Title
ZTF-SEDm Type Ia supernova sample for Twins Embedding spectrophotometric standardization

Journal
Astronomy & Astrophysics, 710

ISSN
0004-6361

Authors
Ganot, C
Copin, Y
Rigault, M
et al.

Publication Date
2026-06-01

DOI
10.1051/0004-6361/202558467

A&A, 710, A315 (2026)
https://doi.org/10.1051/0004-6361/202558467

Astronomy & Astrophysics

ZTF-SEDm Type Ia supernova sample for Twins Embedding spectrophotometric standardization

C. Ganot[1,⋆], Y. Copin[1], M. Rigault[1], G. Dimitriadis[2], A. Goobar[3], K. Maguire[4], J. Nordin[5], M. Smith[2], G. Aldering[6], C. Barjou-Delayre[7], M. Betoule[8], J. S. Bloom[6,9], U. Burgaz[4], L. Galbany[10,11], M. Ginolin[12,1], M. Graham[13], D. Hale[14], J. Johansson[3], M. M. Kasliwal[13], Y.-L. Kim[15], F. J. Masci[16], T. E. Müller-Bravo[4,17], S. Perlmutter[6,18], B. Popovic[19], J. N. Purdum[14], B. Rusholme[16], J. Sollerman[20], J. H. Terwel[4], and A. Townsend[5]

(*Affiliations can be found after the references*)

Received 8 December 2025 / Accepted 18 April 2026

ABSTRACT

Aims. This paper has two aims: the first aim is to build a large homogeneous spectrophotometric sample of Type Ia supernovae (SNe Ia) from the second data release of the *Zwicky* Transient Facility (ZTF DR2). We used the spectrum sample from the low-resolution ($R \sim 100$) SEDmachine (SEDm) Integral Field Spectrograph (IFS) that gathered 3069 spectra. This is one of the largest samples of such collections that can attempt to reproduce the Twins Embedding (TE) spectrophotometric standardization method. This is our second objective. The method was developed based on high-quality spectra from 200 SNe Ia of the Nearby Supernova factory (SNfactory) and led to an exceptionally low value of 0.073 mag for the intrinsic scatter.

Methods. As the SEDm is not designed as a spectrophotometric instrument, we first improved the flux-calibration accuracy of the SN Ia spectrum sample using the ZTF photometric data, which were calibrated at the percent level. We corrected the spectra for second-order polynomials, fitted by comparing the synthetic photometry in the ZTF g, r, i filters with the light-curve (LC) data. We then applied the three steps of the TE parameterization to a subset of 783 ZTF SN spectra near maximum light while comparing results from SNfactory and ZTF. We finally analyzed the standardization methods based on the TE parameters.

Results. The precision of the phase-correction model, which is the first step of the TE, is estimated at 0.01 mag in g band based on ZTF data. Despite the challenge posed by the spectrum-extraction pipeline associated with the SEDm (flux calibration, leftover host signal, low signal-to-noise ratio, and low resolution), we applied a first standardization in color based on the second step of the TE, called read between the lines (RBTL), to the ZTF sample. We reached a Hubble residual scatter of 0.153 mag, all in normalized median absolute deviation, which is to be compared to the ~0.11 mag obtained with the SNfactory data. The SALT color and stretch standardization reaches a scatter of 0.164 mag for the same ZTF SN Ia sample, and its host steps are ~0.1 mag and zero for RBTL. When considering the scatter due to the redshift error and flux calibration error, we estimated a RBTL scatter of ~0.129 mag for this ZTF sample as an upper limit because we identified an additional contribution from a systematic error in color. We tested the standardization based on the nonlinear TE parameters, and, as expected from the low spectrum quality, it did not improve the overall dispersion.

Conclusions. We release 1897 flux calibrated spectra of 1607 SNe Ia with an estimated photometric accuracy of 0.07 mag. We further demonstrate that some amount of spectrophotometric SN Ia standardization can be done with limited-quality spectra. The RBTL standardization is more efficient than that of SALT with one parameter less, and the resulting host steps are consistent with zero. This makes it less prone to astrophysical bias. For future spectroscopic surveys, targeting the extraction pipeline for a thorough flux calibration and good signal-to-noise ratio would enable us to compute the full TE standardization, which would further reduce the scatter in the distance estimate.

Key words. methods: data analysis – surveys – supernovae: general – cosmology: observations – dark energy – distance scale

1. Introduction

Supernovae Ia (SNe Ia) are standardizable cosmological candles that are used to derive the expansion history of the Universe. They allowed the discovery that the expansion of the Universe accelerates, which is thought to be caused by an unknown dark energy (Riess et al. 1998; Perlmutter et al. 1999). Since then, cosmologists have focused on measuring the properties of dark energy, probing its equation of state parameter w_0 and its potential evolution with cosmic time w_a. If dark energy is a simple cosmological constant Λ in Einstein's equation of gravitation, we would expect $w_0 = -1$ at all times (hence, $w_a = 0$). However, recent combinations of baryon acoustic oscillation data, anisotropies in the cosmic microwave background, and the SNe Ia distance-redshift relation have led to a ~4σ tension with Λ, finding $w_0 = -0.727 \pm 0.067$ and $w_a = -1.05^{+0.31}_{-0.27}$ (Adame et al. 2025; Planck Collaboration VI 2020; DES Collaboration 2024).

Data from SN Ia are the most sensitive probe of the recent expansion of the Universe dominated by dark energy (Efstathiou 2025). The tension varies from ~2.5σ to ~4σ depending on the choice of the SN Ia dataset and the cosmological inference pipeline (Brout et al. 2022; Rubin et al. 2025; DES Collaboration 2024; Popovic et al. 2025b).

This new tension in Λ adds to the nearly decade-old Hubble tension. The Hubble constant, H_0, is found to have a higher value, at 5σ, when derived by calibrating the absolute magnitudes of SNe Ia from a direct distance ladder, in comparison to the expectation of the standard model of cosmology, whose parameters are anchored by early-epoch data (Riess et al. 2016, 2024; Planck Collaboration VI 2020).

⋆ Corresponding author: c.ganot@ip2i.in2p3.fr

If they were confirmed, these tensions would prove the current standard model of cosmology to be incorrect, and new fundamental physics should emerge. The physical nature of SNe Ia, however, which are the key cosmological probes for both of these results, is not fully understood so far, and deriving distances from SN Ia data is ultimately an empirical procedure. This affects the accuracy of the derived distances, and therefore, of the inferred cosmology.

The peak magnitudes of SN Ia have a natural scatter of ~0.4 mag, which is reduced to ~0.15 mag when two empirical relations, based on imaging data, are taken into account. These two relations connect their peak magnitudes with properties of their light curves (LCs): SNe Ia with fast-evolving LCs are fainter, as are those with redder colors. These are usually referred to as the brighter-slower and redder-fainter relations (Phillips 1993; Tripp 1998; Riess et al. 1996). The two parameters that are involved can be estimated by fitting spectrophotometric templates on broadband photometric LCs, that is, using the reference SALT model (Guy et al. 2007; Betoule et al. 2014).

When these two relations are accounted for, the residual dispersion is dominated by an intrinsic unexplained scatter of ~0.1 mag (e.g., Betoule et al. 2014; Scolnic et al. 2018), and the standardized SN Ia magnitudes are correlated with the properties of their environment. This can either be global, such as the host stellar mass (e.g., Sullivan et al. 2010; Kelly et al. 2010; Childress et al. 2013; Rigault et al. 2020; Smith et al. 2020), or local to the SN Ia, such as the local color or star formation rate (e.g., Rigault et al. 2013; Roman et al. 2018; Jones et al. 2018; Rigault et al. 2020; Kelsey et al. 2021). It raises questions about the accuracy of this photometric SN Ia standardization (Ginolin et al. 2025b,a; Popovic et al. 2025a), and consequently, of the derived distances and cosmological parameters.

Alternatively, Bailey et al. (2009) first demonstrated that we can standardize SNe Ia using spectrophotometric information alone, with a precision that is competitive with that obtained using photometry. Fakhouri et al. (2015) then introduced the concept of twins, demonstrating based on spectrophotometric timeseries data from the Nearby Supernova factory (SNfactory or SNf, Aldering et al. 2002), that two different SNe Ia can be extremely alike spectroscopically. The authors suggested that these twins went through the same explosion mechanism, which is thus expected to correspond to a similar absolute brightness. The authors indeed showed that by grouping SNe Ia by spectral similarity, a scatter along the Hubble diagram as low as ~0.07 mag can be reached. This leaves much less room for unaccounted scatter or astrophysical biases. The authors demonstrated that using a single spectrum at maximum light is enough to reach this precision. This method would allow us to dispense with photometry when the SN is found and the phase is known, and to only use spectra to infer SN Ia distances.

This initial study was expanded into a formal parameterization of twins using an enlarged dataset, and it was developed as a standardization method by Boone et al. (2021a,b) (hereafter B21a and B21b). The Twins Embedding (TE) is a three-parameter nonlinear space that captures the SN Ia spectral diversity at maximum light after accounting for the estimated reddening from spectral domains that are far from strong and highly variable features. These three parameters along with the estimated color can then be used to standardize SN Ia peak magnitudes, leading to a similar scatter of 0.073 mag. More recently, Stein et al. (2022) used a neural network probabilistic auto-encoder to extend this work using the same dataset. They employed SNe Ia with sparse spectrophotometric time series without the requirement of a spectrum near peak luminosity.

While highly encouraging, these results have been limited to the spectrophotometric time series of the SNfactory dataset so far. We here investigate whether these results can be recovered using another spectral dataset, and we focus on the TE method to do this. We used the recently published ZTF SN Ia DR2 sample (Rigault et al. 2025), which gathered 3628 spectroscopically confirmed SNe Ia acquired between March 2018 and December 2020 by the ZTF survey (Bellm et al. 2019; Graham et al. 2019).

This paper is structured as follows: we first present the characteristics of the ZTF SN Ia DR2, in particular, the SEDm spectroscopic sample, in Sect. 2 . In this section, we improve the flux calibration of the SEDm spectra based on the associated photometric data. We compare the resulting spectrophotometric sample to that of SNfactory, based on which the TE standardization method was developed.

Second, we summarize the TE standardization method in Sect. 3, and we present the ZTF SN Ia subset defined for the TE application in Sect. 4. We apply the first step of the TE, which converts it into a spectrum sample at maximum light. Third, we present the results of the TE spectrophotometric standardization on ZTF: the method called read between the lines (hereafter RBTL) in Sect. 5 and the full TE in Sect. 6. We discuss these results in Sect. 7 and evaluate the dependence of the standardization method on several SN parameters and their environment. We conclude in Sect. 8.

We use the AB-magnitude system and keep the parameter notation from B21a, except for the TE color parameter $\Delta\tilde{A}_V$, which we simplify to ΔA_V for clarity. The dispersions are in nMAD when not specified otherwise.

2. Building the ZTF spectrophotometric dataset

As the ZTF SN Ia DR2 is the largest homogeneous SN Ia dataset to date (Rigault et al. 2025), we wished to take advantage of this large ZTF spectrum sample to apply the spectrophotometric standardization method to the possible extent. The primary objective of the ZTF spectrograph is follow-up and spectral typing, and its dataset quality was not initially sufficient for this study because of the non-spectrophotometric extraction pipeline and residual contaminations.

We exploited the calibrated photometric data to build a spectrophotometric sample, which is essential for deriving distances. We examined the limitations of the flux calibration method due to host signal residuals and low SNR of the sample by quantifying the calibration precision. Section 2.1 describes the ZTF DR2 SN Ia dataset, Sect. 2.2 introduces the flux calibration method and estimates its accuracy and limitations, and Sect. 2.3 presents the spectrophotometric sample of 1897 spectra from 1607 SNe Ia that we release.

2.1. ZTF spectroscopic sample

The ZTF (Bellm et al. 2019; Graham et al. 2019; Masci et al. 2019; Dekany et al. 2020) is a time-domain survey that started scientific operation in March 2018 and is expected to continue until the end of 2026. The ZTF camera is mounted on the 48-inch Schmidt telescope at Mount Palomar Observatory and monitors the northern sky with a typical cadence of one to three days in the g, r, and i optical bands. The 47 deg^2 wide camera reaches a typical depth of 20.5 mag in g and r and of 20 mag in i in 30 seconds.

The ZTF survey has dedicated access to the SEDm (SEDm; Blagorodnova et al. 2018; Rigault et al. 2019), a low-resolution ($R \sim 100$) integral field spectrograph covering the 3776–9223 Å wavelength range that was designed to spectrally classify SN-like transients detected by ZTF. In 30 minutes, the SEDm can acquire spectra with a precision sufficient for a classification of SN type up to ~19 mag in r band. The ZTF collaboration also has access to other spectroscopic facilities through specific follow-up proposals.

As part of the Bright Transient Survey (BTS; Fremling et al. 2020) program, the ZTF collaboration used these spectroscopic facilities to classify all SN-like transients up to 18.5 mag and the majority of those up to 19 mag (Perley et al. 2020). Of these SNe, ~70% are of Type Ia. The ZTF SN Ia DR2 dataset of 3628 SNe Ia largely overlaps with the BTS sample and has 80% of the sources in common (Rigault et al. 2025). The other sources are fainter (Smith et al. 2026). As a consequence, the ZTF SN Ia DR2 sample is free of nonrandom selection (i.e., volume-limited) up to $z \leq 0.06$ (Amenouche et al. 2025; Rigault et al. 2025).

Every SN Ia in the ZTF SN Ia DR2 release has at least one spectrum that led to a secured classification for a total of 5138 spectra, 3069 of which come from the SEDm. Of these SNe Ia, 2630 are non-peculiar SNe Ia with good sampling and pass the usual cosmological quality cuts. The ZTF SN Ia DR2 contains 3652 spectra associated with these SNe Ia, 2273 (64%) of which come from the SEDm.

In this first analysis, in which we study the application of TE on a non-SNfactory dataset, we focused on this homogeneous sample of 2273 SEDm spectra from 1872 SNe Ia. These spectra were extracted using the data-processing procedure described in Rigault et al. (2019), with a typical precision of the wavelength solution of 3 Å and an accuracy in the color calibration of ~5%. The relative and absolute spectrophotometry of the sample is unknown a priori.

The background host contamination was handled using the spectroscopic scene modeling `Hypergal`, described in Lezmy et al. (2022). In this pipeline, a pure-host cube was estimated using SED fitting on public optical photometric images. This host-cube model was then aligned to match actual observing conditions on a given SEDm exposure. The SN spectrum was obtained by using this host-cube model as a background; the whole fit was made using a single forward-modeling procedure. Leftover narrow wavelength features that are not well constrained by the SED fit might still contaminate the extracted SN spectrum, however, as these features are hard to model properly based on only a few photometric images. Furthermore, targets within the core of the host still show a strong background signal despite the `Hypergal` procedure (see details in Lezmy et al. 2022).

2.2. Flux calibration of SEDm spectra

The absolute flux calibration of SEDm spectra is unknown because it was not necessary for the SN typing process. To recover it and correct for the spectral zero point (ZP(λ)) of the flux calibration, we compared the spectra to LC data to set a low-order ZP(λ). We present the flux calibration method and estimate the precision of the flux calibration.

2.2.1. Flux calibration method

Each SN of the spectroscopic sample was initially observed with the ZTF camera. SALT2.4 (Guy et al. 2007; Betoule et al. 2014) interpolates the LC data points in time, and the resulting LCs are flux calibrated to the percent level (Rigault et al. 2025). As the photometric data have a good coverage in time in r and g, the interpolation error in time is marginal. About 19% of the SNe have no data in i, and the uncertainty resulting from the extrapolation to this filter is estimated by SALT. The light-curve amplitudes x_0, computed from ZTF LCs, have an absolute floating ZP known at 9%, but a relative calibration of 1% (Lacroix et al. 2025). We therefore expect a variation of 1% from the LCs calibration.

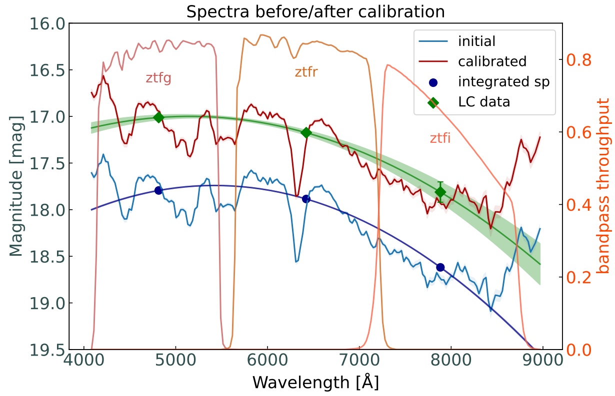

Fig. 1. Spectrum of SN *ZTF20aayvubx* at phase 0.01 day (in blue) and after flux calibration (in red). The camera filter bandpasses are displayed in orange. The blue dots show the spectrum integrated through the filters. The green dots are the LC values. The offsets are 0.780, 0.708, and 0.805 mag in g, r, and i, respectively. The second-order polynomials that pass through the three photometric points are shown in their respective colors.

We used the calibrated LC data to further calibrate the spectra. For each SN spectrum at phase p, we had access to the flux values of the three interpolated LCs at the corresponding time of observation. We compared the synthetic photometry, computed with `sncosmo` (Barbary et al. 2016) in the ZTF camera filter bandpasses to these three interpolated fluxes, as illustrated in Fig. 1.

We obtained three magnitude offsets from the gri observations through which we passed a polynomial of degree two and corrected the spectrum for this polynomial. We took the synthetic photometry error and LC data-point errors into account during the fit. The offset distributions before calibration are shown in Fig. 2 in magnitude. Their dispersion are about ±0.55, ±0.63, and ±0.83 mag on average in g, r, and i, respectively. We noted a strong correlation that propagates as an error on the polynomial of the calibration. We used a degree two, as it is the maximum degree for fitting three data points, including for SNe without i data, because the offset uncertainty in i was taken into account in the filter interpolation.

For numerical stability, we used orthogonal Legendre polynomials and fit the coefficients on the spectra in magnitude. We show the coefficient distributions in the upper panel of Fig. 4, as well as their mean and standard deviation (hereafter STD) values. Here, a_0 corresponds to a global magnitude offset, a_1 to a slope, and a_2 to a quadratic correction.

2.2.2. Precision of the flux calibration

To evaluate the spectrophotometric precision of this procedure, we conducted two studies. First, we provided an estimate of the SEDm observation errors alone, including telluric line residuals and instrumental effects. A second study included SN observation errors, which are mainly residuals from the host galaxy spectrum.

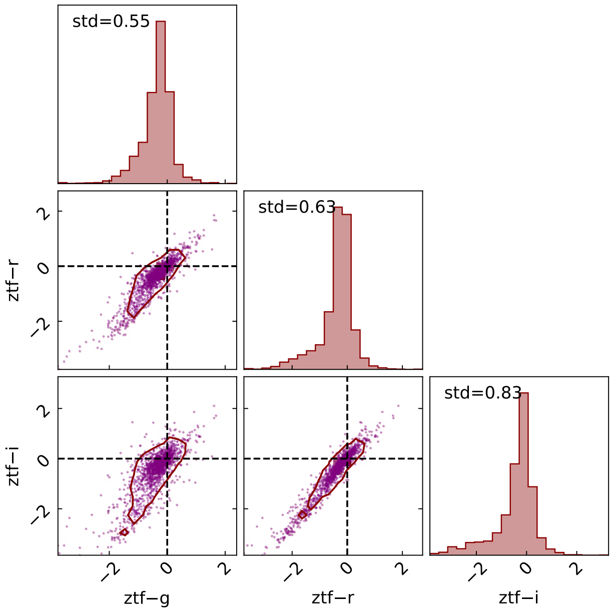

Fig. 2. Magnitude offsets between the 1897 ZTF integrated spectra and the associated LC values before calibration in *g*, *r*, and *i* filters. The STD are shown in the legend. The dashed black lines indicate the LC expectations, and the 2σ density levels of the 2D histograms are shown in dark red.

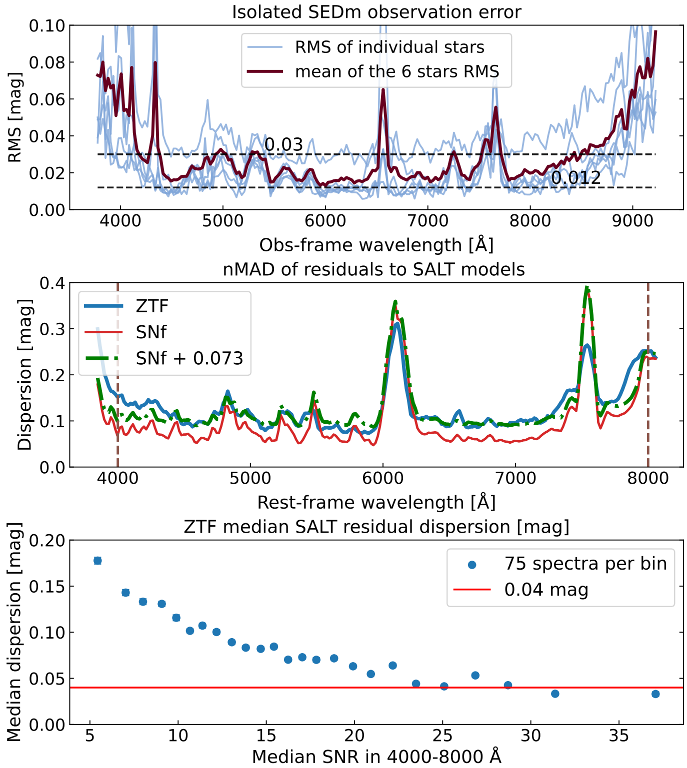

Fig. 3. Top: Precision (RMS) of the flux calibration as a function of wavelength for six standard stars (light purple) and the quadratic average (red). The limits of 0.03 mag and 0.012 mag are shown as dashed lines. Middle: Dispersion (nMAD) of the residuals relative to a SALT 2.4 spectrum model given LC data for 1897 ZTF spectra and 583 SNfactory spectra. The nMAD of SNfactory (light green) plus an additional 0.07 mag noise from ZTF are also shown. The wavelengths delimiting a good precision are shown with dashed lines at 4000 Å and 8000 Å. Bottom: Median SALT model residual dispersion (nMAD) binned in SNR for the ZTF (75 SNe per bin). A floor of 0.04 mag is shown by the red line.

Estimation based on standard stars The SEDm data acquisition includes observations of standard stars at a daily frequency to derive a primary flux calibration. Their fluxes are known and considered constant in time. We applied the calibration procedure on all the spectra of each of six standard stars (Feige34, G191B2B, HZ2, HZ4, HZ44, and LB227) for a total of 518 spectra based on their broadband magnitudes computed from the reference catalog of CALSPEC standard star spectra (Bohlin et al. 2020). Then, we compared the newly flux-calibrated spectra to the reference spectra.

The dispersion of the residuals of the star calibrated spectra with their reference spectra, shown in the upper panel of Fig. 3, gives an estimate of the calibration precision in flux. We computed it for each star and then averaged it out. We obtained a spectral RMS between 0.012 mag and 0.03 mag in the 4000–8500 Å range. The dispersion is larger for lower and higher wavelengths because of the instrumental response and quadratic flux calibration, illustrated in Fig. 1, especially at large wavelengths because the photometry in *i* is less precise. Atmospheric absorption lines are visible: the double peak of the dispersion around 7600 Å corresponds to the Fraunhofer A O_2 atmosphere absorption line, and the smaller peak next to it is the H_2O absorption line at 7200–7300 Å. A varying absorption line of the stars is visible as well: the Hα feature at 6563 Å.

This process allowed us to derive a lower limit of the calibration precision because it does not contain the host subtraction errors and because SNe are fainter than our standard stars; the latter are detected at ~11 to 15 mag in Bessell-*V* and at a mean of ~18 mag for Type Ia SNe. Nevertheless, it is a good approximation for a bright isolated SN.

Estimation based on SNe SALT2.4 (Guy et al. 2007; Betoule et al. 2014) provides SN Ia spectral models for given LC data, a spectrum phase, and a wavelength range. Using the `sncosmo` library (Barbary et al. 2016), we computed a spectral model for a given SN Ia spectrum by providing its associated SALT2 parameters derived from LC data. The method for isolating the ZTF spectrum flux-calibration error is the following. We compared the ZTF SNe Ia spectra to their associated SALT spectrum model and computed the dispersion of their residuals. This dispersion in wavelength originates from the SN spectrum error and the SALT model error. To isolate SN spectrum errors for ZTF, we subtracted the SALT model error. A proxy for this error was derived by computing the dispersion of the residuals of the SNfactory spectra with their SALT models, under the hypothesis that their flux calibration error is negligible.

We computed the dispersion of the spectrum model residuals for the ZTF sample using only 1897 spectra from 1607 SNe Ia after the cuts of the DR2 (Rigault et al. 2025) that are described in Table 1. The LC data we provided to `sncosmo` are described in Smith et al. (2026). This dispersion is shown in Fig. 3. We also computed the SALT residuals for the 583 SNfactory spectra from 203 SNe used in B21a. The residual dispersion shows a similar response in wavelength to that of ZTF, as shown in Fig. 3, because the SALT model is limited at the location of strong SN lines.

To estimate the full ZTF spectrum precision after calibration, we added an error floor in quadrature to the SNfactory SALT residual dispersion. A floor of 0.07 mag causes the dispersions to match at most wavelengths, between 4000 to 8000 Å,

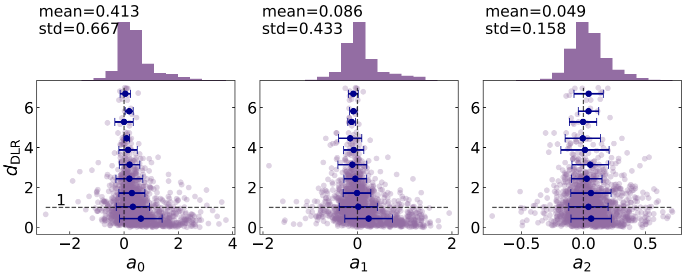

Fig. 4. Top panels: Distributions of the calibration coefficient a_0, a_1, and a_2 for 1897 spectra. Their mean and STD are indicated in the legends. Bottom panels: d_{DLR} against the calibration coefficients. The dark blue data points show the coefficient mean and STD binned on 10 d_{DLR} values. The vertical dashed black lines indicate the values corresponding to no correction, and the horizontal lines correspond to the $d_{\mathrm{DLR}} = 1$ limit.

as illustrated in Fig. 3. As ZTF is flux calibrated using SALT parameters, its residuals are closer to zero by construction than those of SNfactory, and we therefore focused on residual scatter (nMAD) and not on the total RMS. After the calibration process, the remaining ZTF error was almost achromatic, at 0.07 mag on average.

There are two contributions to this additional scatter at 0.07 mag from ZTF compared to SNfactory: a flux-calibration error floor, and an error scatter due to the weak SNR of ZTF spectra (∼18 per wavelength bin). To separate these contributions, we reproduced the aforementioned analyses per bin of SNR, and we show the evolution of the scatter in SNR in the lower panel of Fig. 3. We computed the average (median) nMAD per SNR bin and on a wider wavelength range of [4000, 8000] Å because the median ignores the strong variations from SN features. As expected, the magnitude scatter decreased as the SNR increased, until it reached a floor of 0.04 mag at SNR ∼ 25. This floor corresponds to the precision in flux calibration with respect to SNfactory, and it is consistent with the floor estimated from isolated standard stars (see Sect. 2.2.2). This calibration error adds up quadratically to that of the SNR (therefore 0.057 mag), to explain the total ZTF spectral error after calibration, giving a total of 0.07 mag on average.

2.2.3. Impact of host contamination

To estimate the amount of host galaxy contamination in the total spectrum extraction error, we studied the correlation between the position of the SN in the galaxy and the calibration coefficients, introduced in Sect. 2.2.1. The host subtraction is more critical in the inner parts of the galaxy. We used the d_{DLR} (Sullivan et al. 2006; Gupta et al. 2016), which is defined here as the normalized distance to the galaxy center, where zero is the center, and one is the luminosity radius of the galaxy in the direction of the SN.

The coefficients as a function of the d_{DLR} values are shown in Fig. 4 for 1897 spectra of 1607 SNe after DR2 cuts. We note that the dispersion of the a_0 and a_1 coefficients increases for lower d_{DLR}, and in particular, below 1. As the dispersion spreads, the mean value is also shifted to higher values. This means that the closer the SN to the host galaxy, the more significant the calibration correction. The quadratic coefficient a_2 is poorly correlated with the d_{DLR}, as its Pearson coefficient is $r = -0.073 \pm 0.023$, while $r = -0.238 \pm 0.022$ for a_0 and $r = -0.262 \pm 0.022$ for a_1.

A significant part of the calibration correction, and therefore, of its error, is due to the host galaxy residuals, especially as a

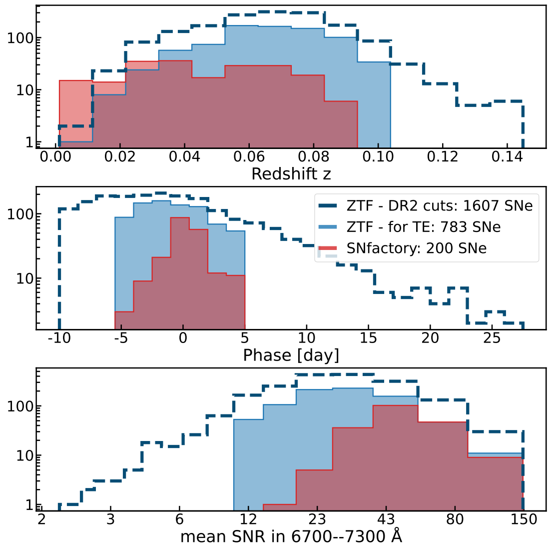

Fig. 5. Distributions of SN redshifts, phases nearest maximum, and SNR defined in the range 6700–7300 Å for the flux-calibrated sample of ZTF SNe Ia after DR2 cuts (dashed blue line), the subsample for TE application (blue), and the SNfactory sample (red).

magnitude offset and a slope. There is a correction bias, correlated with d_{DLR}, and SNe with low d_{DLR} parameter are thus more likely to have a large flux error.

2.3. The spectrophotometric ZTF-SEDm SN Ia sample

We provide a selected sample of spectra after flux calibration. It is an homogeneous sample from the DR2 spectra, only observed with the SEDm. It consists of 1897 spectra from 1607 SNe Ia. We removed the poorly calibrated spectra with a cut on the calibration coefficients, as detailed in Table 1. We propagated the polynomial variance to the calibrated spectra, neglecting non-diagonal covariance terms. Similarly to B21a, we applied a cut on SNR defined in the bluest and reddest values of the spectrum, which we label $\mathrm{SNR}_{\mathrm{poly}}$, to take this polynomial error into account. We defined it in the 1000 Å extrema values, and 7% of the sample were removed as a result. This error was clearly separated from the spectrum error in order to properly maintain the spectrum SNR.

The spectroscopic sample was flux calibrated using photometric data whose accuracy is at the percent level, and the final average dispersion was estimated to be 0.07 mag in the 4000–8000 Å range. This combined a common error floor of 0.04 mag and a contribution from the lowest SNRs. The redshift, phase, and SNR distributions of this sample are shown in Fig. 5. We note that the redshift accuracy varies from $\sim 10^{-5}$ to $\sim 10^{-3}$ depending on its source. At this stage, spectra are still in observer frame and are not corrected for Milky Way extinction. The flux-calibrated sample is available online[1].

3. The spectroscopic standardization method

This section summarizes the TE spectrophotometric standardization method (B21a) to introduce the concepts needed when it is applied to the ZTF SEDm data. Similarly to the photometric SN Ia standardization, the first step of the TE spectroscopic standardization is to estimate parameters that explain the input SN Ia

[1] https://ztfcosmo.in2p3.fr/

data. These are the stretch x_1 and color c in the SALT-based photometric analysis (Guy et al. 2007, 2010), while these are three $\xi_{1,2,3}$, and a color (named ΔA_V) for TE, as detailed in B21a. One advantage of the TE method is the access to the rest frame to derive its parameter. While the TE method aims to standardize SNe Ia using spectra alone, it uses the phase, which is derived from photometric data for now.

The second standardization step infers correlations between the $\xi_{1,2,3}$ and ΔA_V parameters and the SN Ia peak brightness, thereby reducing the scatter along the Hubble diagram by empirically accounting for observed variabilities. We summarize these two steps in the next subsections.

3.1. Spectroscopic parameters ΔA_V and ξ

The TE parameterization consists of three steps: first, to correct for spectral variations in phase to obtain spectra at maximum light, second, to fit a RBTL color term ΔA_V, and third, to extract nonlinear Isomap parameters from spectral features. In principle, steps 2 and 3 might be made simultaneously, but this was not the approach chosen by B21a. As discussed in Sect. 4.2, no TE vector was retrained for this analysis; we used them as given by B21a.

3.1.1. Phase correction

The objective of this first step is to estimate the spectrum at maximum light per SN Ia. B21a approximated the quadratic evolution in phase of SN Ia per wavelength with respect to maximum light within the ±5 days rest-frame phase range,

$$m_i(\lambda, p) - m_i(\lambda, 0) = c_1(\lambda) \cdot p + c_2(\lambda) \cdot p^2, \tag{1}$$

with $m_i(\lambda, p)$ the spectrum of SN i expressed in magnitude at rest-frame phase p as a function of rest-frame wavelength λ; and $c_{1,2}(\lambda)$ two spectroscopic model coefficients to correct for the phase p. This differential time evolution model (hereafter DTEM) only corrects for variability that is common to all SNe Ia and not for variability that is specific to individual targets. B21a reported that for the SNfactory dataset, and averaged over all wavelengths, 84.6% of the variance in the evolution of spectra near maximum light is common to all SNe Ia and is thus captured by the DTEM. The DTEM is independent of photometric LC data; nevertheless B21a also proposed a DTEM with additional stretch corrections. We retained the standard model for our baseline. The extended DTEM, which takes into account the stretch, is analyzed in Appendix C.

Finally, when several spectra for a given SN Ia were available within this ±5-day phase range, they were combined after phase correction to have only one spectrum at maximum light per target. This takes the error of the model into account, which is modeled as a function of the phase (details in B21a). Once the vectors of this model have been trained on SNfactory, the DTEM can simply be applied to other samples, so we used those from B21a.

3.1.2. Read between the lines

After we had one spectrum per SN Ia, corresponding to the spectrum expected at maximum light, we estimated the color component ΔA_V. We used the method called read between the lines introduced by B21a.

This method captures the color term ignoring the strong absorption lines, which exhibit great dispersion from one SN Ia to the next. Outside of these well-identified lines at rest frame, the spectrum of SN i at maximum light ($f_{\max,i}(\lambda)$) can be modeled as $f_{\mathrm{model},i}(\lambda)$,

$$-2.5 \log_{10}\left(\frac{f_{\mathrm{model},i}(\lambda)}{f_{\mathrm{ref}}(\lambda)}\right) = \Delta m_i + \Delta A_{V,i} \cdot CL(\lambda), \tag{2}$$

with a reference SN Ia spectrum at maximum ($f_{\mathrm{ref}}(\lambda)$), reddened by an established dust-color law $CL(\lambda)$ with the coefficient $\Delta A_{V,i}$, and with a gray offset Δm_i.

In practice, $f_{\mathrm{ref}}(\lambda)$ was assumed to be the average SN Ia spectrum at maximum of light of the SNfactory sample. The strong spectral features with large dispersion, common to all SNe Ia, were identified by B21a by estimating a wavelength-dependent intrinsic scatter $\eta(\lambda)$, defined as a fraction of flux, or a difference in magnitude. Wavelengths at which this scatter was strongest were de-weighted during the fit. After $\eta(\lambda)$ has been modeled, strongly varying wavelengths can be fully discarded with a Boolean cut on η. This was not considered in B21a, but we illustrate the benefits of this stricter RBTL approach in Sect. 5.

The RBTL parameters $\Delta A_{V,i}$ and Δm_i were fitted for each SN Ia with a simple χ^2 minimizing $f_{\max,i}(\lambda)$ and $f_{\mathrm{model},i}(\lambda)$ given measurement errors (see details in Sect. 4.1 of B21a). The dereddened SN Ia spectrum is thus defined as

$$f_{\mathrm{dered},i}(\lambda) = f_{\max,i}(\lambda) \cdot 10^{+0.4(\Delta m_i + \Delta A_{V,i} \cdot CL(\lambda))}. \tag{3}$$

3.1.3. Manifold learning

The last step of the TE parameterization estimates the ξ parameters that aim to describe variabilities in SN Ia spectra at maximum, not already accounting for the RBTL parameters. Following the initial concept of twin spectra (Fakhouri et al. 2015), the TE is constructed based on spectral distances between SNe Ia that are input to an Isomap algorithm (Tenenbaum et al. 2000). This algorithm computes a nonlinear low-dimensional reduction known as manifold learning, which was shown to outperform a classical (linear) PCA in its ability to capture data variance in a smaller number of eigenvectors.

For the SNfactory data, the fraction of explained variance reaches a limit at the third component (see Sect. 5.1 of B21a). With three components $\boldsymbol{\xi}$, this manifold technique can account for 86.6% of the intrinsic SN Ia dispersion. This explained dispersion is mostly contained within strong and highly variable spectral regions of SN Ia spectra, reducing the scatter on Ca II and Si II features from ~0.3 mag to less than 0.1 mag (see B21a Fig. 6), for example.

In practice, when the Isomap was trained on SNfactory data, the process of estimating the parameters $\boldsymbol{\xi}$ for another SN Ia sample was the following: the spectrum $f_{\mathrm{dered},i}(\lambda)$, corrected for RBTL color and magnitude, was input to the trained Isomap algorithm from the `sklearn` library (Pedregosa et al. 2011), and three parameters $\boldsymbol{\xi}_{1,2,3}$ were returned per SN.

3.1.4. Modeling of spectral variation

An extension of the spectral model described in Eq. (2), making use of the $\boldsymbol{\xi}_i$ spectral impact, was developed by training a Gaussian process (hereafter GP). It returns a single value per wavelength as a ratio of the flux for given Isomap coordinates. This allowed us to model the spectrum of SN i at maximum light more accurately, and Eq. (3) becomes

$$f_{\mathrm{model},i}(\lambda) = f_{\mathrm{ref}}(\lambda) \cdot 10^{-0.4(\Delta m_i + \Delta A_{V,i} \cdot CL(\lambda))} \cdot (1 + \mathrm{GP}(\lambda)(\boldsymbol{\xi}_i)). \tag{4}$$

This model was not developed as a standardization method in B21b because it does not intend to decrease the overall gray dispersion. Instead of modeling the spectral variation through this GP per wavelength, B21b therefore used an alternative method to explain the gray offset, as presented in the next subsection.

3.2. Twins-embedding standardization

After we defined the RBTL (Sect. 3.1.2) and manifold learning (Sect. 3.1.3), we introduce the TE standardization method from B21b. This spectroscopic standardization relies on the Isomap parameters $\boldsymbol{\xi}$ from Sect. 3.1.3, and is implemented using a GP (B21b) as

$$\delta m_i^{\mathrm{GP}} = GP\left(\mu(m_{\mathrm{ref}}, \omega \cdot \Delta A_{V,i}); \sigma(\boldsymbol{\xi}_i, \boldsymbol{\sigma}^2_{p.v.,i}, \sigma_u^2)\right). \quad (5)$$

For a given SN i, the mean function of the GP uses a ΔA_V correction ($\omega \cdot \Delta A_{V,i}$) alongside a gray magnitude offset m_{ref}; the variance function of the GP uses the Isomap parameters $\boldsymbol{\xi}_i$ and the peculiar velocity $\sigma^2_{p.v.,i}$, along with GP hyperparameters common to the set of SNe, such as the intrinsic dispersion σ_u. This model yielded an unexplained dispersion of 0.073 mag (σ_u) from the SNfactory sample of 134 SNe that met the standardization requirement cuts (detailed in B21b). In the following study, we fixed $\omega = 0$, in contrast to B21b, as we wished to correct for any additional color variation in a separate step.

4. ZTF sample for Twins Embedding

The TE method was developed and tested using 173 spectrophotometric SN Ia spectra with a high SNR from the SNfactory sample (Aldering et al. 2002), rebinned to a $R \sim 300$ resolution (initially $R \sim 2000$). The method was trained by B21a on half of the sample, and it was tested on the other half at the un-blinding of the results to validate the method. It has not been applied to another dataset so far.

In this section, we apply relevant cuts to the ZTF sample for TE application and compare the resulting subsample with that from SNfactory. We ran the preliminary data processing common to SNfactory, whose full framework can be found in B21a. Finally, in this section, we interpolate the spectra to maximum light using the DTEM.

4.1. Cuts and processing of the sample

Table 1. Cuts applied to different ZTF samples: the selection, the removed fraction, and the number of SNe remaining in the sample.

Cut	Interval	% removed	# SNe
	Flux-calibrated sample		
Source	only SEDm	40%	1872
	$\lvert a_0 \rvert < 4$	1 %	1853
Flux coefs	$\lvert a_1 \rvert < 2$		
	$\lvert a_2 \rvert < 0.75$		
Calib. error	$\mathrm{SNR_{poly}} > 10$	7%	1721
	Evaluation of the precision		
	$-0.2 \leq c \leq +0.8$	7%	1607
	$-3 \leq x_1 \leq +3$		
ZTF DR2	$x_{1,\mathrm{err}} \leq 1$		
	$c_{\mathrm{err}} \leq 0.1$		
	$t_{0,\mathrm{err}} \leq 1$		
	$\mathrm{fitprob} > 10^{-7}$		
	RBTL sample		
Phase	±5 days	41%	939
Redshift	$z < 0.1$	5%	888
Peculiar	cosmo	8%	824
Quality	SNR > 10	3%	783
	Standardization		
Redshift	$z > 0.03$	3%	744
RBTL	$-0.5 < \Delta A_V < 0.5$	9%	688

Table 1 details the cuts we applied to the sample, for the full flux-calibrated sample and for the sample on which we evaluated its precision. We created two other samples: a sample of 783 SNe for the application of the TE method, and a sample of 688 SNe for the standardization using the TE.

For the TE sample of 783 SNe, we kept good-quality data, that is, data with a cut on redshift above 0.1 and a cut on SNR < 10 that we defined in the wavelength range of 6700–7300 Å as the mean of the ratio of the flux and the STD of the fluctuations, so corresponding to an SNR per pixel. The largest cut was on the phase because the TE applies to spectra within ±5 days from peak. The spectra outside this interval represent 41% of the sample.

Distributions of redshift, phase, and SNR are presented in Fig. 5 for the ZTF flux-calibrated sample, the TE sample, and for the SNfactory sample. The figure shows that the ZTF sample has more high-redshift SNe than the SNfactory sample. The phase distribution is more uniform for ZTF, with more SNe observed before maximum. Most SNe in the SNfactory sample have a spectrum at one day from maximum. The SNR is greater for the SNfactory sample, with a median of 49, while the median is 29 for ZTF.

We prepared the ZTF spectra before applying the TE. We corrected for the recession speed by shifting the spectra to the rest frame. As the rest-frame wavelength range is different for each SN, we interpolated them to a common range of 4103–8388 Å by keeping the same binning as B21a at 1000 km.s^{-1}. The spectra were also shifted in flux assuming some cosmology to a common redshift, $z_{\mathrm{ref}} = 0.05$, in order to capture the dispersion of the magnitudes uncorrelated with redshift. At this redshift range, the cosmological model used to de-redshift does not significantly affect the result. The spectra were corrected for the Milky Way extinction using the Cardelli extinction law (Cardelli et al. 1989) and extinction maps from Schlafly & Finkbeiner (2011), also assuming a total-to-selective extinction ratio $R_V = 3.1$.

The SNe with a low redshift accuracy ($\sim 10^{-3}$) have an additional spectral uncertainty when the spectra are shifted to rest frame and corrected in flux to the common redshift of 0.05. It was not propagated in the spectrum variances and will need to be taken into account separately.

4.2. Wavelength range for TE

As the ZTF sample is defined in the 4103–8388 Å wavelength range, we truncated the TE vectors as $c_{1,2}(\lambda)$ or $f_{\mathrm{ref}}(\lambda)$. These were trained on the SNfactory spectra, which were defined in rest frame on the 3305–8586 Å wavelength range. To compare the TE application results between ZTF and SNfactory, we computed the results in parallel for SNfactory spectra in the common wavelength range.

By applying the TE method to the ZTF sample of 853 spectra from 783 SNe Ia step by step, we wished to analyze the method, benefiting from the large sample size, and to determine whether

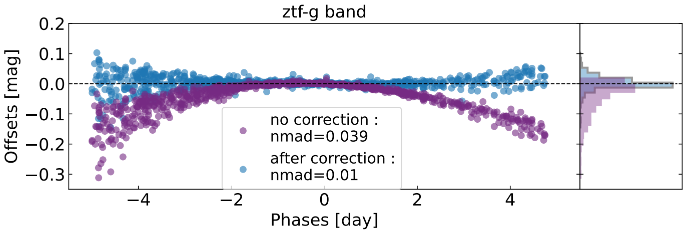

Fig. 6. Magnitude offsets between the LC value at $t_{\max}$ in the g band and integrated spectrum before (purple) and after (blue) the DTEM correction for 783 SNe as a function of the initial phase of the spectrum. The two distributions are shown in the right plot.

the accuracy of our data is sufficient for a spectrophotometric standardization.

4.3. Phase correction

The sample contains spectra at rest-frame phases between −5 and +5 days. To correct for spectral time evolution, we used the DTEM presented in Sect. 3. It provides one spectrum at $t_{\max}$ per SN, using the wavelength-dependent phase evolution described in Eq. (1). The DTEM error in phase is propagated to the spectrum uncertainty. When multiple spectra were available for one SN, the spectrum at maximum was computed by combining the corrected spectra (see details in B21a). After this procedure, the ZTF sample contained 783 spectra, one for each of the 783 SNe Ia.

To estimate the precision of the DTEM, we compare in Fig. 6 LCs values at maximum light, interpolated by SALT2.4, to the synthetic magnitude from spectra before and after phase correction as a function of the initial phases of the spectra. We computed it in the g filter because it is closest to the Bessell-B bandpass on which the maximum luminosity is defined. This allowed us to test for evolution in phase: before DTEM correction, we therefore expect the photometric flux to reach a maximum around $t_{\max}$. After correction with the DTEM, a spectrum integrated in this bandpass should return a constant value for any initial phase. When there were several spectra per SNe, we computed the offset for the spectrum at the phase nearest to the maximum.

Before correction, the offsets evolved with phase as expected, reaching a maximum around $t_{\max}$. This evolution was correctly removed by the phase-correction procedure, which displays flat residuals with phase. The dispersion (nMAD) evolves from ~0.039 mag to ~0.008 mag, although the model is less accurate near the phase extrema. Current use of the DTEM does not include any parameter to describe the variation in the time evolution between SNe Ia, such as the SALT2.4 stretch parameter x_1, which we left out to be independent of photometry.

By keeping only the 512 spectra (out of the 783) with a phase between ±3 days, we achieved a phase-correction residual scatter in g as low as 0.007 mag.

5. Read between the lines

We applied the RBTL method, the concept of which is summarized in Sect. 3.1.2, to the 783 ZTF spectra at maximum light. We compared the results to the result obtained using the 200 spectra from SNfactory[2] in the same common wavelength range (see Sect. 4.2). We describe the method of application in Sect. 5.1, and we show the results in Sect. 5.2. We show spectral residuals in Sect. 5.3, and we discuss its connection to photometry. In Sect. 5.4 we explore an initial spectrophotometric standardization method, which consists of correcting for the RBTL color parameter.

5.1. RBTL parameter estimation

We fit the two RBTL parameters (the color term ΔA_V and the magnitude offset Δm) for each SN i independently.

The fit was made by minimizing a χ^2 for each SN. We compared the observed SN i spectrum at maximum light $f_{\max,i}(\lambda)$, and its model, which was computed using Eq. (2). The reference flux is the SNfactory mean flux, $f_{\mathrm{ref}}(\lambda)$, and is shown in Fig. 7. Following B21a, we used the Fitzpatrick (1999) extinction-color relation $CL(\lambda)$ assuming $R_V = 2.4$. This is close to $R_V = 2.8\pm0.3$ found by Chotard et al. (2011). The total error is

$$\sigma^2_{\mathrm{tot},i}(\lambda) = \sigma^2_{\max,i}(\lambda) + \sigma^2_{\mathrm{poly},i}(\lambda) + \left(f_{\mathrm{model},i}(\lambda)\cdot\eta(\lambda)\right)^2, \quad (6)$$

with $\sigma^2_{\max,i}(\lambda)$ the variance of the spectrum at maximum light of SN i, and $\eta(\lambda)$ the chromatic intrinsic dispersion, assumed to be a fraction of flux (i.e., close to a magnitude scatter). Both $\eta(\lambda)$ and $f_{\mathrm{ref}}(\lambda)$ were derived from the SNfactory training (B21b) (see Fig. 7 and Sect. 3). Because of the calibration process, we took an additional uncertainty into account, $\sigma^2_{\mathrm{poly},i}(\lambda)$, propagated from the calibration uncertainty (see Sect. 2.2). Most of these uncertainties diverge at low and high wavelengths, as illustrated in Fig. 1, and therefore act as de-weighting the fit of RBTL parameters in these parts of the spectra.

The Δm parameter is the gray offset left over by the color-law correction, $\Delta A_V \cdot CL(\lambda)$. For instance, Δm receives contributions from SN Ia common intrinsic gray scatter (estimated at ~0.073 mag in B21b), peculiar velocity scatter, and from the flux calibration precision, which we estimated for ZTF in Sect. 2.2. Because the low spectral SNR affects the estimate of the overall spectral amplitude very little, we considered the contribution of the SNR to be negligible, and we only considered the spectrophotometric precision as safely contributing to Δm. The remaining variance between $f_{\mathrm{model},i}(\lambda)$ and $f_{\max,i}(\lambda)$ includes the SN Ia chromatic variations (e.g., absorption line scatter), the chromatic flux-calibration error, and a potential miscorrection for the host contamination or atmospheric features. The SN Ia absorption lines are the main sources of variance that is not explained by the RBTL model, as illustrated in Fig. 7.

To prevent this chromatic dispersion from being included in the SN Ia color ΔA_V and the gray offset, Δm, B21a used $\eta(\lambda)$ to implement inverse variance weighting. Instead, we used it as a binary mask, passing only the wavelengths with $\eta(\lambda) < \eta_{\mathrm{lim}} = 0.073$ mag, shown in Fig. 7. This scatter amplitude is higher numerically for $\lambda > 6500$ Å. In comparison to photometry, the strength of the spectroscopic approach is the ability to remove these high SN Ia variance regions while estimating the SN Ia color.

5.2. RBTL parameters

We applied the RBTL method to the SNfactory sample of 200 SNe and the ZTF sample of 783 SNe. The distributions of the RBTL parameters are shown for both samples in Fig. 8. The parameter Δm has a similar distribution, but with a narrower scatter for SNfactory (0.13 mag) than for ZTF (0.17 mag), and a shift from zero is expected for ZTF as the LCs absolute ZP

[2] We cut three SNf SNe of the 203 at $z > 0.1$ to compare targets at the same redshift range as ZTF.

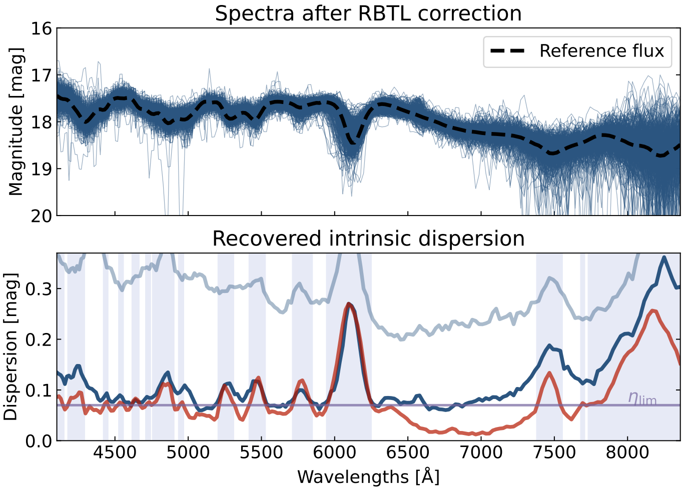

Fig. 7. Spectral correction with RBTL. Top: 783 ZTF SNe Ia spectra, corrected for RBTL parameters (dark blue), in magnitude, and the SNfactory reference spectrum (dashed black line). Bottom: Dispersion (nMAD) before correction (light blue) and after correction (dark blue). The model vector $\eta(\lambda)$ from B21a, trained on SNfactory, is shown in red. The high-dispersion areas, defined as $\eta(\lambda) > \eta_{\rm lim}$, and masked during the RBTL fit, are shown as vertical light purple bands in the bottom plot.

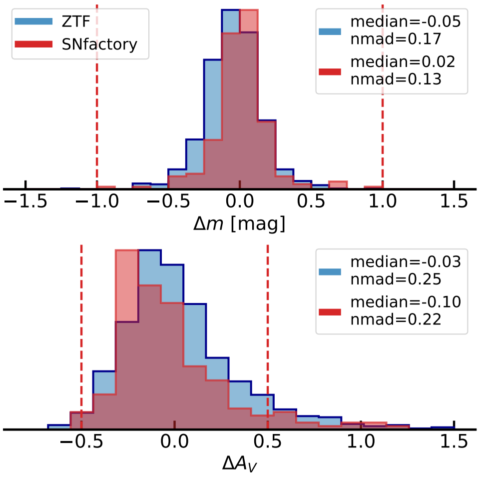

Fig. 8. Normalized RBTL parameter Δm and ΔA_V distributions for ZTF (783 SNe; blue) and SNfactory (200 SNe; red). The cuts applied for the standardization sample of those parameters alone are shown with dashed pink lines. The median and nMAD of the distributions are shown in the legends.

is unknown (see Sect. 2.2.1). However, the ΔA_V distributions differ slightly between the two samples: ZTF contains a relatively higher number of red SNe. In comparison, the SALT2 color distributions, shown in Fig. 9, seem to be more similar in the two samples. We conducted statistical tests (e.g., k-sample Anderson-Darling test, hereafter A.D.) to compare the sample distributions for both colors: they show a more significant difference for the RBTL color (e.g., A.D. p-value: 0.003) than for the SALT color (e.g., A.D. p-value: 0.04). This suggests that the ΔA_V difference is also caused in part by leftover calibration issues; we discuss this further in Sect. 7. This might in turn explain the larger scatter on Δm from ZTF because Δm and ΔA_V were fitted simultaneously.

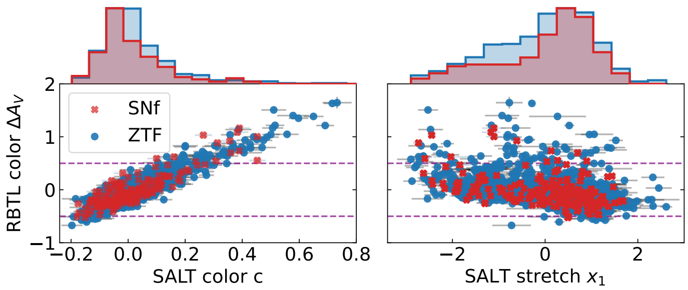

Fig. 9. Read-between-the-lines color ΔA_V as a function of the SALT color c and SALT stretch x_1 for the SNfactory sample (red) and the ZTF sample (blue). The cuts on ΔA_V for standardization at ±0.5 are shown with dashed lines.

We compare the RBTL color ΔA_V with the SALT color c and stretch x_1 parameters in Fig. 9. As expected, ΔA_V, as an SN Ia color derived from spectroscopy, is tightly correlated with c, which is derived from photometry. Interestingly, ΔA_V is also correlated with x_1: the bluest SNe are significantly slower because the stretch values are above zero, while starting at $\Delta A_V \sim -0.3$, the two stretch modes (Rigault et al. 2020) are present. This is consistent with the photometric sample results from Ginolin et al. (2025a,b), based on the c parameter, which showed that SNe in (locally) red environment are distributed equally in the two modes when the locally blue SNe are concentrated in the high-stretch mode.

5.3. RBTL dependence on wavelength

To estimate the spectral residuals of this RBTL fit, we applied the correction described in Eq. (3) to both samples. We show the residual spectroscopic dispersion for ZTF and SNfactory in the bottom panel of Fig. 7. The dispersion corresponds to the variance that is not explained by the RBTL color+magnitude offset model. For comparison, we also show the original ZTF scatter prior to the application of the RBTL. At wavelengths bluer than 6500 Å, the unexplained scatter for ZTF and SNfactory agrees remarkably well. SNfactory exhibits a net reduction in the 6500–7300 Å region, as the RBTL procedure captures nearly all the SN Ia variance at ∼7000 Å. However, the ZTF dispersion is not reduced in this region, as it remains stable with ∼0.1 mag dispersion and rises to a higher dispersion at the redmost wavelengths, as does SNfactory. The ZTF spectral behavior can be reproduced using the SNfactory curve by adding a 0.057 mag error floor in quadrature (see Appendix A), caused by the lower SNR in ZTF spectra, as detailed in Sect. 2.2. We expect that the dispersion originating from the flux calibration residuals is captured by Δm, as discussed in Sect. 5.1.

We compared Δm to a magnitude derived from photometry, as m_B. We added a term λ_0, which we placed at the Bessell-B effective wavelength, at 4385 Å (Bessell 1990), so that the right side of Eq. (2) was rewritten for SN i as

$$\Delta m_{4385,i} + \Delta A_{V,i} \cdot \left[CL(\lambda) - CL(\lambda_0 = 4385\text{Å})\right], \quad (7)$$

with $\Delta m_{4385,i}$ the magnitude offset of the spectrum $f_{\max,i}(\lambda)$, compared to $f_{\rm ref}(\lambda)$, at 4385 Å[3]. Using the ZTF sample, we computed the RBTL magnitude parameter in the configuration Δm_{4385}, while the color term ΔA_V did not change by construction. We

[3] λ_0 is interpreted here as a leverage wavelength around which the color law rotates with an amplitude depending on the color factor ΔA_V.

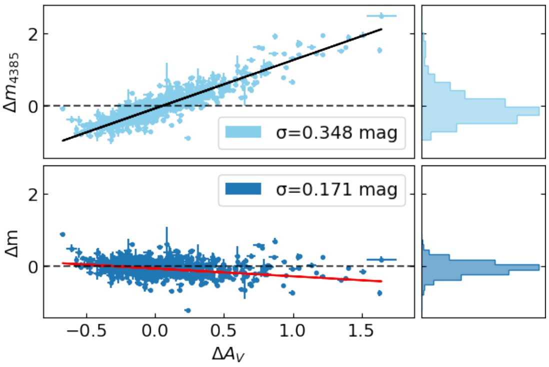

Fig. 10. Top: Magnitude offsets for 783 SNe computed at 4385 Å against the RBTL color ΔA_V. Bottom: RBTL offsets, i.e., gray offsets. Right plots: Corresponding distributions. The magnitude dispersions are provided in the legend. The black and red lines highlight the color correlation discrepancy between the two plots.

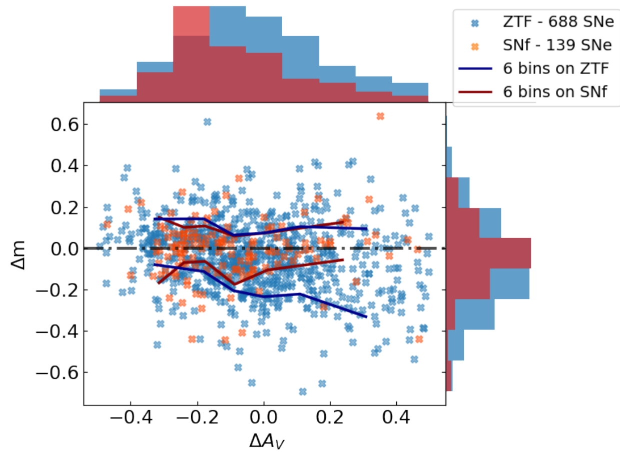

Fig. 11. Read-between-the-lines parameter Δm as a function of ΔA_V for ZTF (light blue) and SNfactory (orange). For both samples, Δm is binned on ΔA_V to display the trend with color. It is shown by lines encompassing the uncertainties of the bins for six bins with an equal number of SNe : 114 SNe per ZTF bin (dark blue lines) and 23 per SNfactory bin (red lines). $\Delta m = 0$ is indicated as a dashed black line.

show both magnitude distributions in Fig. 10: the dispersion of the 4385 Å-magnitude is 0.347 mag, close to the one expected by photometry, while the Δm dispersion is smaller, at 0.171 mag. By plotting both magnitudes against color, we show that the slope is strongly mitigated for RBTL. The next section addresses this remaining correlation with color.

The wavelength range was reduced to 50% in power (ln) compared to B21a, and during the RBTL fit, the weighting was different because we only kept the stable spectral domains: these changes might lead to different RBTL parameters. These impacts would be seen throughout the paper as we compared all ZTF results with those of SNfactory, which we generated under these new conditions.

5.4. RBTL standardization

Similarly to the photometric standardization, we computed a standardization in color of the two SNe samples. To do this, we only considered the ΔA_V parameter derived from spectroscopy.

5.4.1. RBTL parameter correlations

We followed the quality cut recommendation from B21b, which is summarized Table 1. We only considered SNe with $|\Delta A_V| < 0.5$. This is the most significant cut and cuts high-color SNe on which an incorrect R_V would have a high impact. We discarded those at $z < 0.03$ to avoid significant effects from peculiar motions. After these cuts, 688 SNe Ia were left for the ZTF sample and 139 for the SNfactory sample. Fig. 9 shows that the cut of $\Delta A_V < 0.5$ is similar to the usual $c < 0.3$ in cosmology.

As mentioned Sect. 5.1, Δm magnitude corresponds to the gray offset between the (predefined) SN Ia reference spectrum $f_{\rm ref}(\lambda)$ relative to spectral observation after the color-law correction. This gray offset is averaged over the spectrum through low-dispersion areas by construction. It can thus be understood as the Hubble residuals, constructed over a bandpass defined by the $\eta(\lambda) < 0.073$ mag Boolean mask, taking advantage of spectroscopy to specifically exclude areas of high variability. In addition, we worked at peak brightness as the spectra were rescaled at maximum light. To these Hubble residuals, the associated uncertainties are relative to spectral variations, propagated during the Δm fit, mostly due to SNR and spectral uncertainties.

To verify whether a correlation with ΔA_V remains after the color-law correction, we compare the residuals Δm as a function of ΔA_V in Fig. 11. The ZTF sample has a larger number of high ΔA_V SNe than SNfactory, and these SNe show a broader dispersion in magnitude. By computing the Pearson coefficient for the two samples, we obtained $r = 0.076 \pm 0.085$ for SNfactory and $r = -0.229 \pm 0.036$ for ZTF, demonstrating that the two samples behave differently in terms of color.

This correlation between the Hubble residuals and ΔA_V could be absorbed by changing R_V, but ZTF and SNfactory present similar population characteristics, and therefore, there is no reason for the color law to be different. We instead fit for two separate slopes to correct for this correlation, and we interpret it in Sect. 7.

5.4.2. RBTL standardization method

Since the ΔA_V–Δm correlation is non-negligible, we fit for the linear relation between them with a $\beta_{\rm RBTL}$ slope coefficient. The RBTL Hubble residual of SN i is thus defined as

$$\Delta\mu_{{\rm RBTL},i} = -\left[\Delta m_i - \beta_{\rm RBTL} \cdot \Delta A_{V,i}\right]. \quad (8)$$

We fit $\beta_{\rm RBTL}$ by minimizing a total χ^2 , which procedure is described by Ginolin et al. (2025a). It takes the two parameter errors $\sigma_{\Delta {\rm m,i}}$ and $\sigma_{\Delta {\rm A_V,i}}$ into account and also the errors on RBTL magnitude and RBTL color of SN i.

We obtained $\beta_{\rm RBTL} = -0.219 \pm 0.017$ for ZTF, corresponding to CL(λ_0 = 13950 ± 700 Å) (see Sect. 5.3). The $\beta_{\rm RBTL}$ obtained for SNfactory is 0.096 ± 0.039, which is much closer to zero than ZTF, as expected from Fig. 11. Simulations matching the c distributions of the two samples showed that the $\beta_{\rm RBTL}$ discrepancy cannot be explained by the small difference between the photometric color distributions. This rather suggests an incorrect color calibration of the ZTF spectral sample, which affects the redder SNe. This spectral color systematic uncertainty affects the Δm errors and causes the increase in scatter. This hypothesis is analyzed in Sect. 7. When fitting β on the volume limited sample (see Sect. 2.1), we obtained $\beta = -0.203 \pm 0.022$. We continued with the complete sample to keep good statistics.

Another approach is to fit $\beta_{\rm RBTL}$ directly on the spectra between the lines instead of on the parameters Δm and ΔA_V. The results are consistent, and this approach does not improve

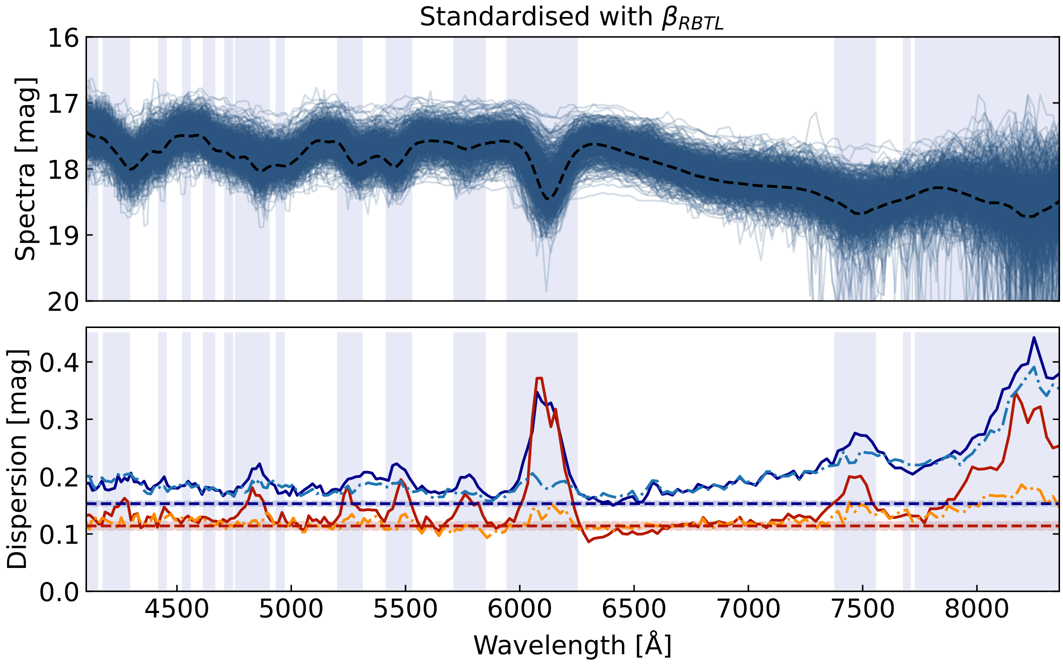

Fig. 12. Top: 688 ZTF spectra in magnitude, after color standardization with $\beta_{\rm RBTL}$. The reference flux, derived from SNfactory training, is shown with a dashed black line. Bottom: Dispersions (nMAD) in wavelength for ZTF (blue) and the 139 SNe from SNfactory (red). The dispersion (nMAD) of the scalar residuals is shown with dashed lines at 0.153 mag and 0.114 mag. The residual dispersion in wavelength after an additional line correction for ξ, described by Eq. (4), is displayed with dashed lines in light blue (ZTF) and light orange (SNfactory).

the results. Therefore, we continued with the primary approach to fit $\beta_{\rm RBTL}$ in the Δm, ΔA_V plane.

5.4.3. RBTL Hubble diagram

After we defined the standardization parameter $\beta_{\rm RBTL}$, we corrected the spectra at maximum light $f_{\max,i}(\lambda)$ for this standardization in color,

$$f_{{\rm stand},i}(\lambda) = f_{\max,i}(\lambda) \cdot 10^{+0.4 \cdot \Delta A_{V,i} \cdot [CL(\lambda) - \beta_{\rm RBTL}]}. \quad (9)$$

The standardized spectra of the ZTF sample are shown in Fig. 12 with the residual dispersion in wavelength for SNfactory and ZTF. We show the dispersion of Hubble residuals, which are defined by Eq. (8), as dashed lines. These dispersions are lower than the spectrum dispersions for both samples because of the SNR variations, which are propagated on the spectral dispersion, but are averaged in the Δm dispersion. This effect is indeed mitigated for the SNfactory results because their SNR is higher: the ZTF scalar dispersion matches the lowest dispersion values in wavelength, while for SNfactory, it matches the average of the dispersion in wavelength. The absorption line dispersion is again shown to be the main chromatic dispersion error. This highlights the advantage of spectrophotometry to select more standard parts of the spectrum through an averaged gray offset.

The resulting Hubble diagram is shown in Fig. 13. The RBTL residual dispersion is 0.153 ± 0.004 mag (nMAD) with a single color term, and the STD is 0.171 ± 0.005 mag. Of this scatter, a ~0.071 mag (nMAD ; ~0.078 STD) contribution is expected from the redshifts fitted on SN features (see Sect. 7). Another scatter floor of 0.04 mag at least is expected due to spectrophotometric inaccuracy, as explained in Sect. 5.1. We estimated the scatter of the RBTL standardization to be 0.129 ± 0.004 mag (nMAD) for the ZTF sample of 688 SNe (see the further discussion in Sect. 7). By applying the RBTL standardization for the 139 SNe of SNfactory (after standardization cuts), we obtained a scatter of 0.114 ± 0.007 mag (nMAD, and 0.145 ± 0.009 mag STD). The 0.006 mag difference to the result by B21b comes from the difference in wavelength range discussed in Sect. 5.3. All results are summarized in Table 2. We conclude that the ZTF data confirm the low Hubble residual scatter associated with this single-parameter spectroscopic standardization, which was first presented using SNfactory data (Fakhouri et al. 2015; Boone et al. 2021a,b).

We obtained a scatter of 0.163 ± 0.004 mag (nMAD) from these same ZTF SNe Ia using the SALT standardization, using the standardization parameters values estimated by Ginolin et al. 2025a. Therefore, the RBTL scatter is lower than the scatter obtained with the photometric standardization, even though it is based on fewer parameters.

Table 2. ZTF and SNf Hubble residual dispersion for several analyses in nMAD.

	SNf (139 SNe)	ZTF (688 SNe)
RBTL color standardization		
B21b	0.108	
This paper (in 4100–8400 Å)	0.114	0.153
Impact of σ_z	~0	0.071
Impact of flux-calibration	–	≥0.04
Host galaxy contamination	–	>0
Corrected dispersion	0.114	≤0.129
TE standardization		
B21b	0.083 [a]	
This paper (in 4100–8400 Å)	0.097	0.156

Notes. STDs are in the text. [a]This dispersion does not take into account the GP error, otherwise the intrinsic dispersion is 0.073 mag.

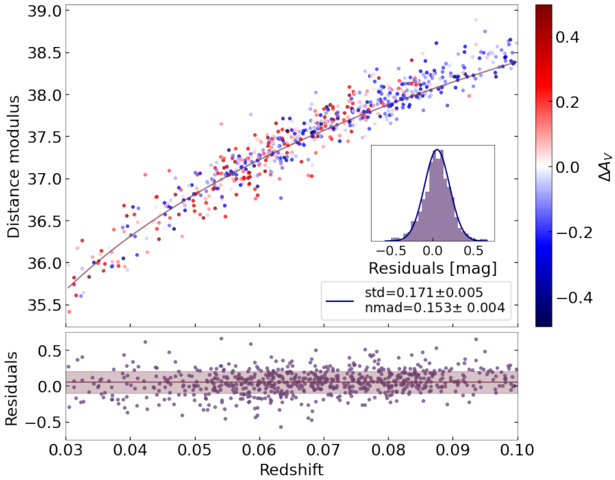

Fig. 13. Zwicky Transient Facility Hubble diagram of 688 SNe Ia (top) and their residuals (bottom), computed between the lines of high variation after a color ΔA_V standardization. The color bar at the right indicates the ΔA_V values. The brown lines correspond to the distance modulus values expected by the ΛCDM model from Planck Collaboration VI (2020) in the top plot, and zeros by definition in the bottom plot. The residual distribution is shown in the inset, and its dispersion (nMAD) is shown in the legend. This dispersion interval is colored in the bottom plot.

6. Twins Embedding

After the DTEM and RBTL, the third step to describe SN Ia dispersion is the manifold reduction introduced by B21b and summarized in Sect. 3. It consists of an Isomap parameterization that captures variabilities in absorption line regions and a GP that uses these variability parameters $\boldsymbol{\xi}$ to predict the overall spectral amplitude, that is, to standardize the SN Ia magnitude. Isomap and GP were both initially trained by B21a on 134 selected SNe from SNfactory.

In this section, we apply the method to the SNfactory dataset (139 SNe Ia) and the larger ZTF samples (688 SNe Ia). We first derive the three components of the Isomap in Sect. 6.1, and we test the associated standardization method in Sect. 6.2.

6.1. Manifold learning: Deriving the Isomap parameters

The Isomap algorithm we presented in Sect. 3 was trained by B21a on selected SNfactory data to capture the SN Ia dispersion that is not considered by the simple RBTL color + magnitude offset model. It typically corresponds to spectral features that were explicitly identified and masked in Sect. 5 to isolate the color and that contain further information on the SN Ia variability. The Isomap provided by B21a was computed on SNfactory data over the 3305–8586 Å wavelength range. To apply it on ZTF data (SEDm), we simply truncated the Isomap vectors to match its 4103–8388 Å wavelength range. We also tested retraining the Isomap on SNfactory data while limiting the data to that wavelength range, and the results were consistent. We therefore used the B21a Isomap, which we truncated and applied it to ZTF and SNfactory for self-consistent comparisons.

The distributions of the three Isomap parameters ξ_i are shown in Fig. 14. They agree between the ZTF and SNfactory samples, even though the color-calibration issue potentially affects the ZTF sample, as discussed in Sect. 5. This is expected because this Isomap was designed to capture narrow spectral features and is therefore less sensitive to the large-scale calibration

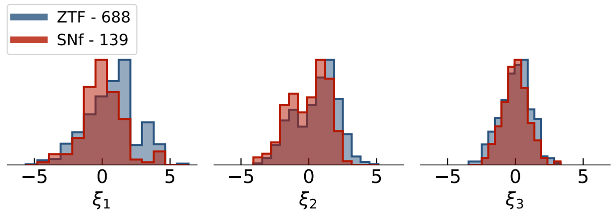

Fig. 14. Isomap coordinate distributions for the two standardization samples: ZTF (688 SNe) in blue and SNfactory (139 SNe) in red.

issue. We compared the Isomap parameters to SALT stretch x_1: as expected, ξ_2 is strongly correlated with stretch x_1, with a Pearson correlation of 0.59 ± 0.025 for ZTF and 0.83 ± 0.026 for SNfactory.

In addition to the Isomap, B21a provided a functionality (based on GPs) to convert Isomap parameters into spectral distortions[4]. The resulting spectral modeling is described by both ΔA_V and $\mathrm{GP}(\lambda)(\boldsymbol{\xi})$ in Eq. (4). As illustrated in Fig. 12, it captures some of the main SN Ia chromatic variabilities. However, the two lines that vary most in the initial Isomap training (see Fig. 6 of B21a) are not captured with ZTF data: the Ca II H&K feature is below the wavelength range, and the Ca II IR triplet is not reduced for ZTF, as shown in Fig. 12, which might be due to the flux calibration procedure. In addition, the continuum dispersion is not reduced and remains above the RBTL residual dispersion for ZTF and SNfactory: this GP was trained after the full RBTL correction, and model $\eta(\lambda)$. It was therefore not trained to reproduce the overall continuum, which is already captured by Δm.

6.2. Twins-embedding standardization

Given the Isomap parameters $\boldsymbol{\xi}$ plus the RBTL color ΔA_V, B21b trained a different Gaussian process to predict the remaining gray offsets (see Eq. (10)). Thus, the residual magnitude $\Delta\mu_{\mathrm{TE}}$, which corresponds to the standardized SN Ia Hubble residuals of the TE method, can be expressed as

$$\Delta\mu_{\mathrm{TE},i} = -\left[\Delta m_i - \beta_{\mathrm{RBTL}} \cdot \Delta A_{V,i} - \delta m^{\mathrm{GP}}(\boldsymbol{\xi}_i)\right]. \quad (10)$$

We show in Fig. 15 the TE Hubble residuals $\Delta\mu_{\mathrm{TE}}$ for SNfactory and ZTF. They are compared with those computed using only the color standardization $\Delta\mu_{\mathrm{RBTL}}$. For SNfactory, the dispersion reduction from 0.114 mag ± 0.007 to 0.097 mag ± 0.006 is smaller than the results from B21a, although compatible, as its final dispersion is 0.083 ± 0.010 mag. This might be due to the difference in wavelength range, which withdraws the two Ca II lines. These lines vary most (see B21a Fig. 6), and they therefore contain most of the variation information. The residual dispersion does not improve for the ZTF sample: the nMAD changes from 0.153 mag to 0.156 mag. This is expected because the GP conversion of $\boldsymbol{\xi}$ into magnitude offsets relies on the precision of these Isomap parameters. In addition to the fact that ZTF lacks information on the Ca II lines, the Isomap requires a high spectral quality, and the spectral uncertainty of the ZTF sample is significant. Figure 15 highlights a wider $\Delta\mu_{\mathrm{RBTL}}$ range for the ZTF, for example, above 0.2, which will not be captured by the training on SNfactory. This correction does not remain for the following steps.

[4] Conceptually, this is similar to converting x_1 into its corresponding spectral distortions in the SALT modeling (i.e., $M_1(\lambda)$), except that the $\boldsymbol{\xi}$ space is nonlinear.

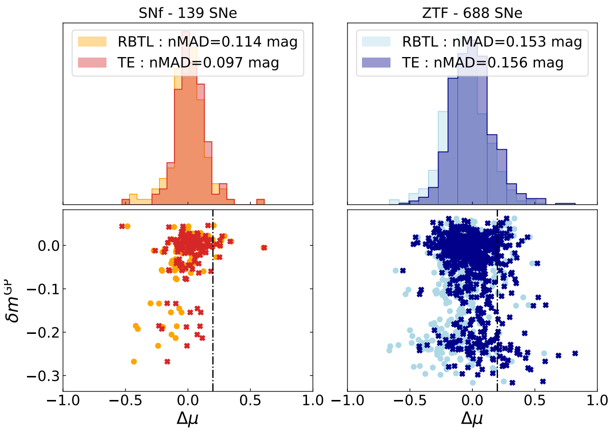

Fig. 15. Top panels: Histograms of Hubble residuals $\Delta\mu$ after RBTL standardization for SNfactory (left; light orange) and ZTF (right; light blue) and after $\delta m^{\rm GP}$ correction for SNfactory (dark orange) and ZTF (dark blue). The dispersions (nMAD) are provided in the legends. Bottom panels: Predicted magnitude $\delta m^{\rm GP}$ against the magnitude residuals. We highlight the lack of SNfactory data above 0.2 mag with a dashed black line. This region is strongly populated in the ZTF sample.

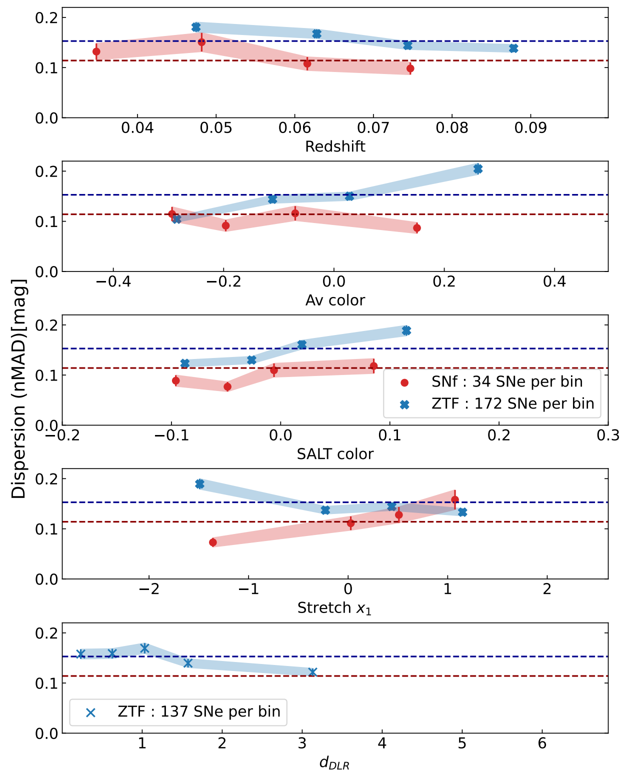

Fig. 16. Read-between-the-lines residual dispersion (nMAD) binned into four bins on the number of SNe as a function of (top to bottom) redshift, RBTL color ΔA_V, SALT color c, SALT stretch x_1, and $d_{\rm DLR}$ for the ZTF in blue, with 172 SNe per bin (or 5 bins of 137 SNe for $d_{\rm DLR}$ parameter), and SNfactory in red, with 34 SNe per bin. The x-axis limits are those of the underlying parameters. The RBTL standardization dispersion for the full samples is shown with dashed lines for ZTF (0.153 mag in blue) and for SNfactory (0.114 mag in red).

7. Discussion

In this section, we discuss the consistency of our results. We study the dependence of the RBTL standardization on SN properties in Sect. 7.1 and its dependence on the host galaxy properties in Sect. 7.2. In all results presented in this section, application of the TE standardization with $\boldsymbol{\xi}$ only increased the ZTF scatter, even for subsamples identified with a low $\Delta\mu_{\rm RBTL}$ scatter. We hence focused on studying the origin of the increased variance found in ZTF in comparison to SNfactory.

7.1. Parameter dependences

To investigate the RBTL standardization dependences on parameters such as the SN Ia redshift, the RBTL color, the LC stretch and color parameters, and the distance to the host galaxy, we binned our data into four bins with an equal sample size: 173 SNe Ia per bin for ZTF, and 34 for SNfactory. The $\Delta\mu_{\rm RBTL}$ dispersion (nMAD) as a function of the parameters is shown in Fig. 16.

First, ZTF SNe Ia with a negative stretch have a larger scatter (0.189 ± 0.01 mag) than those with $x_1 > 0$ (0.138 ± 0.007 mag). This difference likely arises because the stretch parameter was not used in the DTEM (see Sect. 3.1.1): the DTEM has a higher dispersion for low-stretch SNe Ia especially, as shown in Appendix C. The SNfactory dataset is less affected because most SNe Ia have at least one spectrum close to maximum light. Additional work on the DTEM might reduce this stretch dependence.

Second, the residual scatter increases along with the spectroscopic color ΔA_V, and the bluest ($\Delta A_V < -0.2$) bin has a scatter of ~0.1 mag, which is as low as that from SNfactory. We showed in Sect. 5.2 that the ΔA_V distribution is significantly redder for ZTF than for SNfactory, unlike that of SALT c. This suggests that a flux calibration problem causes the spectra to redden, as a left-over host galaxy signal contamination in SEDm spectra despite the `Hypergal` procedure (Lezmy et al. 2022). This is supported by Appendix D, which shows that the SNe whose color differs depending on SALT or RBTL, the latter being redder, correspond to those with the largest calibration slope correction ($a_1 > 0.5$). These SNe are shown in Fig. 4 as corresponding to be close to the host galaxy, with $d_{\rm DLR} < 1$. This systematic color error would result in a poor fit for Δm, which would consequently add dispersion to the residuals.

Third, the scatter decreases as the $d_{\rm DLR}$ increases, which supports the hypothesis that leftover host contamination scatters the residuals. No significant dependence on redshift is detected. As discussed in Sect. 5.4.3 and illustrated in Fig. 7, we expect no dependence on SNR as it is absorbed by Δm.

Finally, we studied the dependence on the redshift measurement method because ~43% of the SN redshifts were derived from SN features, with a precision of $\sigma_z = 3 \times 10^{-3}$ (see Rigault et al. 2025). We show in Appendix B that for a given sample, using these redshifts instead of galaxy redshifts adds a residual scatter of ~0.1 mag. In our sample, the residual scatter is nevertheless lower for the SNe with low-quality redshifts because they are bluer on average. Therefore, instead of removing these SNe, we estimated the additional scatter due to those redshifts by taking each SN redshift error into account. We found a scatter (nMAD) of 0.071 mag.

7.2. Host correlations

We finally studied the correlations between the Hubble residuals $\Delta\mu_{\rm RBTL}$ and host galaxy properties. Using ZTF SN Ia DR2 data, Ginolin et al. (2025a) reported a large magnitude offset, also known as the step, between the SALT2 standardized SN Ia

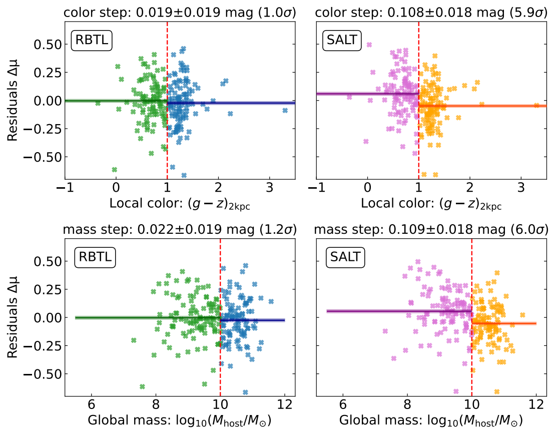

Fig. 17. Local color step (top panels) and global mass step (bottom panels) for the volume-limited 209 SNe (of 663) of the ZTF sample after RBTL standardization (left panels) and for the SALT residuals (right panels). The dashed red lines separate the two populations at $(g-z) = 1$ and $M_{\rm host} = 10^{10}\,M_\odot$. The means of the different populations and their uncertainties are shown by plain lines.

magnitude split as a function of their global host stellar masses or their local environmental colors $(g-z)$, defined within 2 kpc around the SN. When they are measured after stretch and color standardization, these steps are ~0.1 mag, and they are higher when they are fitted simultaneously (see details in Ginolin et al. 2025a).

We replicated these results using the ZTF spectrophotometric subsample and compared the step amplitude after the usual stretch and color standardization (as provided by Ginolin et al. 2025a) and after the RBTL standardization alone. As in Ginolin et al. 2025a, we limited our study to the ZTF volume-limited sample ($z < 0.06$, 219 out of the 688 ZTF SNe Ia) to cancel selection function effects, and we split the sample at $M_{\rm host} = 10^{10}\,M_\odot$ and $(g-z)_{2\,\rm kpc} = 1$ to define low and high host stellar mass or local color environments, respectively. The results are shown in Fig. 17.

We recovered the steps of Ginolin et al. (2025a) using our subsample and found the SALT2 photometric stretch and color standardization with a leftover 0.11 mag environmental step (similar for both environment tracers). However, the amplitude of the astrophysical bias was strongly reduced using the ΔA_V RBTL spectroscopic standardization, to 0.019 ± 0.019 mag for the local color and to 0.022 ± 0.019 mag for the global host mass. The two steps are two to three times smaller than their photometrically standardized counter part, and they are altogether only measured at the ~1.2σ level. These results are consistent with those previously reported by B21b using the SNfactory dataset. The host steps would be caused by the variability of strong spectral features, which are ignored by the RBTL fit.

When the volume-limited cut was removed, the steps were more pronounced: the local color step was 0.036 ± 0.010 for RBTL residuals, and it was 0.098 ± 0.009 for the SALT standardization; and the mass step was 0.036 ± 0.010 mag for the RBTL residuals, and it was 0.101 ± 0.009 for the SALT standardization.

8. Conclusion

The SEDm was designed as a typing spectrograph rather than a spectrophotometric instrument. Therefore, we first calibrated the spectra of the SN Ia DR2 sample in flux using the associated g, r, i photometric LCs, calibrated at the percent level. We computed the synthetic photometry of the spectra, compared them to the photometric data interpolated in time by SALT, and then corrected the spectra for a polynomial of second order. We estimated the final flux dispersion floor at the level of ~0.07 mag, with a contribution of a flux-calibration residual of ~0.04 mag and a contribution from low SNR. We release the final sample of 1897 flux-calibrated spectra from 1607 SNe Ia.

We have presented the first application of the TE spectrophotometric standardization method to a dataset other than the SNfactory sample. The ZTF subset contained 783 SNe after we applied the quality cuts, corresponding to about four times the SNfactory sample. The first step of TE is a spectral correction for the phase, and we estimated its accuracy using the ZTF sample to be at 0.011 mag in g band. The second step, fitting a color ΔA_V per SN on stable spectral domains, presented redder SNe for ZTF sample. A first standardization, based on this RBTL color parameter ΔA_V alone, resulted in a dispersion of 0.153 mag for the ZTF sample and of 0.114 mag for the SNfactory sample. The resulting dispersion for the ZTF sample falls at 0.129 mag when the contributions of redshift error and spectrophotometric error are taken into account. In addition, ZTF ΔA_V are significantly redder than SNfactory, while their photometric color distributions are similar. The 25% bluest ZTF SNe and the 20% farthest from the host galaxies reached an SNfactory dispersion of ~0.1 mag, indicating that the remaining host galaxy contamination within the ZTF data contributes to increase the scatter.

While the RBTL standardization was based on only one parameter, the SALT photometric standardization reached a dispersion of 0.164 mag for the same SN sample for the stretch and color correction and a dispersion of ~0.151 mag with the additional step correction. We computed host steps for the RBTL residuals, and the mass step and local color steps were both consistent with zero. Thus, a cosmological analysis using the RBTL method would be less affected by astrophysical biases than an analysis using SALT.

The third step of the TE is a nonlinear parameterization of chromatic variations based on three parameters $\xi_{1,2,3}$. The results for the ZTF spectral dataset were sufficient to explain some line variations, but not the two main lines that vary most in the TE training. The standardization based on these nonlinear TE parameters, which predicts the magnitude offset Δm per SN, is not effective on this dataset: this was expected as the instrumental error (SNR, galaxy subtraction, and flux calibration) of the ZTF spectrum dataset is large.

The Nancy Grace Roman Space Telescope, scheduled for launch in 2026, has a resolution prism of $R \sim 100$ and a wavelength range of 7500–18 000 Å. The wider wavelength range is interesting as the dust will not contaminate the signal in the infrared, and it reaches the low dispersion domain around 7000 Å for redshifts up to $z = 1.6$. The presence of atmospheric water bands, perceived by the ground-based telescopes of ZTF and SNfactory, while mitigated in the case of SNfactory, are the main contaminant of the signal at wavelengths above 7500 Å. When we focus on galaxy subtraction and redshift quality, the RBTL standardization of the spectral sample would result in a Hubble residual scatter of 0.129 mag (upper limit). In addition, the full TE standardization could be applied if the spectral data present a good spectral quality, which would further reduce the Hubble residual dispersion.

Acknowledgements. Based on observations obtained with the Samuel Oschin Telescope 48-inch and the 60-inch Telescope at the Palomar Observatory as part of the *Zwicky* Transient Facility project. ZTF is supported by the National Science Foundation under Grant No. AST-1440341 and a collaboration including

Caltech, IPAC, the Weizmann Institute of Science, the Oskar Klein Center at Stockholm University, the University of Maryland, the University of Washington, Deutsches Elektronen-Synchrotron and Humboldt University, Los Alamos National Laboratories, the TANGO Consortium of Taiwan, the University of Wisconsin at Milwaukee, and Lawrence Berkeley National Laboratories. Operations are conducted by COO, IPAC, and UW. SED Machine is based upon work supported by the National Science Foundation under Grant No. 1106171. The ZTF forced-photometry service was funded under the Heising-Simons Foundation grant #12540303 (PI: Graham). This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement n 759194 – USNAC). KM, U.B and T.E.M.B. acknowledges funding from Horizon Europe ERC grant no. 101125877. This work has been supported by the research project grant "Understanding the Dynamic Universe" funded by the Knut and Alice Wallenberg Foundation under Dnr KAW 2018.0067 and the *Vetenskapsrådet*, the Swedish Research Council, project 2020-03444. L.G. acknowledges financial support from AGAUR, CSIC, MCIN and AEI 10.13039/501100011033 under projects PID2023-151307NB-I00, PIE 20215AT016, CEX2020-001058-M, ILINK23001, COOPB2304, and 2021-SGR-01270. Y.-L.K. was supported by the Lee Wonchul Fellowship, funded through the BK21 Fostering Outstanding Universities for Research (FOUR) Program (grant No. 4120200513819) and the National Research Foundation of Korea to the Center for Galaxy Evolution Research (RS-2022-NR070872, RS-2022-NR070525). G.A. and S.P. were 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. MG is supported by the European Union's Horizon 2020 research and innovation programme under ERC Grant Agreement No. 101002652 (BayeSN; PI K. Mandel). In the development of our pipeline, we acknowledge use of the Python libraries Numpy (Harris et al. 2020), Scipy (Virtanen et al. 2020), Matplotlib (Hunter 2007), Astropy (Astropy Collaboration 2013, 2018, 2022) and Pandas (The pandas development Team 2025).

References

Adame, A. G., Aguilar, J., Ahlen, S., et al. 2025, J. Cosmol. Astropart. Phys., 2025, 021
Aldering, G., Adam, G., Antilogus, P., et al. 2002, in Survey and Other Telescope Technologies and Discoveries (SPIE), 4836, 61
Amenouche, M., Rosnet, P., Smith, M., et al. 2025, A&A, 694, A3
Astropy Collaboration (Robitaille, T. P., et al.) 2013, A&A, 558, A33
Astropy Collaboration (Price-Whelan, A. M., et al.) 2018, ApJ, 156, 123
Astropy Collaboration (Price-Whelan, A. M., et al.) 2022, ApJ, 935, 167
Bailey, S. J., Aldering, G., Antilogus, P., et al. 2009. AAS Meet. Abstr., 213, 321.05
Barbary, K., Barclay, T., Biswas, R., et al. 2016, Astrophysics Source Code Library [record ascl:1611.017]
Bellm, E. C., Kulkarni, S. R., Graham, M. J., et al. 2019, PASP, 131, 018002
Bessell, M. S. 1990, PASP, 102, 1181
Betoule, M., Kessler, R., Guy, J., et al. 2014, A&A, 568, A22
Blagorodnova, N., Neill, J. D., Walters, R., et al. 2018, PASP, 130, 035003
Bohlin, R. C., Hubeny, I., & Rauch, T. 2020, ApJ, 160, 21
Boone, K., Aldering, G., Antilogus, P., et al. 2021a, ApJ, 912, 70
Boone, K., Aldering, G., Antilogus, P., et al. 2021b, ApJ, 912, 71
Brout, D., Scolnic, D., Popovic, B., et al. 2022, ApJ, 938, 110
Cardelli, J. A., Clayton, G. C., & Mathis, J. S. 1989, ApJ, 345, 245
Childress, M., Aldering, G., Antilogus, P., et al. 2013, ApJ, 770, 108
Chotard, N., Gangler, E., Aldering, G., et al. 2011, A&A, 529, L4
Dekany, R., Smith, R. M., Riddle, R., et al. 2020, PASP, 132, 038001
DES Collaboration (Abbott, T. M. C., et al.) 2024, ApJ, 973, L14
Efstathiou, G. 2025, MNRAS, 538, 875
Fakhouri, H. K., Boone, K., Aldering, G., et al. 2015, ApJ, 815, 58
Fitzpatrick, E. L. 1999, PASP, 111, 63
Fremling, C., Miller, A. A., Sharma, Y., et al. 2020, ApJ, 895, 32
Ginolin, M., Rigault, M., Copin, Y., et al. 2025a, A&A, 694, A4
Ginolin, M., Rigault, M., Smith, M., et al. 2025b, A&A, 695, A140
Graham, M. J., Kulkarni, S. R., Bellm, E. C., et al. 2019, PASP, 131, 078001
Gupta, R. R., Kuhlmann, S., Kovacs, E., et al. 2016, ApJ, 152, 154
Guy, J., Astier, P., Baumont, S., et al. 2007, A&A, 466, 11
Guy, J., Sullivan, M., Conley, A., et al. 2010, A&A, 523, A7
Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, Nature, 585, 357
Hunter, J. D. 2007, Comput. Sci. Eng., 9, 90
Jones, D. O., Riess, A. G., Scolnic, D. M., et al. 2018, ApJ, 867, 108
Kelly, P. L., Hicken, M., Burke, D. L., Mandel, K. S., & Kirshner, R. P. 2010, ApJ, 715, 743
Kelsey, L., Sullivan, M., Smith, M., et al. 2021, MNRAS, 501, 4861
Lacroix, L., Regnault, N., de Jaeger, T., et al. 2025, A&A, submitted [arXiv:2509.04073]
Lezmy, J., Copin, Y., Rigault, M., Smith, M., & Neill, J. D. 2022, A&A, 668, A43
Masci, F. J., Laher, R. R., Rusholme, B., et al. 2019, PASP, 131, 018003
Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2011, JMLR, 12, 2825
Perley, D. A., Fremling, C., Sollerman, J., et al. 2020, ApJ, 904, 35
Perlmutter, S., Aldering, G., Goldhaber, G., et al. 1999, ApJ, 517, 565
Phillips, M. M. 1993, ApJ, 413, L105
Planck Collaboration VI. 2020, A&A, 641, A6
Popovic, B., Rigault, M., Smith, M., et al. 2025a, A&A, 694, A5
Popovic, B., Shah, P., Kenworthy, W. D., et al. 2025b, MNRAS, submitted [arXiv:2511.07517]
Riess, A. G., Press, W. H., & Kirshner, R. P. 1996, ApJ, 473, 88
Riess, A. G., Filippenko, A. V., Challis, P., et al. 1998, ApJ, 116, 1009
Riess, A. G., Macri, L. M., Hoffmann, S. L., et al. 2016, ApJ, 826, 56
Riess, A. G., Anand, G. S., Yuan, W., et al. 2024, ApJ, 962, L17
Rigault, M., Copin, Y., Aldering, G., et al. 2013, A&A, 560, A66
Rigault, M., Neill, J. D., Blagorodnova, N., et al. 2019, A&A, 627, A115
Rigault, M., Brinnel, V., Aldering, G., et al. 2020, A&A, 644, A176
Rigault, M., Smith, M., Goobar, A., et al. 2025, A&A, 694, A1
Roman, M., Hardin, D., Betoule, M., et al. 2018, A&A, 615, A68
Rubin, D., Aldering, G., Betoule, M., et al. 2025, ApJ, 986, 231
Schlafly, E. F., & Finkbeiner, D. P. 2011, ApJ, 737, 103
Scolnic, D. M., Jones, D. O., Rest, A., et al. 2018, ApJ, 859, 101
Smith, M., Sullivan, M., Wiseman, P., et al. 2020, MNRAS, 494, 4426
Smith, M., Rigault, M., Dimitriadis, G., et al. 2026, A&A, submitted
Stein, G., Seljak, U., Bohm, V., et al. 2022, ApJ, 935, 5
Sullivan, M., Le Borgne, D., Pritchet, C. J., et al. 2006, ApJ, 648, 868
Sullivan, M., Conley, A., Howell, D. A., et al. 2010, MNRAS, 406, 782
Tenenbaum, J. B., de Silva, V., & Langford, J. C. 2000, Science, 290, 2319
The pandas development Team 2025, https://doi.org/10.5281/zenodo.3509134
Tripp, R. 1998, A&A, 331, 815
Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, Nat. Methods, 17, 261

[1] IP2I Lyon/IN2P3, CNRS, Universite Claude Bernard Lyon 1, UMR 5822, F-69622 Villeurbanne, France
[2] Department of Physics, Lancaster University, Lancaster LA1 4YB, UK
[3] The Oskar Klein Centre, Department of Physics, AlbaNova, SE-106 91 Stockholm, Sweden
[4] School of Physics, Trinity College Dublin, College Green, Dublin 2, Ireland
[5] Institut für Physik, Humboldt-Universität zu Berlin, Newtonstr. 15, 12489 Berlin, Germany
[6] Physics Division, Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, US
[7] Université Clermont Auvergne, CNRS/IN2P3, LPCA, F-63000 Clermont-Ferrand, France
[8] Sorbonne Université, CNRS/IN2P3, LPNHE, F-75005 Paris, France
[9] Department of Astrophysics, University of California, Berkeley, CA 94720-3411, USA
[10] Institute of Space Sciences (ICE-CSIC), Campus UAB, Carrer de Can Magrans, s/n, E-08193 Barcelona, Spain
[11] Institut d'Estudis Espacials de Catalunya (IEEC), 08860 Castelldefels (Barcelona), Spain
[12] Institute of Astronomy and Kavli Institute for Cosmology, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK
[13] Division of Physics, Mathematics, and Astronomy, California Institute of Technology, Pasadena, CA 91125, USA
[14] Caltech Optical Observatories, California Institute of Technology, Pasadena, CA 91125, USA
[15] Department of Astronomy & Center for Galaxy Evolution Research, Yonsei University, Seoul 03722, Republic of Korea
[16] IPAC, California Institute of Technology, 1200 E. California Blvd, Pasadena, CA 91125, USA
[17] Instituto de Ciencias Exactas y Naturales (ICEN), Universidad Arturo Prat, Inquique, Chile
[18] Department of Physics, University of California, Berkeley, 366 Physics North, MC 7300, Berkeley, CA 94720, USA
[19] University of Southampton, Highfield, Southampton, England, UK
[20] The Oskar Klein Centre, Department of Astronomy, AlbaNova, SE-106 91 Stockholm, Sweden

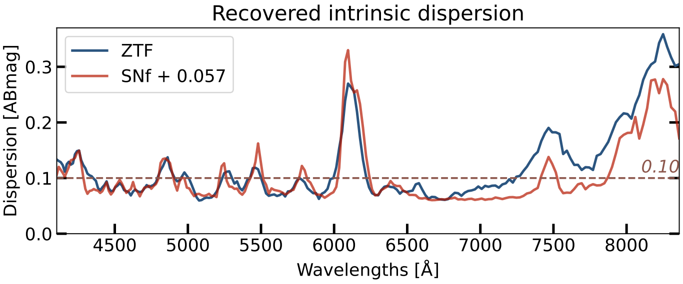

Fig. A.1. Spectral dispersion after correction for both RBTL parameters, for both 783 SNe of ZTF (dark blue) and 200 SNe of SNfactory (dark red), to which we added a 0.057 mag dispersion in quadrature.

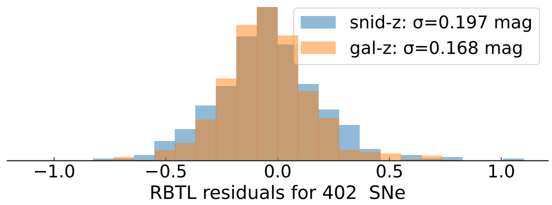

Fig. B.1. RBTL residual distributions for the same 402 SNe, using either their gal-z (orange) or their snid-z (blue).

Appendix A: RBTL spectral dispersion

To compare the residual dispersion in wavelength of SNfactory and ZTF after RBTL correction, we add the ZTF dispersion floor due to the larger SNR, quantified in Sect. 2.2, of 0.057 mag to SNfactory spectral dispersion, in quadrature, as shown in Fig. A.1. We can see that SNfactory dispersion is comparable to the ZTF one, below about 7000 Å.

Appendix B: Redshift source impact on standardization scatter

Redshift uncertainty depends strongly on the redshift source: there is either a high-quality ($\sim 10^{-5}$) redshift derived from host galaxy, which we label gal-z, or a lower quality redshift ($\sim 10^{-3}$) derived from SN features, labeled snid-z. The sample of 783 SNe on which we computed the RBTL standardization includes 402 SNe for which we have access to both redshift sources. To estimate the impact of the redshift quality on the Hubble residuals, we compute the RBTL standardization of this subsample, using either the gal-z redshifts or the snid-z redshifts. We show both residual distributions in Fig. B.1: their dispersions (nMAD) are 0.168 mag (gal-z) and 0.197 mag (snid-z). This dispersion discrepancy corresponds to a difference of 0.1 mag, in quadrature.

Appendix C: DTEM with stretch correction

In practice, B21a assumes $c'_1(\lambda) = c'_2(\lambda) = 0$ to make DTEM independant of photometric data. We now apply the DTEM taking into account the stretch to the 783 ZTF SNe:

$$m_i(\lambda, p) - m_i(\lambda, 0) = \left(c_1(\lambda) \cdot p + c_2(\lambda) \cdot p^2\right) + x_1 \cdot \left(c'_1(\lambda) \cdot p + c'_2(\lambda) \cdot p^2\right), \quad \text{(C.1)}$$

with the vectors $c'_1(\lambda)$ and $c'_2(\lambda)$, that corrects for the stretch x_1, provided by the SN Ia LCs.

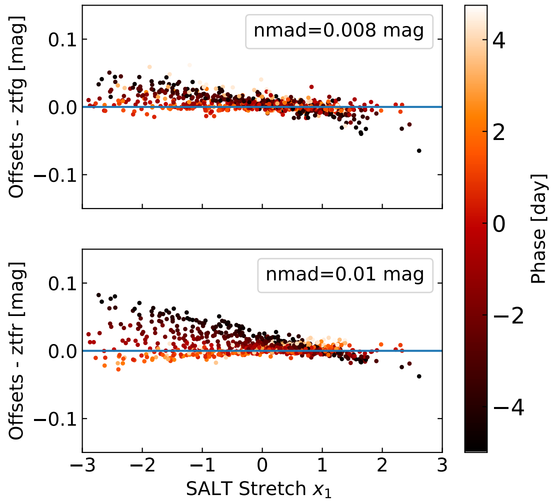

Fig. C.1. Magnitude offsets for 783 ZTF SNe, between LCs at peak and integrated ZTF spectra corrected for DTEM with stretch correction, in both g (top plot) and r (bottom plot), as a function of the SN stretch x_1. The colorbar indicates the initial phase of the spectrum. A blue line indicates the 0 value. The offsets nMAD are given in the legends.

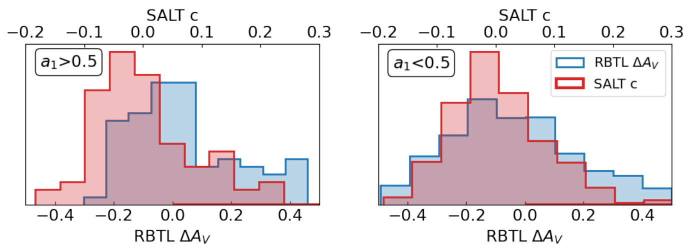

Fig. D.1. SALT color (red) and RBTL color (blue) distributions of 82 SNe with a slope coefficient $a_1 > 0.5$ (left) and 600 SNe with $a_1 < 0.5$ (right).

We integrated the corrected spectra of 783 SNe through the ZTF camera filters, as in Sect. 3.1.1, and compared the obtained photometric fluxes with those of the LCs at peak. Figure C.1 shows the offsets with the LC, in the g and r filters, as a function of the stretch x_1, and the colorbar indicates the phase. The model does not properly correct the spectra for the phase, as the offsets are not exactly zero. We see a strong correlation with the stretch, especially in r, and for low stretch. As shown in Fig. 9, SNfactory has a few low-stretch (i.e., fast) explosions; therefore, the DTEM could be less effective for those SNe. It can also be due to the quadratic evolution not being sufficient to catch all the information of fast SNe. This effect could also be reduced with the ξ_2 parameter, shown to be strongly correlated with x_1 (see Sect. 6).

Appendix D: Calibration residual effects on color

We show in Fig. D.1 the distributions of both RBTL color and SALT color for the 82 ZTF SNe with large flux-calibration correction in slope, that we define with $a_1 > 0.5$, and 600 ZTF SNe with $a_1 < 0.5$. Spectra with large slope corrections display redder RBTL color than SALT color. This suggests that the excess of high-ΔA_V SNe in ZTF could be due to the flux calibration, which adds a systematic error in color.