Non-Parametric Photometric Identification of Rising SN Light Curves in the pSNid II Package

B. Connolly[1] and N. Connolly[2],

1 Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, PA 19104-6396, USA,
brianco@sas.upenn.edu
2 Physics Department, Hamilton College, Clinton, NY 13323, USA

Abstract

pSNid II is a public software package used to identify and type (rising) supernova (SN) light curves. Until recently, typing and identification of SNe have been dependent on light curve models or templates. However, to produce an unbiased SN sample, one may want to simply select light curves that are on the rise without regard to the identity of the underlying astronomical object. We present a method that achieves this by computing a Bayes factor that quantifies how much more probable it is that a measured light curve is derived from one which is rising as opposed to one that is not. The method relies upon the following assumptions: (1) there is a set of "pre-explosion" SN fluxes that fluctuate about zero, (2) at the SN explosion there is a statistically significant deviation from zero flux, and (3) a rising light curve is a light curve in which the mean flux for some measurement is larger than the mean flux in the previous measurement. The technique has been validated using simulated data.

1 Motivation

Type Ia supernovae (SNe Ia) have become a staple for experimental cosmology. Identifying SNe Ia among supernovae of other types is traditionally done using spectroscopy, typically performed when the SN is at its maximum brightness. However, large upcoming supernova surveys (LSST, DES, Pan-STARRS) will most likely not have sufficient resources to perform a spectroscopic follow-up of each of its thousands of SN candidates. Therefore, photometric identification and classification of SNe, along with an estimate of when a SN candidate has reached or will reach its maximum light, are becoming increasingly important.

Photometric studies of SNe are typically done by comparing the measured SN light curves with SN models or templates. However, as larger data samples appear to point to an increasing diversity of SNe (for a recent example see [1]) it is important to impose minimal selection criteria on the SN candidates in order to minimize selection biases in cosmology and studies of SN properties.

The Bayesian order-restricted literature is rich with methods that can be used to identify rising light curves. However, they either require integrations over an infinite parameter space (e.g., [4]) or require that the light curve fit within a limited number of shapes (e.g., [6]). To mitigate these limitations, we take a slightly different approach, directly calculating the posterior probability that the light curve begins to rise at some undetermined time vs. the hypothesis that it does not satisfy this rising constraint.

Specifically we parametrize the Bayes factor

$$B = \frac{P(D|\mathcal{H}_{rising})}{P(D|\overline{\mathcal{H}}_{rising})}, \qquad (1)$$

as the ratio of the probability of obtaining the photometric measurements, D, given that they are derived from a rising light curve to the probability of obtaining the same measurements given that they are derived from a non-rising light curve. The hypotheses are defined so that

$\mathcal{H}_{rising}$: includes all light curves where the $(i-1)^{th}$ flux is less than the i^{th} flux

$\overline{\mathcal{H}}_{rising}$: includes all light curves not in $\mathcal{H}_{rising}$.

Note that only one mean flux needs to be less than the previous mean flux for the light curve to be considered non-rising. Further, no binning in time or flux is required, as the time between flux measurements is not factored into our calculation.

2 Method

General Idea

In order to decide whether or not a light curve is rising the problem must be parametrized in such a way as to render a non-trivial Bayes factor. The sets of rising and non-rising light curves are different, but infinite. Therefore, one must be careful to parametrize the problem in such a way as to have priors that properly cancel.

This is done by effectively marginalizing over the magnitudes of the differences between the mean fluxes of the i^{th} measurement and the $(i-1)^{th}$ measurements for both the $\mathcal{H}_{rising}$ and $\overline{\mathcal{H}}_{rising}$ hypotheses, and then calculating the ratio of the probability that the mean differences are positive to the probability that they are not through the Bayes factor.

Details

If μ_i is the hypothesized mean of the i^{th} flux, then it can be parametrized as follows:

$$\mu_i = \left(\sum_{j=1}^{i} (-1+2a_j) f_i \right) b \qquad (2)$$

where $a_j \in \{0,1\}$, $f_i = [0,1]$ and $b = [0,\infty)$.

We then trivially define the "rising" hypothesis, $\mathcal{H}_{rising}$, and its alternative, $\overline{\mathcal{H}}_{rising}$, solely in terms of the restrictions on a_i. That is, roughly speaking,

$$\mathcal{H}_{rising} \equiv a_i = 1 \qquad (3)$$

and

$$\overline{\mathcal{H}}_{rising} \equiv \text{ at least one } a_i = 0 \qquad (4)$$

Figure 1 illustrates the meaning of f_i, a_i and b.

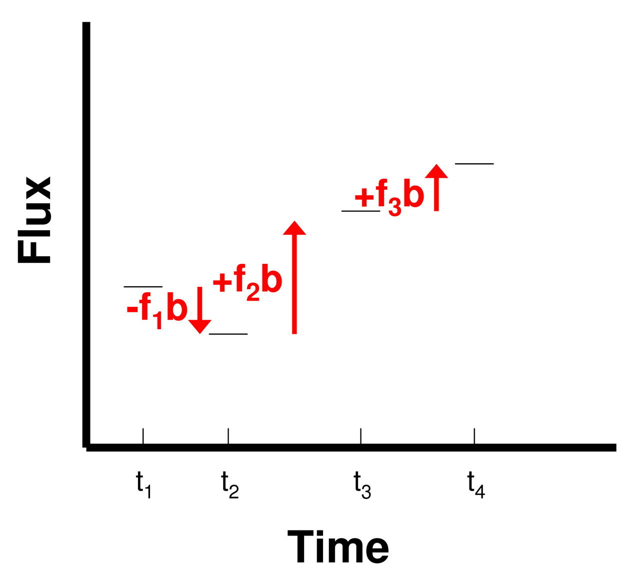

Fig. 1: Example of a non-rising light curve hypothesis where each flux is parametrized by $(-1+a_i)f_ib$. The values of a_i in this case are $a_1 = 0$, $a_2 = 1$, and $a_3 = 1$. Note that the measurements do not need to be made at regular intervals. The data are not shown on this diagram.

The Bayes factor is then expanded in terms of θ and θ_a defined as:

$$\theta \equiv \{\vec{f}, b, n_{pre}\} \qquad (5)$$

and

$$\theta_a \equiv \{\vec{a}, \vec{f}, b, n_{pre}\} \qquad (6)$$

where $\vec{a} \equiv \{a_i\}$, $\vec{f} \equiv \{f_i\}$, and n_{pre} is the number of pre-explosion flux measurements. The Bayes factor becomes:

$$B_{10} = \frac{\int d\theta P(\vec{D}|\theta, \mathcal{H}_{rising}) P(\theta|\mathcal{H}_{rising})}{\int d\theta_a P(\vec{D}|\theta_a, \overline{\mathcal{H}}_{rising}) P(\theta_a|\overline{\mathcal{H}}_{rising})} \qquad (7)$$

In practice, the non-rising $\overline{\mathcal{H}}_{rising}$ hypothesis would require the summation of $2^N - 1$ possible choices for $\{a_i\}$. However, using Monte Carlo integration techniques (e.g., the MISER algorithm [8, 9]) the summation over a_i quickly converges to the correct result with ~ 1000's of iterations.

3 pSNid II Package

pSNid II (http://web-facstaff.sas.upenn.edu/ brianco/psnid) is a software package designed to photometrically identify supernovae both parametrically and non-parametrically, assuming the SN candidate

- fits within a known set of SN templates (see [2, 5, 7, 10]).
- is either one of a known set of SN templates or does not fit within any known model. Note that the SN can be classified using color only, allowing for an elegant formulation of the "anything else" hypothesis (see [3]) that does not require that the SN candidate be one of a finite set of known astronomical objects.
- is simply consistent with a rising light curve.

The package is user-friendly and flexible, allowing for a variety of inputs and interchangeable supernova templates.

4 Results

Figure 2 shows the log of the Bayes factor for simulated rising light curves (red) and non-rising light curves (blue) using 5% Gaussian uncertainties for 14 flux measurements with random explosion times. **It is apparent that the Bayes factor provides good discrimination between the two hypotheses, showing the statistic performs as advertised.**

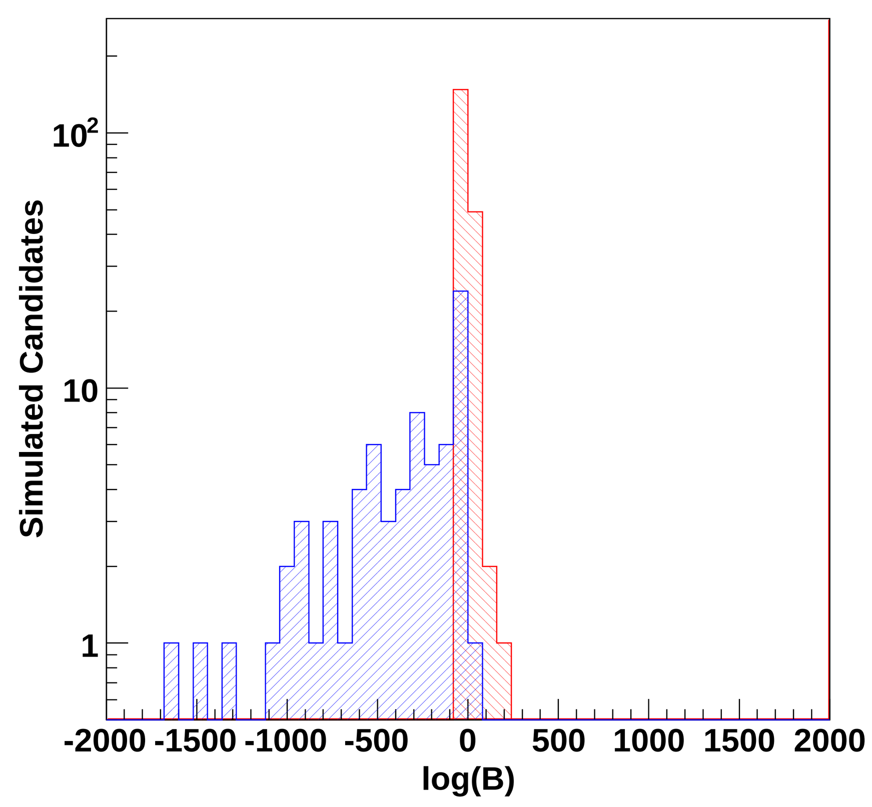

Fig. 2: Distribution of $\log(B)$ for simulated rising (red) and non-rising (blue) candidates with random explosion times (n_{pre})

Future studies will concentrate on quantifying the sensitivity of the Bayes factor with real data, increasing the speed of the calculation, incorporating multiple filters for the photometric measurements, and including information on the supernova magnitudes.

References

[1] Barbary, K., et al., ApJ, 690, 2, 1358 (2009)

[2] B. J. Barris and J. L. Tonry, ApJ 613, L21 (2004).

[3] N. Connolly and B. Connolly, "A Bayesian Approach to Classifying Supernovae With Color," arXiv:0909.3652.

[4] Hoijtink, H., Klugkist, I., & Boelen, P. A. (Eds.). (2008). Bayesian evaluation of informative hypotheses. New York: Springer, Chapter 12.

[5] Kuznetsova, N. & Connolly, B., 2007, ApJ, 659, 530

[6] K. Nashimotoa and F.T. Wright, Computational Statistics and Data Analysis Volume 52, Issue 12, 15 August 2008, Pages 5143-5153

[7] Poznanski, D., Maoz, D., and Gal-Yam, A., 2007, AJ, 134, 1285

[8] W. H. Press and G. R. Farrar. Recursive stratied sampling for multidimensional Monte Carlo integration. Computers in Physics, 4:190-195, 1990.

[9] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. Numerical Recipes in C. Cambridge University Press, second edition, 1992.

[10] Masao Sako *et al.* Publication in progress.