Astronomy & Astrophysics manuscript no. usnac_snana
March 29, 2022

Implementation of the redshift evolution of type Ia supernovae in SNANA

N. Nicolas⋆[1], M. Rigault⋆⋆[1], M. Smith[1,2], D. Scolnic[3], B. Popovic[3], G. Aldering[4], M. Briday[1], Y. Copin[1], J. Lezmy[1], J. Nordin[5], and Saul Perlmutter[4]

[1] Univ Lyon, Univ Claude Bernard Lyon 1, CNRS, IP2I Lyon / IN2P3, IMR 5822, F-69622, Villeurbanne, France
[2] University of Southampton, Southampton, UK
[3] Duke University, FILL IN AFFILIATION
[4] Physics Division, Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA, 94720, USA
[5] Institut fur Physik, Humboldt-Universität zu Berlin, Newtonstr. 15, 12489 Berlin, Germany

Submitted to A&A the 19th of May 2020 / Accepted 26 February 2021

ABSTRACT

The use of Type Ia supernovae (SNe Ia) as standard candles lead to a huge improvement on the constraints on cosmological parameters. In order to continue refining these values in a era of increasingly bigger datasets, systematic uncertainties have to be dealt with. Following previous work on the link between supernovae and their environment, we aim at finding the precise impact of a redshift-evolving underlying stretch distribution of SNe Ia compared to the currently-used distributions in cosmological tool SNANA. We reproduced previous studies of the BBC team, then modified the underlying stretch distribution and implemented an age step that logically follows the use of an age-evolving population. We find that WHAT

Key words. Cosmology – Type Ia Supernova – Systematic uncertainties

1. Introduction

General talk about SNe Ia, Hubble diagram, constraints and assumptions from BBC to fit HD.

Standardizing SNe Ia led to the discovery of the acceleration of the Universe's expansion (**Riess98, Perlmutter99**): thanks to their color (c) and stretch (x_1) parameters, the systematic uncertainty on their absolute magnitude and thus their distance were greatly reduced. Dark energy is commonly thought to be the cause of this expansion, and its equation-of-state parameter, w, is currently know down to $\approx 4\%$ by combining constraints from both SNe Ia and the Cosmic Microwave Background (**Scolnic18, Jones18, Brout19b**). With the current effort to lower statistical uncertainties, the systematic uncertainties are gaining in importance in the total error budget and need to be addressed.

With that goal in mind, some recent studies try to improve the known correlations. In **NR21**, we presented an initial study of the drift of the underlying SNe Ia stretch distribution as the function of redshift. We based this study on the assumption of two populations of SNe Ia, young and old, and implemented different modelings of their distributions to better represent the observed data. For that end we took SNe Ia data and created a complete sample by applying redshift cuts corresponding to the expected distance where no selection effects should impact each survey. This approach was at the time the simplest one to implement but does not allow to ponder the impact such a modeling has on the cosmology, and mainly on w, and skips all selection effects altogether.

Other studies tend to define and refine other correlations between a SN's brightness and some observable properties, such as their host-galaxy's mass (**Sullivan10, Betoule14, Smith20**) or Star Formation Rate (**Uddin17, Kim18, Rigault20**). One of the most common ways to test these correlations is to use simulations (**Kessler09a, Scolnic18, Brout19**). The SuperNova ANAlysis package (SNANA, **Kessler09b**) allows to simulate surveys by using underlying distributions of parameters and correlations between them with the aim to find which ones come closest to the actual data. **Scolnic & Kessler 16** describes the procedure of finding these distributions, implementing accurate selection effects for each survey, noise and intrinsic scatter that shift them to the observed distributions. In this framework, it's a galaxy's stellar mass ($M_{\rm stellar}$) that is assumed as the main observable responsible for its environmental effects on SNe Ia. The goal is to correct the values of SNe affected by Malmquist bias by computing the average of the simulated ones, as opposed to our first approach of getting rid of the affected SNe.

These simulated surveys have been used to study the dependence with redshift of the bias correction (**Kessler 09a, Betoule14**), but showed remnant correlations in the computed distance moduli (SK16). The analysis thus went from 1-dimensional to a 5-dimensional parameter space, studying the correlations with redshift, color, stretch, color-luminosity and stretch-luminosity relationship parameters (α, β) as developed by **Kessler & Scolnic 17** with the BEAMS with Bias Corrections (BBC) method, hereafter refered to as BBC5D. This method is used in both the Dark Energy Survey 3 years (DES3YR, **Abbott19, Brout19b**) and Panoramic Survey Telescope and Rapid Response System (PAN-STARRS, **Scolnic18, Jones18**) analyses. This formalism was further improved based on the work of **Smith20**

⋆ n.nicolas@ip2i.in2p3.fr, equal contribution
⋆⋆ m.rigault@ip2i.in2p3.fr, equal contribution

2. Modeling

In our first analysis, we based our work on a two-populations modeling based on age: a young (log(LsSFR) ≥ -10.82) and an old one (log(LsSFR) < -10.82). We determined the underlying stretch distribution for both these populations, and used the evolution of the fraction of young SNe Ia given by

$$\delta(z) = \left(K^{-1} \times (1+z)^{-2.8} + 1\right)^{-1} \quad (1)$$

we determined a relationship between SNe Ia stretches and redshift. Supposing age is the driving phenomenon behind the different systematics seen in SNe Ia cosmology also implies that SNe Ia have an age step of 0.130 mag rather than a mass step of 0.050 mag. In this work we are trying to generate simulations to ponder the impact of these systematics on the determination of cosmological parameters, and notably w.

In order to simulate SNe, we require a host galaxy to follow the distributions of what has been observed by the different simulated surveys. While we argue that LsSFR is a better tracer of a SN's environment (**Briday 21**), most survey characterize galaxies using their stellar mass. Therefore, to compare the implication of our modeling based on LsSFR with what other surveys observed, we need to modelize galaxy masses with respect to LsSFR using the same sample as described in Section 2 of **NR21**.

2.1. Modeling the mass

Following **NR21**, we use the LsSFR as the tracer of the age of a SN on the mass estimates from SNf, then model the young and old population through a series of different parameterizations and pick the lowest AIC one. However, SNf masses were computed using Eq. 8 of **Taylor 2011** (see **Rigault 20**) while other surveys from the Pantheon catalog use different techniques of mass estimation that might give different output values for a same galaxy. The estimate from Taylor involves the absolute i-band AB-magnitude M_i. It is deduced from the apparent magnitude m_i knowing the galaxy's redshift but assumes that the observed i band is close to the restframe one, which is true for the SNf redshifts which are below z 0.05. Surveys from the Pantheon sample are at higher redshifts and used SED to avoid K-corrections in that procedure. In order to maintain coherence between the mass modeling based on SNf data and the masses measured in the Pantheon surveys we will simulate, we needed to use the same method for each object. We thus chose to use SED fitting for everyone.

A few words about SED fitting

Based on the shape of the M_{host} vs LsSFR scatter plot, different modelings were implemented. There are referred to by the number of Gaussian functions, Mean values and Sigma values that compose their mathematical behavior. For instance, a modeling having 1 symmetric Gaussian with different means for each population is labeled 2G2M2S. A total of 8 modelings have been tested for this study:

- 1G1M1S also simply named *Gaussian*, a pure redshift-independent and age-independent Gaussian modeling;
- 1G1M2S or *Asymmetric*, using a unique asymmetric Gaussian;
- 2G2M2S or *Howell*, based on the work from **HOWELL07** and previously described;
- 2G2M3S with one Gaussian for the young population and one asymmetric Gaussian for the old one;

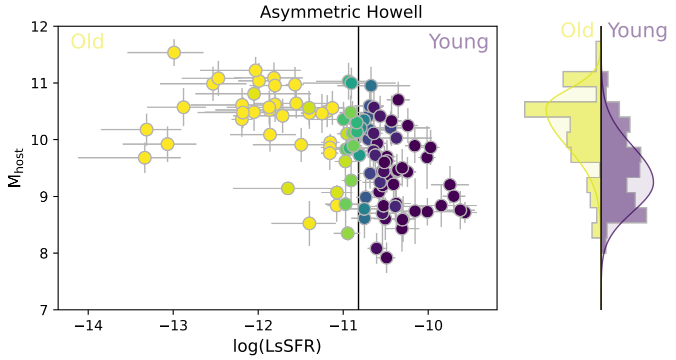

Fig. 1. *Main*: SED-fitted host mass (M_{host}) as a function of the LsSFR for SNfactory SNe. The color corresponds to the probability, p_y, for the SNe Ia to be young, i.e., to have log LsSFR ≥ -10.82 (see **?**). *Right*: p_y-weighted histogram of the SN masses, as well as the adjusted model; contributions of the younger and older population are shown in purple and yellow, respectively.

- 2G2M4S or *Asymmetric Howell* with one asymmetric Gaussian per population;
- 3G3M3S or *Free Base*, close to the base model of **NR21** with a single Gaussian for the young population and a Gaussian mixture for the old population, each having a different mean;
- 3G3M4S using Free Base with a mixture of Gaussian and asymmetric Gaussian;
- 4G4M4S or *Double Howell*, having one Gaussian mixture for each population.

2.2. Mass modeling results

Each one of these modelings is adjusted on the whole Pantheon dataset, following the procedure in Section 3 of **NR21** depending on the presence of LsSFR measurments in each sub-survey. The results are gathered in Table 1 where

$$-2\ln(L) = -2\sum_i \ln\mathcal{P}\left(x_1^i \mid \boldsymbol{\theta}; \mathrm{d}x_1^i, y^i\right). \quad (2)$$

and

$$\mathrm{AIC} = -2\ln(L) + 2k, \quad (3)$$

Due to unrealistically small errors on the last three modelings thus holding no physical meaning, we discarded them in the analysis regardless of their fitting quality. We used the Akaike Information Criterion (AIC, **BURNHAM01**) to compare each model's ability to properly describe the observations, penalizing additionnal degrees of freedom to avoid overfitting the data. After computation, the best, lowest-AIC model is the Asymmetric Howell modeling, in which the mass distribution of the younger population (log(LsSFR) ≥ -10.82) and the older population (log(LsSFR) < -10.82) are modeled as asymmetric Gaussians $\mathcal{N}\left(\mu, \sigma^2_{-\mathrm{y,o}} \text{ if } x_1 < \mu, \text{ else } \sigma^2_{+\mathrm{y,o}}\right)$. A graphical representation is shown Fig. 1. The parameters value we found are summarized Table 2.

3. SNANA

SNANA is a combination of tools to simulate SNe Ia data and exploit these to put constraints on cosmological parameters by fitting a Hubble diagram with them. There are 3 main parts to this procedure: (1) generate a lightcurve by modeling a SN and

Table 1. Comparison of the relative ability of each model to describe the data. For each considered model, we report whether the model is drifting, its number of free parameters, $-2\ln(L)$ (see Eq. 2), the AIC and the AIC difference (ΔAIC) between this model and the base model used as reference because it has the lowest AIC.

			Fiducial sample (569 SNe)			Conservative sample (422 SNe)		
Name	drift	k	$-2\ln(L)$	AIC	ΔAIC	$-2\ln(L)$	AIC	ΔAIC
Asym Howell	$\delta(z)$	6	1538.7	1550.7	–	1197.4	1209.4	–
Howell	$\delta(z)$	4	1546.6	1554.6	−4.0	1205.0	1213.0	−3.6
Asym+Howell	$\delta(z)$	5	1546.5	1556.5	−5.8	1204.8	1214.8	−5.4
Asymmetric	–	3	1593.1	1599.1	−48.5	1248.6	1254.6	−45.2
Gaussian	–	2	1608.3	1612.3	−61.6	1258.2	1262.2	−52.8

Table 2. Best-fit values of the parameters for the mass distribution model when applied to the SNfactory dataset only (114 SNe Ia).

Sample	μ_y	$\sigma_{-,y}$	$\sigma_{+,y}$	μ_o	$\sigma_{-,o}$	$\sigma_{+,o}$
SNFactory	9.34 ± 0.10	0.51 ± 0.07	0.95 ± 0.07	10.74 ± 0.48	0.48 ± 0.06	0.39 ± 0.06
Fiducial	9.34 ± 0.10	0.51 ± 0.07	0.95 ± 0.07	10.74 ± 0.48	0.48 ± 0.06	0.39 ± 0.06
Conservative	9.23 ± 0.10	0.47 ± 0.07	0.96 ± 0.07	10.61 ± 0.48	0.41 ± 0.06	0.44 ± 0.06

how each survey would acquire it, taking selection effects into account; (2) fit said lightcurve with `SALT2.4` and extract the m_b, x_1, c and t_0 parameters; (3) infer distance modulus values using bias correction.

These tools begin by generating a Spectral Energy Distribution to which are applied astrophysical effects such as lensing, dimming, redshifting or galactic extinction, giving a first simulated magnitude. Then this magnitude is converted into an observed one (flux) by reproducing the effects of the instruments: zero point, sky noise and PSF. Selection effects are finally applied to mimick the triggering of which events are analyzed; these include the needed number of detections and needed band(s) and the spectroscopic efficiency function of the simulated surveys. This leads to simulated lightcurves, and after a `SALT2.4` fit we can get to distance modulus values following the BBC framwork from **KS17**, which is defined as:

$$\mu = m_B + \alpha x_1 - \beta c - M_{z_i} + \delta\mu_{\mathrm{host}} + \delta\mu_{\mathrm{bias}} \tag{4}$$

3.1. Environmental impact and HOSTLIB

However, the definition of a realistic model is the be questioned. In the pioneer work of **SK16**, there were no relationship between SNe and their host galaxy. **P21** and **M20** introduced a link between the two thanks to a HOSTLIB: a table of 100,000 galaxies made to mimic the actual surveyed galaxies by the different samples. To each galaxy is associated a SN through its main fitted properties such as x_1 or c, which are generated by models of underlying distributions. Yet, this process was directed at guessing what relationship would fit more the data, by using bins of host galaxy masses and minimizing asymmetric Gaussian distributions for each in a backward-modeling way. It's a non-direct method to infer an evolution of an underlying distribution as a function of mass.

Our approach was to use independent data from the SNf sample that uses LsSFR to characterize a galaxy, as explained in Section 2, and make an evolving, analytical model that can then describe higher-redshift SNe in a forward-modeling approach. This method has the aim to be predictive and to better fit the data, as is discussed in Section 4. In order to implement our modeling in this framework, we had to modify the HOSTLIBs on which the simulations are based.

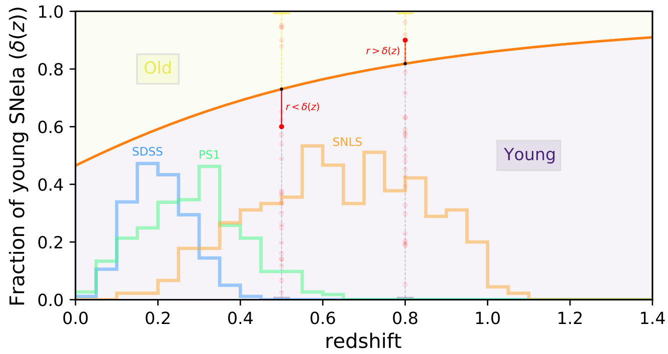

Fig. 2. *Orange*: Estimated fraction of young SNe as a function of redshift. *Histograms*: Number of SNe in bins of redshift for the 3 main surveys of the Pantheon sample, not scaled. *Red vertical lines*: For each z, a random number a between 0 and 1 is picked: if it's lower (higher) than $\delta(z)$ at that redshift then the SN will be young (old), and the distributions of mass and stretch to pick from will follow this flag.

3.2. Input generation

Having the two modelings for mass and stretch allows us to pick a redshift and generate a mass and a stretch value. This will in turn give us the ability to match the HOSTLIB's masses and replace the stretch values by our **N21** modeling's predictions. This is done with the *SNprop* Python module[1]. Given a redshift or lists of redshift, it takes the expected fraction of young stars using $\delta(z)$ from Eq. 1 then sets a "young" or "old" flag by picking a random value r between 0 and 1 and comparing it to said fraction. If $r < \delta(z)$, then the simulated SN will be young and conversely. The higher the z, the higher the $\delta(z)$, and thus the higher the probability to flag it young. A graphical explanation is given Fig. 2. Stretch and mass values are then generated with the previous modelings depending whether the progenitor is considered young or old.

[1] https://github.com/MickaelRigault/snprop

Table 3. Number of data in Pantheon and relative ratio with respect to the smallest sample, LOWZ

Survey	Number	Ratio/LOWZ
LOWZ	172	1.00
SDSS	335	1.95
PS1	279	1.62
SNLS	236	1.37

4. Results of modeling

In order to see the impact of our modeling in the inferred w value, we first needed to get on-par with the current best work on that matter, working closely with the SNANA team to recreate the results of **P21** before using our HOSTLIB.

4.1. Description of the tests

We defined 4 ways to simulate SNe depending on the assumptions that can be made. First, "SK" based on **SK16** where there are no link between a SN and its host galaxy (in practice using a HOSTLIB without taking the additionnal correlations into account). Then, "BP" based on **P21** where the x_1 and c parameters of the HOSTLIB are generated by asymmetric Gaussian modelings adjusted to reproduce the observed distributions in nature. The first new HOSTLIB, dubbed "NN", is based on BP but replacing the x_1 values following the previously described procedure, effectively adding one of the consequences of supposing age as the leading factor between SNe and their environment. The second new HOSTLIB is dubbed "NR", but adds an age column defined by $\text{age} = \begin{cases} 0 & \text{if} \quad r < \delta(z) \\ 1 & \text{if} \quad r > \delta(z) \end{cases}$ as discussed in Section 3.2, and an step column in which we associate a step of ± 0.065 mag for young and old SNe respectively. This step value stems from the other implication of age as the driving phenomenon under the SNe Ia correlations, defined in **R21? B21?**.

In order to replicate the actual datasets on which our stretch model was based on, i.e. the Pantheon dataset what do we do of LOWZ?, we implemented two approaches of simulation:

1. using one DATA sample with (FIND NUMBER) SNe and a unique huge BIASCOR sample of (FIND NUMBER), simulations stamped "FULL" hereafter;
2. and a collection of 500 already sample-sized DATA samples that are each corrected with the same BIASCOR sample.

For each of them, finding the scaling factor of NGEN (describe) for each survey proved crucial to the study. They were computed to have approximately the same ratio of DATA between surveys at the end of the fitting stage than in Pantheon, that is shown Table 3.

Melting pot of things to talk about, maybe some in appendix:

- Different HOSTLIBs depending on the survey (lowz, highz);
- Plot HOSTLIBs parameters
- Weigthmaps, plot them
- Talk about NGENs and how they are made to fit to the ratio in Pantheon
  - Check differences between NN/NR and SK/BP
  - It might be that having to up the NGENs of non-LOWZ surveys in NN/NR wrt SK/BP while keeping LOWZ's NGEN to 20.0 could be a result about the impact of this modeling on the generated number of low redshift SNe

Table 4. Comparison of the relative ability of each HOSTLIB implementation to describe the data. For each HOSTLIB a 2D KDE is computed from the simulated data and used to determine said χ^2.

HOSTLIB	χ^2 x_1 v z	χ^2 x_1 v M_{host}
P21	????	????
N21	????	????

- Inputs:
  - $H_0 = 70.0\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$
  - $\alpha = 0.145$;
  - $\beta = 3.1$;
  - $w = -1$;
  - $\Omega_m = 0.315$;
  - $\Omega_\Lambda = 0.685$
- Spectroscopic efficiencies
- In inputs:
  - GENMODEL: SALT2.JLA-B14 for all but LOWZ: SALT2.WFIRST-H17
  - Intrinsic scatter: G10
  - SNR > 4 or 2
  - CUTWIN_NEPOCH = 5 -5
  - GENFILTERS? GENRANGE_REDSHIFT?
  - GENMEAN, RANGE, SIGMA for x1, c, alpha beta?
- For LCFIT stage, SNRMAX = 5 for LOWZ, not the others, beware cosmological parameters in that stage
- Role of biascor samples
- Flavors of biascor samples:
  - 1D
  - 5D
  - 7D

4.2. Comparing tests on simulations

Before comparing the w v Ω_m values, we looked at the correspondence between simulated data and actual data to ensure the improvement of our LsSFR-based approach. The main idea being the evolution of the underlying stretch distribution as a function of redshift, we represented x_1 v z in log-scale using a 2D hexagonal colored histogram for the simulated data and dots for the Pantheon values. We also looked at the x_1 v Mhost plot to ponder the relationship these two main characteristics of the SNe Ia on one hand and the host galaxies on the other (see 3). We then computed a 2D kernel of each set of parameters based on the simulations and determined the associated χ^2 between the data and the kernels. The results are summarized in Table **??** and the code in the *SNprop* module.

The numerical values follow what was already clear on the figures: it fits best.

5. Impact on cosmology

What we want is not so much w than Δw wrt. best current work. We find x% and here are the contours.

6. Discussion

We expected to have a higher/lower, and we got that.

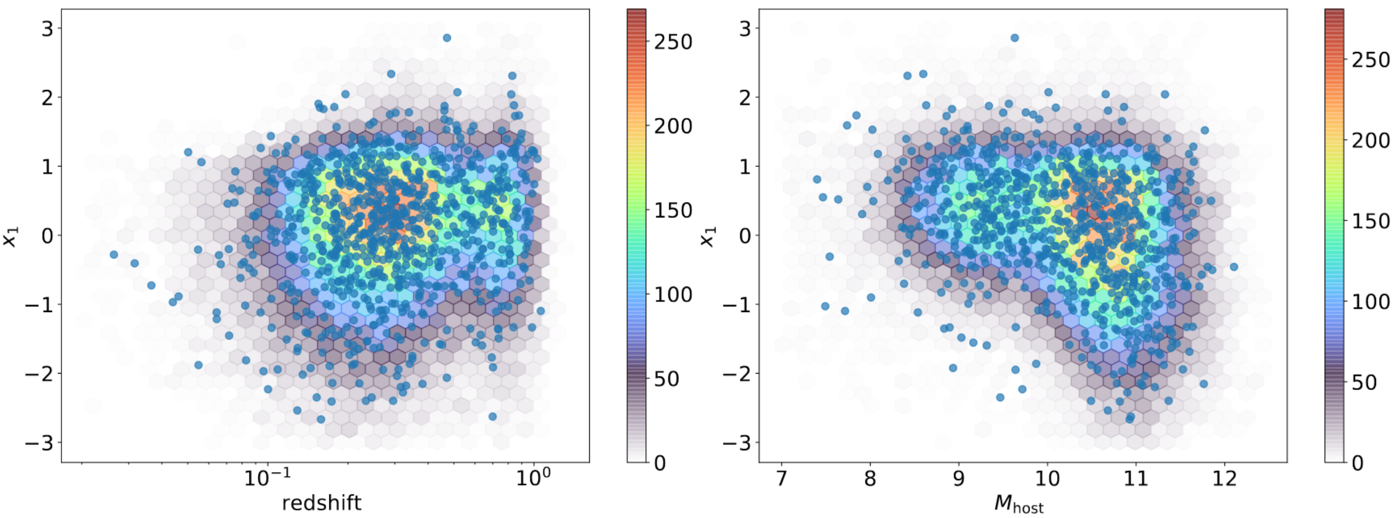

Fig. 3. *Top*: 2D hexagonal histograms of the simulated data using the P21 setup in color (left: x_1 v z, right: x_1 v M_{host}) and actual Pantheon data in blue points. *Bottom*: same data but 2D hexagonal histograms of the simulated program using the N21 improved HOSTLIB.

7. Conclusion

Should be nice.

Acknowledgements. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 Research and Innovation program (grant agreement no 759194 - USNAC). This work was supported in part by the Director, Office of Science, Office of High Energy Physics of the U.S. Department of Energy under Contract No. DE-AC025CH11231. This project is partly financially supported by Région Rhône-Alpes-Auvergne.