition in the intensity of a star recorded at regular intervals in the presence of noise. If the noise is small relative to the intensity change, there will be exactly one interval that captures the transition event as an intermediate intensity. The fraction of recorded intensity corresponds to the fractional timing of the event in the frame and, with interpolation, can yield sub-frame accuracy of the timing. With greater noise, however, even the interval of the event is not known with certainty, leading to error much greater than the timing interval. This is a formal study of one way to estimate the most likely time of the transition event, along with error estimates in the presence of noise. There is a separate problem of identifying whether or not a brief occultation actually occurred, since a brief dip in a noisy trace is hard to interpret. That problem is not addressed here.

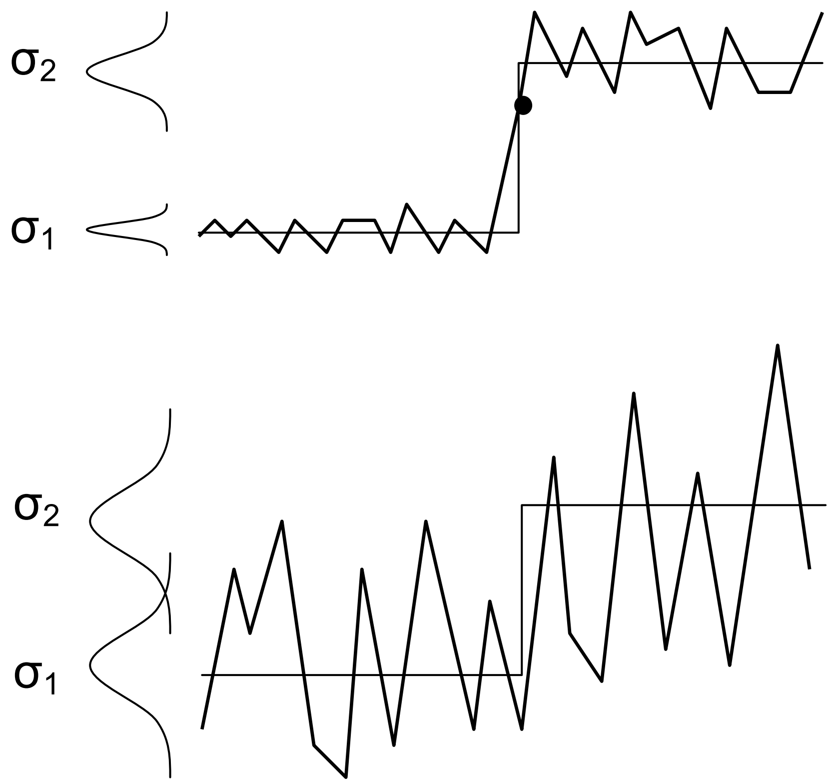


This figure shows an ideal occultation event with no noise. I have scaled the plot to a standard form that shows the intensity switching from 0 to 1, and with unit time intervals. Any occultation trace can be reduced to this form by appropriate scaling. The corresponding measured intensity plot then looks like:



Note that there are three distinct measurements: 0, 1 and a single point of intermediate intensity at the transition event. Since the transition went high early in the frame, the intensity is close to 1. The height of that transition intensity is key to interpolating the exact time of the event.

In the presence of noise, the situation is less clear:



The upper figure shows a transition where the noise increases with intensity, so there is less noise at 0 than at 1. In this case, the transition event is still well-defined to a frame, but the interpolation within the frame has lost accuracy due mainly to the noise in the intermediate intensity value. In the lower case the noise is equal, but large on the scale of the intensity change – making even the interval of the event difficult to determine. Clearly the noise in the signal, vertically, affects the timing accuracy of the event, horizontally.

The model in this paper has several assumptions:

1. The signal has been linearized
2. The behavior of noise over the intensity range is known
3. The timing of the intervals is accurate and there is minimal blanking

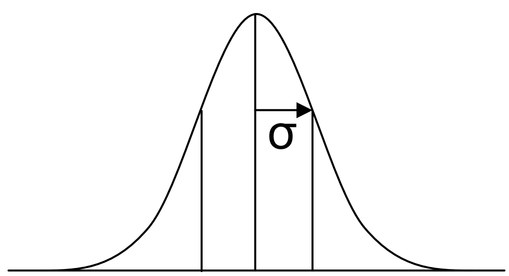
With these assumptions, one needs several more properties to determine the time of the event:

1. A measure of fit for a proposed event time to the actual observation
2. An estimate of the noise in the intermediate frame associated with the intermediate intensity
3. A long run of measurements before and after the event to determine the upper and lower intensities accurately on the scale of noise, and to flatten any slow modulations or drift.

Linearization is critical since the intermediate frame interpolation assumes linear accumulation of intensity over time. Typical video camera nonlinearities may correspond mostly to an imposed gamma response, but regardless of the nonlinearity, each camera can in principle be calibrated and linearized. After linearization, one can then characterize the noise as a function of intensity in a controlled manner. The noise may have an arbitrary functional form that should be included in the model, but in this work it is assumed to be Gaussian. The noise in the intermediate value is the sum in quadrature of the corresponding noise contributed by the high and low fractions of the event interval.

With the data pre-processed appropriately, the goal of occultation timing is to determine the most likely time of the event, along with an estimate of the timing accuracy. Determining the time amounts to fitting a model of the event to the data and finding the parameterized model that fits with least error. There are many ways to do this and some may be better than others in different circumstances. In this work, I do a least-squares fit of the proposed true intensity plot to the observed data, and find the best fit by brute force interval search. I use a brute force search because the search landscape is “rough” due to many local minima, and the computational cost is acceptable. I tried weighting the error based on the known noise levels, but this did not help. I also tried a chi-square approach, but it also was not as good as a non-weighted minimization of sum-squared error. One could weight the fit in arbitrary ways such as to emphasize the fit near the assumed transition point – but I did not find any improvement with such approaches.

In all this work, noise and errors are based strictly on the standard deviation of the values, which as a reminder, looks like:

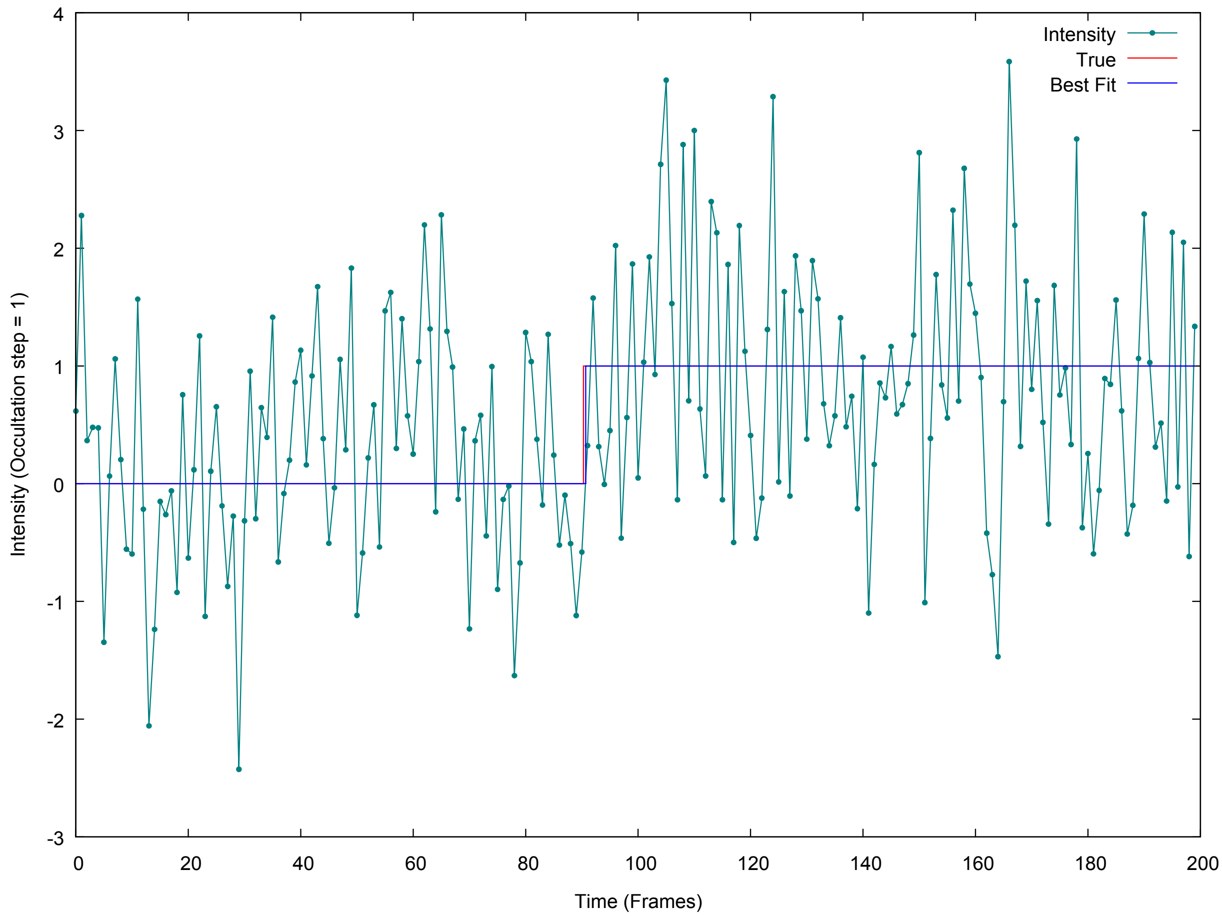


This shows that, both to characterize noise and to quantify error in the estimate, a value of +/-σ will capture 68% of the values, and is not meant as a much higher confidence interval. In fact, the true value is expected to fall outside of this range 32% of the time. It is important to state accuracy in terms of standard deviation so that derivative values, such as a fit of an asteroid profile to the measurements, can consistently weigh the observations according to their accuracies in a quantitative and statistical manner.

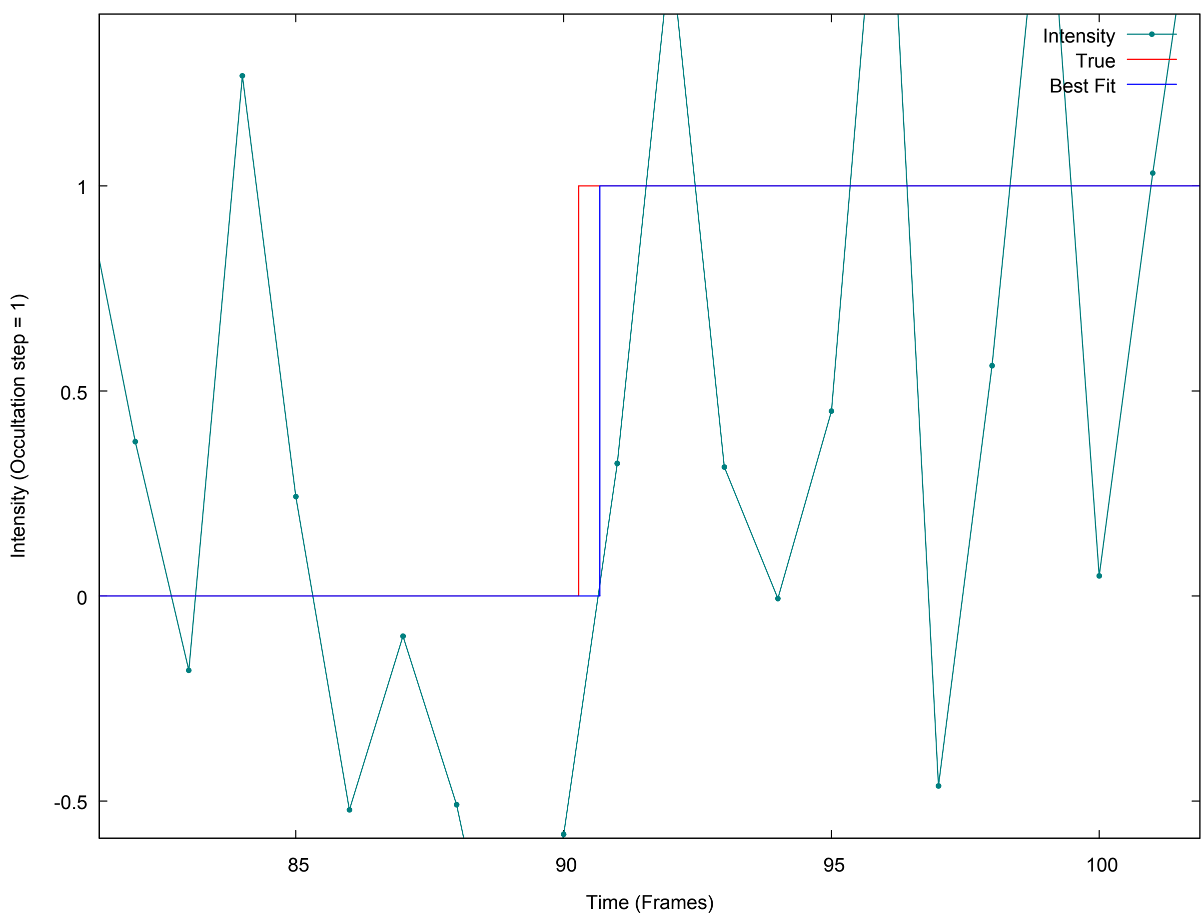
In general, an occultation will have one amount of noise at the lower intensity and a different amount at the higher intensity, with some intermediate values in between. To simplify the problem I focus on two realistic scenarios: equal noise at each level, and negligible noise at the lower level.

To estimate the error for a given amount of noise, I do repeated simulations and fits based on random occultation times. Since I know exactly when the occultation happens I can apply the noise model and interpolation scheme to generate a simulated intensity trace with imposed noise. I then, without knowledge of the true event time, find the value that minimizes the squared error. I repeat this many times and calculate the standard deviation of the found value from the true value, which corresponds to the accuracy of the fit. Finally I use bootstrap statistics on the ensemble of results to estimate the standard error of the accuracy, shown as error bars in the resulting plots.

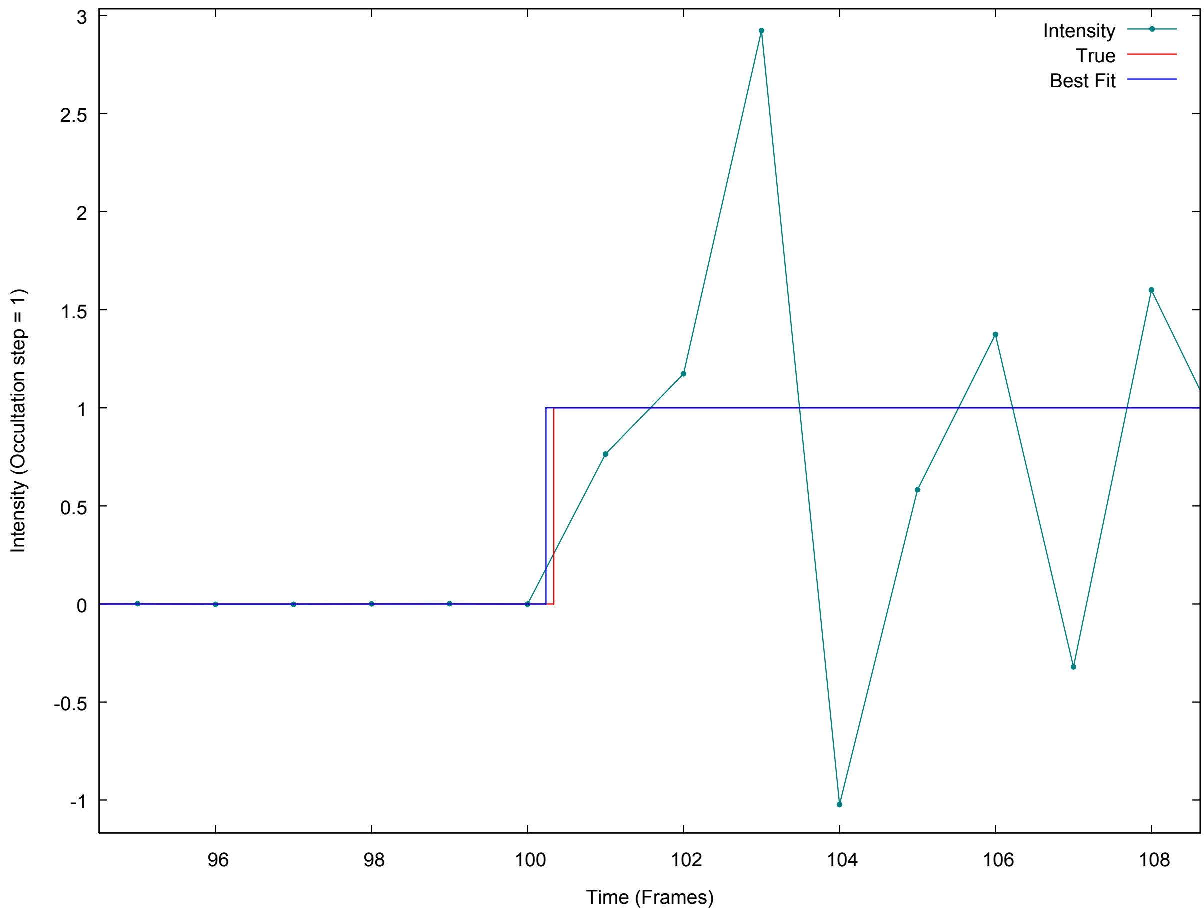
To see what these traces look like at noise levels of 0.001 and 1.0, I show plots of simulated traces with a known transition time, along with the best fit solution found.



The above figure shows the true curve in red, with a transition near frame 90. This is shown with simulated noise with $\sigma_1 = \sigma_1 = 1.0$. Note that since the noise corresponds to the standard deviation around 0 or 1, rather than “peak to valley,” the actual intensity ranges all the way from -2.5 to 3.5. Although the noise buries the signal, the best fit solution, shown in blue, is very close to the true curve.

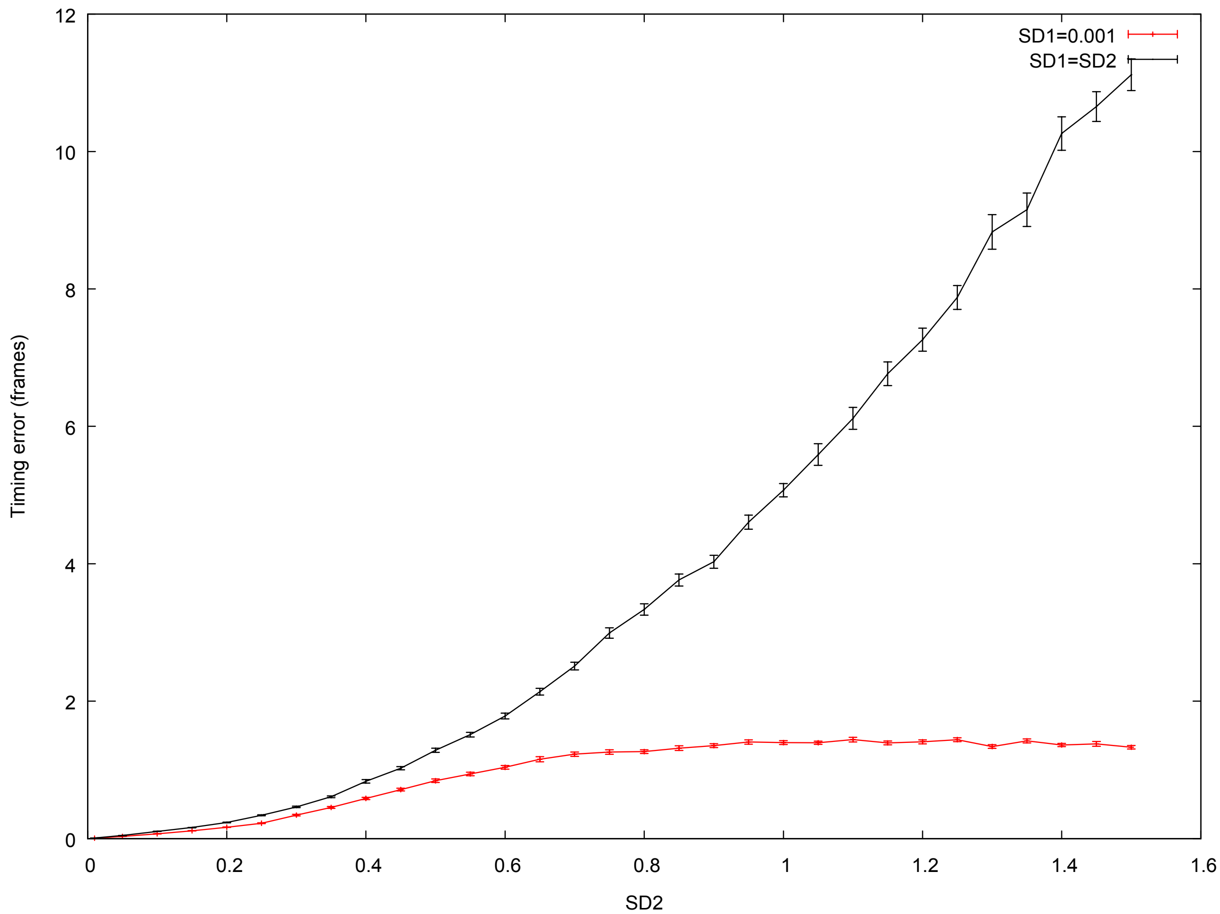


This is a close-up of the above curve, showing the sub-frame accuracy of this fit – despite the noise.

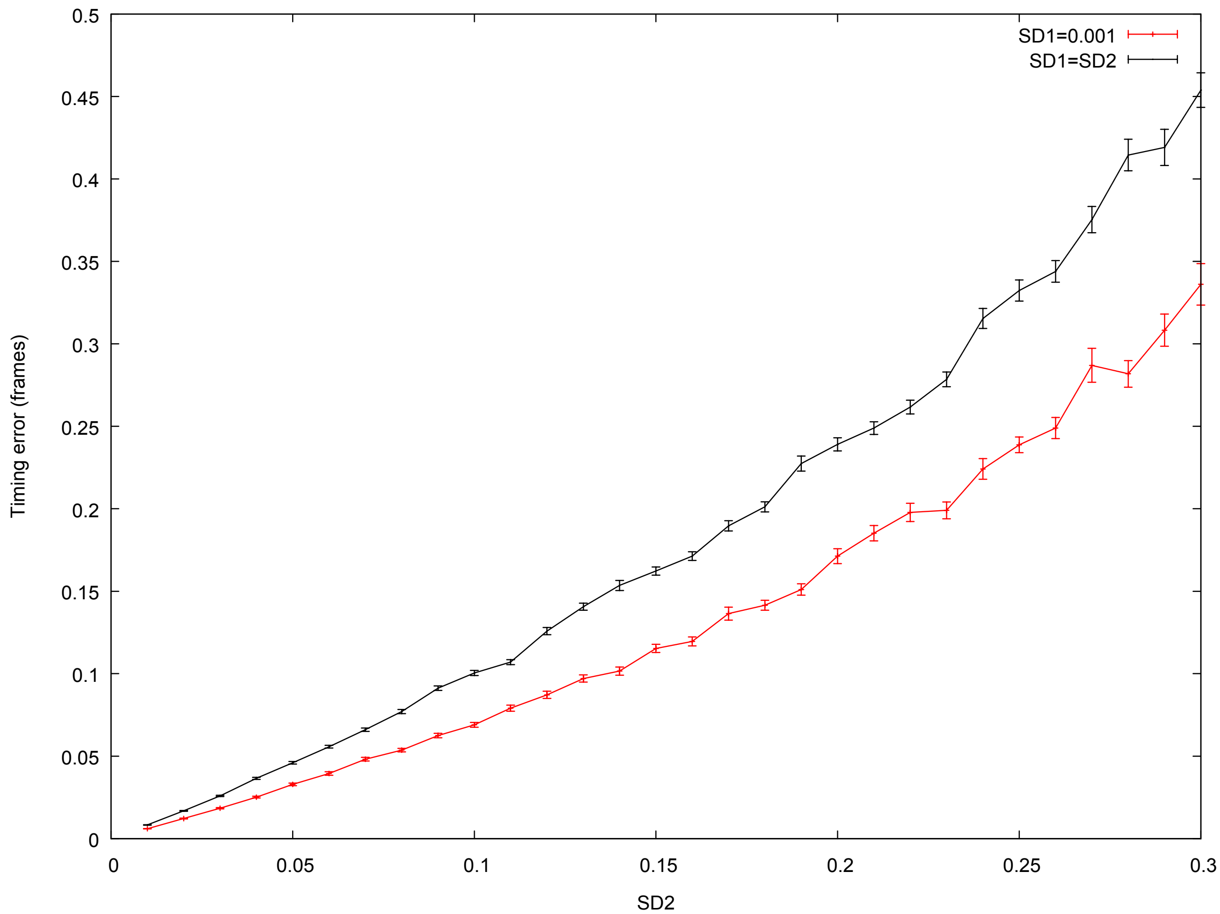


Above shows a curve with σ_1=0.001, and σ_2=1.0. In this case the small noise on the left helps pin the location of the transition, and the accuracy is much greater despite the noise in the upper intensity.

Here are the results of several simulations. For each point in the plot, and associated error bars, I perform 5000 simulation trials using an intensity trace of 200 points. I randomly set the transition time to a point near the center of the trace, and then find the best fit based on the given noise settings. The results show interesting and different behavior at small and intermediate values of noise, so two plots are shown below.



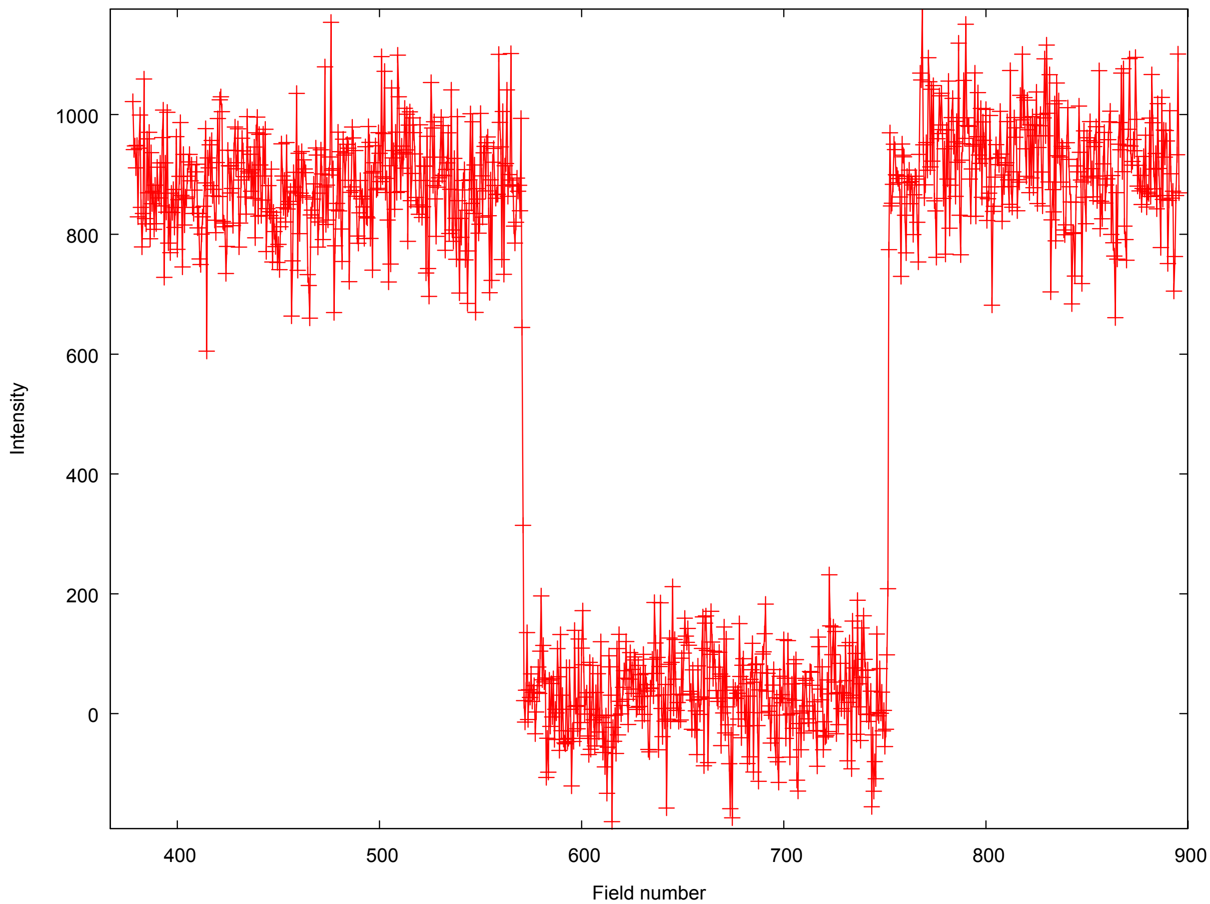
This plot shows two sets of simulations, where $\sigma_1 = \sigma_2$ (upper) and $\sigma_1 = 0.001$ (lower). Both plots show a clear increase from sub-frame timing error at small noise, but the upper curve continues to increase quadratically while the lower curve plateaus near an error of about 1.3 frames. This means that if the upper intensity has large noise equal to the the lower intensity, the timing error increases quadratically with the noise. But if the lower intensity is much less noisy than the upper, the timing error reaches a maximum of about 1.3 frames due to the "pinning" of the event on the left, low noise, side of the trace.



This figure is the same as above, but in the very low noise region, showing that for small noise the timing error increases linearly, and for $\sigma_1 = \sigma_2$ (upper plot) the timing error is nearly equal to the noise. As the noise decreases in the lower level (σ_1) the timing error approaches that of the lower plot. The linearity makes sense here because the interpolated time depends so heavily on the intensity of the transition point, and any vertical shift in that point correspondingly shifts the time of the event, when the noise is small. For greater noise, the behavior of the upper and lower curves plays more of a role in the best fitting time.

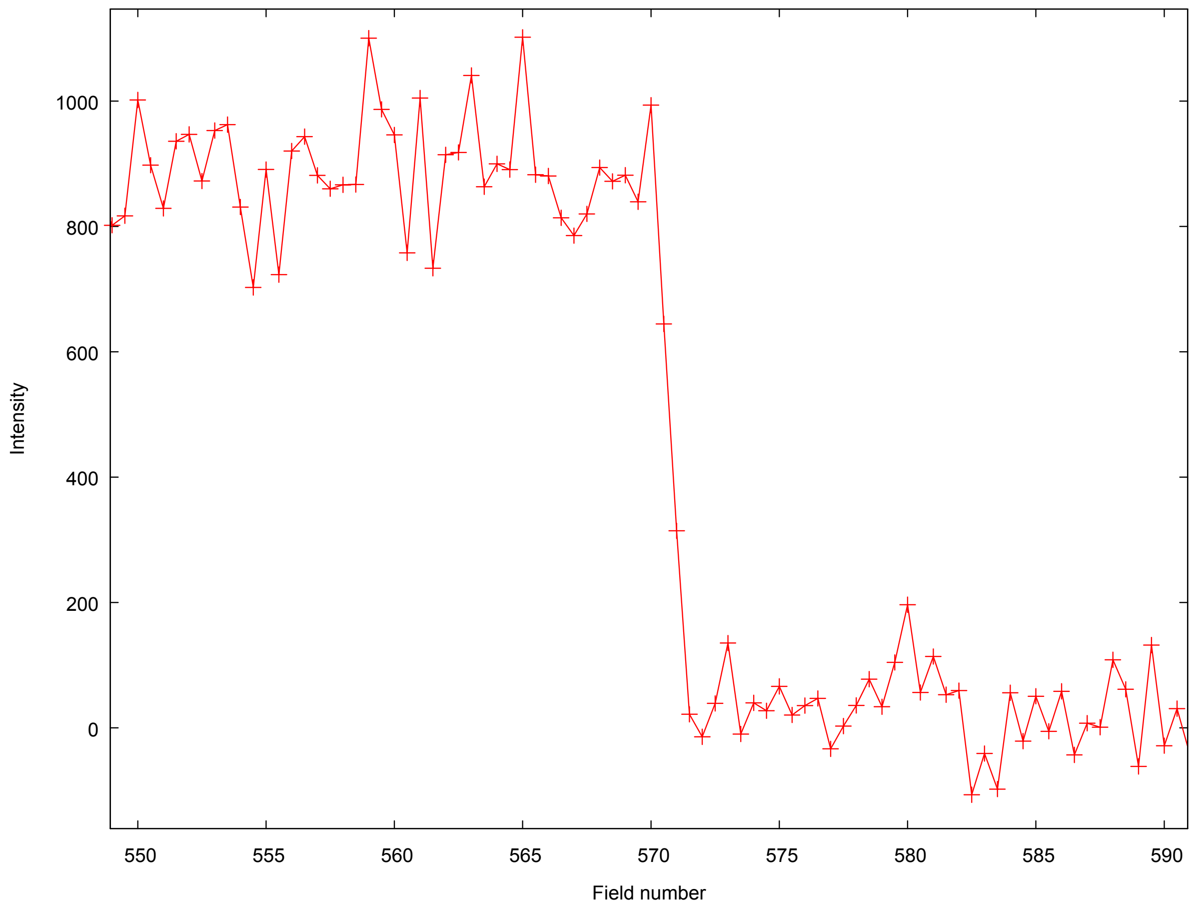
Note that this study is based on the noise values as standard deviations around a transition that has been scaled from 0 to 1. The timing error is naturally related to this noise and less easily expressed in terms of signal to noise ratio, S/N. But if the "signal" is the occultation intensity change, then the S/N is simply $1/\sigma$. For the small noise range shown on the left of the above plot, this S/N value would be very large.

To put the above simulations in perspective I now discuss a recent measurement of an occultation of a 9.9 Mag (R) star by the asteroid Henrietta. I measured with an 11-inch SCT at f/6.3 and a PC164-ex2 video camera. The star was high up at nearly 50 degrees altitude, which combined with the large aperture and clear skies yielded the high S/N trace shown below.

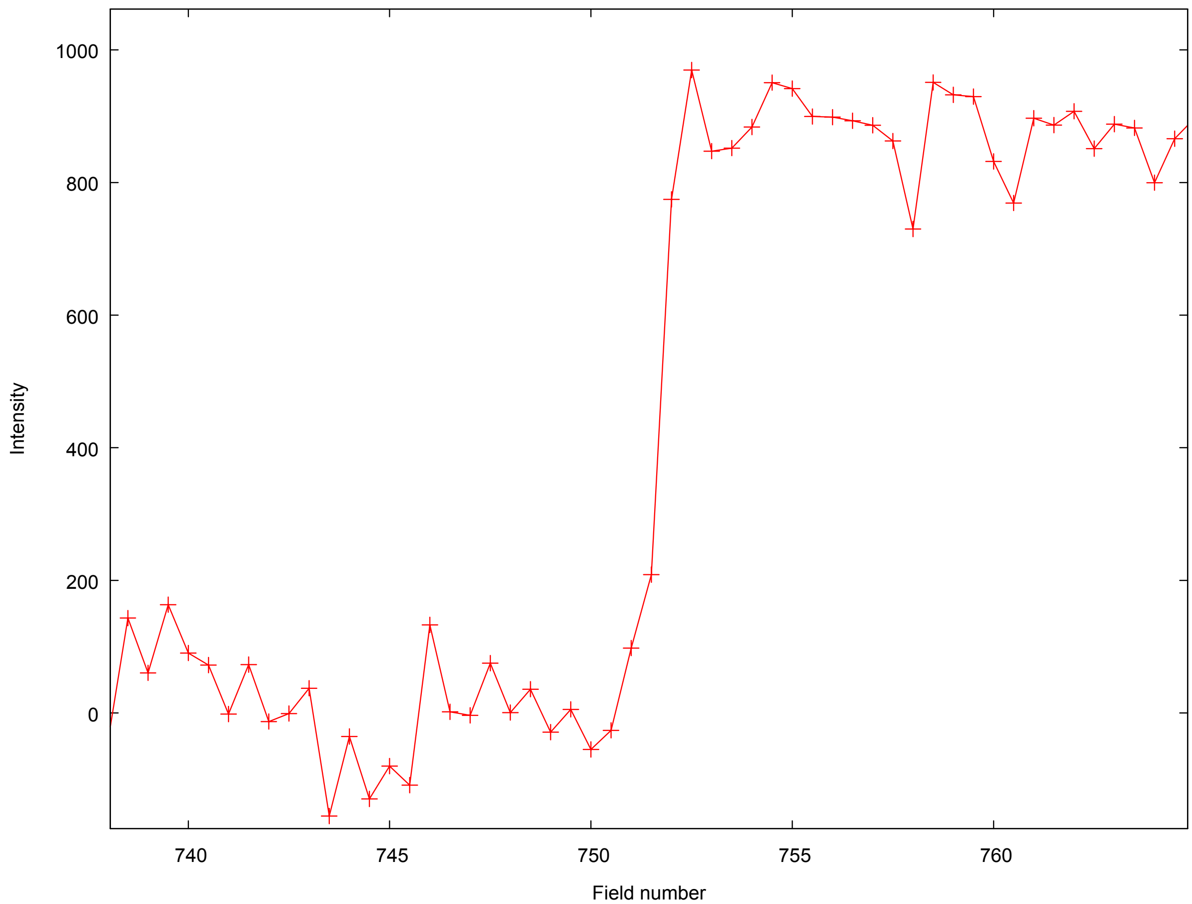


This figure is from an LiMovie analysis of the resulting AVI file, showing the star intensity after background subtract. I slaved the measurement aperture to a nearby star to track the intensity carefully even during the transition frames. The plot shows timings for each field, or at 17 ms intervals. The Disappearance (D) shows two possible transition times, while the Reappearance (R) shows one candidate. On the D side, the upper transition point is within the noise, while the lower one is outside it. In fact, with Gaussian noise there is no way to know which one is a “true” transition – one can only assign probabilities to a given interpretation.

Here are close ups of the dis- and reappearances:



1000
800
600
400
200
0
Intensity
550
555
560
565
570
575
580
585
590
Field number

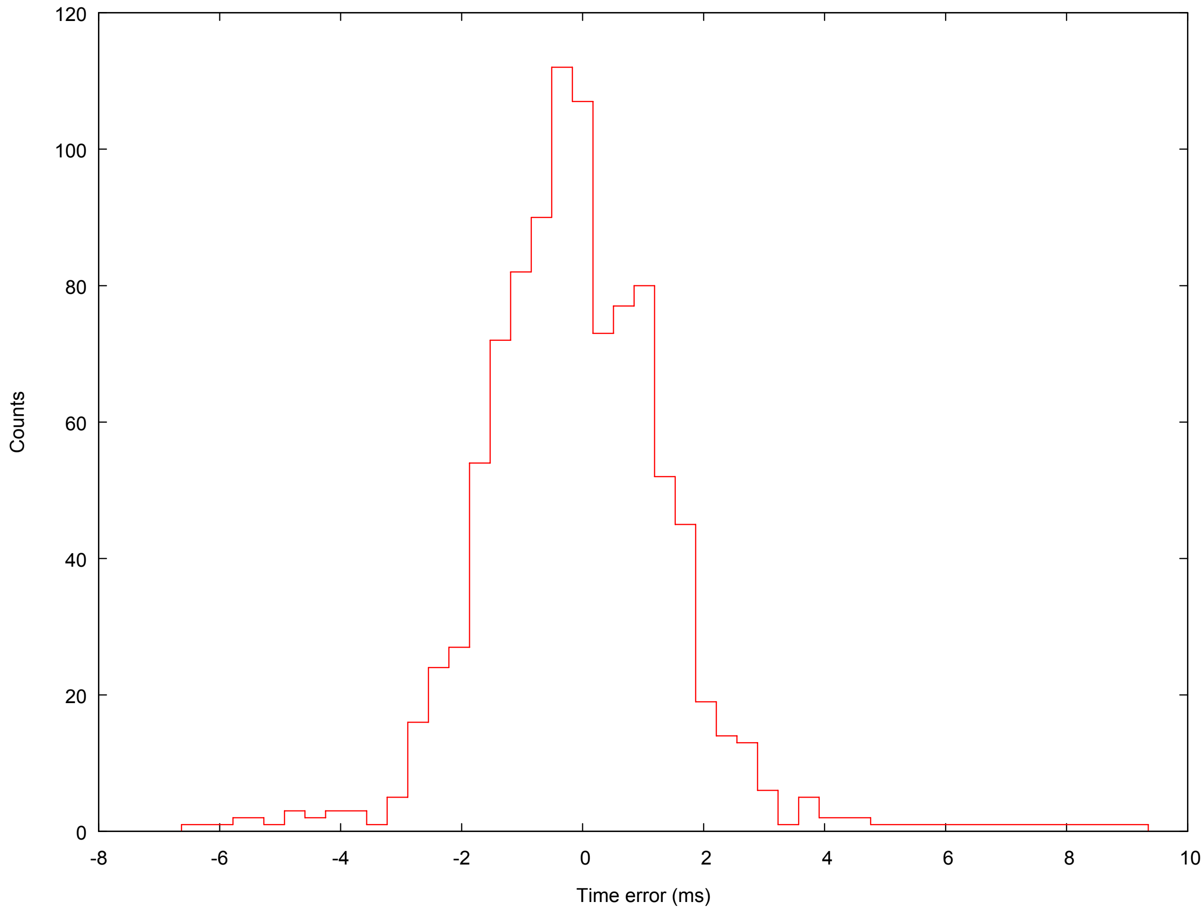


When doing least squares fits to these traces, there is no need to interpret the curves at all. One simply assigns values to the times of D and R, calculates putative transition frames, and finds the corresponding error relative to the true case. One then uses the values of D and R that minimize this error.

How does one determine the accuracy of the fit, and what does that accuracy mean? One way to make this clear is to imagine 1000 occultationists observing the same occultation at the same location with the same equipment and the same noise levels – and each then does a least squares fit to find values of D and R. If you know the true values of D and R, you can calculate the spread of error over all 1000 observers and determine the standard deviation of the true error. That would give the accuracy of each person's individual measurement. Some people will be very near the true values, while others might be several frames away. But the standard deviation of the error will be well defined and directly interpretable. Note that this definition of accuracy does not depend at all on the actual trace someone got, or how well it appeared to show transition events. It only assumes the same level of Gaussian noise for each observer.

Since I don't have 1000 observers to query, I instead do 1000 simulations with random event times – but using the exact same noise characteristics as the measurement. Since I know the exact time of each simulation I can directly measure the standard deviation, hence accuracy, of the actual trace.

My Henrietta measurement, when normalized to a range of 0 to 1 in intensity, has SD1=0.084 and SD2= 0.101. When I plug that those parameters into a simulation of 1000 independent traces, I get the following distribution of errors in a single event:



When 1000 observers record the occultation with these noise parameters, and then do a least squares fit to the transition times, the spread of reported errors will match the histogram above. Most will be within 2 ms of the exact time, but some could be as far as 9 ms off. The single metric to describe this spread is the standard deviation of the errors, which in this case is 1.46 ms. 68%, or 680 of the observers will be within 1.46 ms of the correct time – but 32% will be beyond it. Stating the error as a standard deviation describes the spread in a well-defined way that allows direct statistical interpretation.

To help analyze traces such as these, which show a pretty clear disappearance and subsequent reappearance, I wrote a utility that loads the LiMovie file, searches for the occultation, estimates the noise levels, and then does simulations to determine the accuracy. The entire utility operates without a user interface or any user input or judgement. For the Henrietta occultation it identified the time of the D and R and gave accuracy values of +/- 1.46 ms based on simulations at the same levels of noise. This standard deviation is much less than a field time of 17 ms, and likely approaches limitations of the accuracy of the KIW time insertion, which would need to be determined empirically and are at least +/- 0.5 ms just due to the discreteness of the time measurement.

Putting it all together, an observer with inexpensive equipment can measure asteroid occultation times to an accuracy limited only by the noise in the measurement. For an 11" telescope and a Mag 10 star, this accuracy is shown to be on the order of 1.5 ms. What good is this accuracy? For an asteroid moving with a skyplane velocity of 10 km/s, each millisecond is only 10 m of height on the asteroid's profile. As a result, careful timing with good statistical analysis and fitting of the event could yield detail of the asteroid profile, much as grazing lunar occultations do for the moon. In this case, the analysis was done with a black-box piece of software that required no user input that might skew or allow misinterpretation of the results and corresponding error bars.