The Astrophysical Journal, 923:265 (20pp), 2021 December 20 https://doi.org/10.3847/1538-4357/ac30d8

SALT3: An Improved Type Ia Supernova Model for Measuring Cosmic Distances

W. D. Kenworthy[1], D. O. Jones[2,10], M. Dai[1,3], R. Kessler[4], D. Scolnic[5], D. Brout[6,10], M. R. Siebert[2], J. D. R. Pierel[7,11], K. G. Dettman[3], G. Dimitriadis[2], R. J. Foley[2], S. W. Jha[3], Y.-C. Pan[8], A. Riess[1,9], S. Rodney[7], and C. Rojas-Bravo[2]

[1] Department of Physics and Astronomy, The Johns Hopkins University, Baltimore, MD 21218, USA; wkenwor1@jhu.edu
[2] Department of Astronomy and Astrophysics, University of California, Santa Cruz, CA 95064, USA; david.jones@ucsc.edu
[3] Department of Physics and Astronomy, Rutgers, The State University of New Jersey, 136 Frelinghuysen Road, Piscataway, NJ 08854, USA
[4] University of Chicago, Kavli Institute for Cosmological Physics, Chicago, IL, USA
[5] Department of Physics, Duke University, Durham, NC 27708, USA
[6] Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA
[7] Department of Physics and Astronomy, University of South Carolina, 712 Main Street, Columbia, SC 29208, USA
[8] Graduate Institute of Astronomy, National Central University, 300 Zhongda Road, Zhongli, Taoyuan 32001, Taiwan
[9] Space Telescope Science Institute, Baltimore, MD 21218, USA
Received 2021 April 26; revised 2021 October 14; accepted 2021 October 16; published 2021 December 28

Abstract

A spectral-energy distribution (SED) model for Type Ia supernovae (SNe Ia) is a critical tool for measuring precise and accurate distances across a large redshift range and constraining cosmological parameters. We present an improved model framework, SALT3, which has several advantages over current models—including the leading SALT2 model (SALT2.4). While SALT3 has a similar philosophy, it differs from SALT2 by having improved estimation of uncertainties, better separation of color and light-curve stretch, and a publicly available training code. We present the application of our training method on a cross-calibrated compilation of 1083 SNe with 1207 spectra. Our compilation is 2.5× larger than the SALT2 training sample and has greatly reduced calibration uncertainties. The resulting trained SALT3.K21 model has an extended wavelength range 2000–11,000 Å (1800 Å redder) and reduced uncertainties compared to SALT2, enabling accurate use of low-z I and iz photometric bands. Including these previously discarded bands, SALT3.K21 reduces the Hubble scatter of the low-z Foundation and CfA3 samples by 15% and 10%, respectively. To check for potential systematic uncertainties, we compare distances of low ($0.01 < z < 0.2$) and high ($0.4 < z < 0.6$) redshift SNe in the training compilation, finding an insignificant 3 ± 14 mmag shift between SALT2.4 and SALT3.K21. While the SALT3.K21 model was trained on optical data, our method can be used to build a model for rest-frame NIR samples from the Roman Space Telescope. Our open-source training code, public training data, model, and documentation are available at https://saltshaker.readthedocs.io/en/latest/, and the model is integrated into the `sncosmo` and `SNANA` software packages.

Unified Astronomy Thesaurus concepts: Observational cosmology (1146); Type Ia supernovae (1728)

1. Introduction

Type Ia supernovae (SNe Ia) have been used as cosmological distance indicators for more than two decades, providing early evidence of the accelerating expansion of the universe (Riess et al. 1998; Perlmutter et al. 1999). Today, SN Ia distances are used at low redshift ($z \leqslant 0.15$) for distance ladder measurements of the Hubble constant (H_0; Riess et al. 2021; Freedman 2021). The Hubble constant is currently the subject of a $4-6\sigma$ tension when multiple independent late-time measurements are combined (Verde et al. 2019; Di Valentino et al. 2021), although measurements by Freedman (2021) do not find a significant discrepancy. Competitive constraints of the dark energy equation-of-state parameter, w, can be made based on SNe Ia distances measured at $z \leqslant 2.2$ (currently consistent with $w = -1$; Scolnic et al. 2018; Abbott et al. 2019; Jones et al. 2019). Recent measurements of H_0 (Riess et al. 2016), as well as most large studies of SNe across the observed redshift range for the last decade (Guy et al. 2010; Conley et al. 2011; Betoule et al. 2014; Sako et al. 2018; Riess et al. 2018; Scolnic et al. 2018; Brout et al. 2019a; Jones et al. 2019), have relied upon the SALT2 light-curve model (Guy et al. 2007, 2010) for the brightness standardization of SNe Ia in their analyses.

SN Ia distances are typically estimated by fitting their light curves with a model to determine an overall flux, a color, and one (or more) light-curve shape parameters. The apparent magnitude (as computed from the flux) is standardized with a linear combination of color and light-curve parameters (referred to as the Tripp estimator; Tripp 1998) to produce a standardized apparent magnitude relative to a fiducial SN Ia. The SALT2 (Spectral Adaptive Light-curve Template) model describes SN Ia light curves as a combination of component spectral energy distributions (flux surfaces defined in wavelength and time), multiplied by a color-dependent term described by a color law that is similar to that of the Milky Way. These components are determined through a "model training" process; the last trained model to be used in a published cosmological analysis was SALT2.4 (which we hereafter refer to as SALT2.JLA[12]), presented in Betoule et al. (2014)

[10] NASA Einstein Fellow.
[11] Now at Space Telescope Science Institute.

[12] JLA refers to "Joint Light-curve Analysis," which included the SDSS-SN and SNLS teams, and produced cosmology results in Betoule et al. (2014)

and Mosher et al. (2014), although a retrained model has recently been presented in Taylor et al. (2021). The ubiquity of the SALT2 model in cosmology analyses of the last decade can be attributed to the following:

1. The spectrophotometric model can be integrated over filter bands using the appropriate rest-frame wavelengths, which removes the need for explicit *k*-corrections.
2. The training process is cosmology-independent, as the overall normalization of a light curve is a fitted parameter.
3. The training set incorporates photometric data from multiple surveys, reducing the sensitivity of the model to the calibration of any one survey.
4. The training sample incorporates high-redshift photometric data, allowing the use of well-calibrated observer-frame optical data to extend the model into the rest-frame ultraviolet (UV).
5. Publicly available analysis tools such as `SNANA` (Kessler et al. 2009a) and `sncosmo` (Barbary et al. 2015, 2016) include infrastructure to fit light curves and generate simulated data using the SALT2 model.
6. The model has been tested by many independent groups, as a consequence of its use in cosmology analyses.

Despite these advantages, Scolnic et al. (2018) (hereafter referred to as the "Pantheon" analysis) found that the calibration of the training sample used to create the SALT2.JLA model was the largest single systematic uncertainty in their measurement of w ($\sigma_w = 0.014$, i.e., 30% of the total systematic uncertainty), though a new analysis methodology could somewhat reduce this uncertainty (Brout et al. 2021). Achieving the science goals of the Vera C. Rubin Observatory's Legacy Survey of Space and Time (LSST) will require that systematic uncertainty in the calibration of the light-curve model be decreased by a factor of 5 for the year-one analysis (The LSST Dark Energy Science Collaboration et al. 2018).

The SALT2.JLA model does not fully reproduce observed spectral features such as varying absorption line velocities, which may affect inferred cosmological distances (Wang et al. 2009a; Foley & Kasen 2011; Siebert et al. 2020). Similarly, studies have found evidence that standardized distances based on SALT2 with the Tripp estimator (SALT2+Tripp) are dependent on SN Ia host galaxy properties (Kelly et al. 2010; Sullivan et al. 2010), although the best way to characterize this effect is still in question (Rigault et al. 2013; Betoule et al. 2014; Jones et al. 2018; Rigault et al. 2020; Brout et al. 2021; Smith et al. 2020). Since these astrophysical effects are not explicitly included in the training process, differences between the training sample and cosmology samples can lead to subtle biases in the cosmology results. To characterize and correct for such untrained effects, it is essential to perform training studies on simulated data that incorporate a broad range of physical effects. Therefore, in parallel with the SALT3 development, our team has developed a more general SED-simulation framework described in Pierel et al. (2021).

Extending the wavelength range of the SALT model shows promise in the reduction of statistical uncertainty. Distance standardization based on SALT2+Tripp results in scatter about the Hubble diagram that is ~ 0.1mag larger than expected from photometric uncertainties (which we refer to as "intrinsic scatter"). Numerous studies over the last two decades have found that NIR peak magnitudes show smaller Hubble residuals than SALT2+Tripp distances (Krisciunas et al. 2004, 2007; Wood-Vasey et al. 2008; Burns et al. 2011; Mandel et al. 2011; Dhawan et al. 2018; Avelino et al. 2019; Mandel et al. 2020). However, the SALT2.JLA SED model extends only to 9200 Å (extrapolated further for use in simulations in Pierel et al. 2018), with substantial model uncertainties past ~7000 Å that preclude the use of existing low-redshift optical data. By extending the wavelength range to reliably include existing rest-frame *i*- and *z*-band photometry, we can improve on cosmology constraints on current data sets. Further extension of the model into the NIR would allow future SN Ia cosmology programs to make use of a wavelength range in which SNe Ia are intrinsically more precise.

As a step toward these goals, we have defined a SALT3 spectrophotometric model formalism and developed `SALT-shaker`, a flexible and open-source Python-based code for training a SALT3 model, accepting both spectra and photometry in the training process. The SALT3 formalism has been defined similarly to SALT2, to retain the compatibility of our model with existing analysis frameworks. As part of an overall testing and validation framework (M. Dai et al. 2021, in preparation), `SALTshaker` enables new SN Ia light-curve models to be quickly trained for new samples of cosmological SNe, allowing uncertainties from the modeling process to improve as larger and more accurately calibrated samples are collected. In Section 2, we define our SALT3 model, and describe the procedure of our training code. In Section 3, we apply `SALTshaker` on training data similar to those of the SALT2.JLA model, allowing a direct comparison of our training process to that of SALT2. Next, we add recalibrated data from past and recent SN Ia surveys to our training sample, described in Section 4, to increase the size of the training sample by a factor of ~2.5. Finally, we build the SALT3.K21 model and present our results in Section 5.

2. The SALT3 Model and `SALTshaker`

In the absence of precise physically motivated analytic or simulation-based models, the goal of phenomenological, parametric SN Ia models is to reduce the variability of the population to a limited number of parameters that minimize the unexplained variability and maximize the information available to cosmology analyses. As is well-known, SN Ia light curves show two dominant kinds of variability; a simple time-independent variation in broadband colors, and time-dependent variations first observed as a change in the rate of decline from peak in visual wavelengths (Phillips & Phillips 1993) or a "stretch" of the light curve. The latter variation has also been observed as changes in the strength and timing of a second peak in red optical and NIR bands at ~30 days after *B*-band maximum, or as changes in the phase at which the $B - V$ color is reddest (Burns et al. 2014). We use three light-curve parameters to model SN behavior; x_0 is simply defined as the overall amplitude of the light curve. The parameter c models a time-independent change in broadband colors across the entire wavelength range, while x_1 measures any time-dependent variation.

Given the three light-curve parameters for a given SN, the spectral flux[13] (in units of power per unit area per unit wavelength) at a rest-frame phase p (defined relative to *B*-band

[13] In all cases, we use F to refer to a spectral flux and f to refer to a flux integrated over an appropriate bandpass

maximum) and wavelength λ is defined as

$$F(p, \lambda) = x_0[M_0(p, \lambda; \boldsymbol{m_0}) + x_1 M_1(p, \lambda; \boldsymbol{m_1})] \cdot \exp(c \cdot CL(\lambda; \boldsymbol{cl})), \quad (1)$$

where $M_0(p, \lambda; \boldsymbol{m_0})$ and $M_1(p, \lambda; \boldsymbol{m_1})$ are flux surfaces similar to principal components, representing respectively the SED of a fiducial SN Ia and the time-dependent variability of the population. The two flux surfaces are defined on a basis of second-order B-splines, with $N_{\rm Phase} \times N_\lambda$ basis functions for each surface, with coefficients $\boldsymbol{m_0}$, $\boldsymbol{m_1}$ that are model parameters. The number of basis functions along each axis are determined based on the desired wavelength and phase resolution; in our work, we have chosen wavelength resolution 70 Å and phase resolution of 3 days. Consequently, for the trained model described in Section 5, both flux surfaces are controlled by 2520 free parameters determined by the training code. $CL(\lambda; \boldsymbol{cl})$ is a single color law that defines the effect of the color parameter c on the light curve, thought to combine the effects of some intrinsic color variation and host galaxy dust extinction. As c is intended to represent a phase-independent effect, the color law is independent of phase. The color law is polynomial with N_{CL} coefficients $\boldsymbol{cl}$ within a specified wavelength range λ_- to λ_+ and linearly extrapolated for the rest of the SED wavelength range of the model. Thus, the color law is defined as

$$CL_p(\lambda) = \sum_{i=0}^{N_{CL}} \boldsymbol{cl}_i \lambda^i,$$

$$CL(\lambda) = \begin{cases} CL'_p(\lambda_-)(\lambda - \lambda_-) + CL_p(\lambda_-) & \lambda < \lambda_- \\ CL'_p(\lambda_+)(\lambda - \lambda_+) + CL_p(\lambda_+) & \lambda > \lambda_+ \\ CL_p(\lambda) & \text{otherwise} \end{cases}$$

The objective of the model training process is the determination of the model parameters $\{\boldsymbol{m_0}, \boldsymbol{m_1}, \boldsymbol{cl}\}$ and estimation of unmodeled variability in the flux surfaces; the model training is detailed in Section 2.5.

For a given photometric observation of an SN Ia at heliocentric redshift $z_{\rm Hel}$ in a filter X with observer-frame transmission function $T_X(\lambda)$ in photon units, the broadband flux as a function of phase is

$$f_X(p) = \int T_X(\lambda/(1 + z_{\rm Hel}))F(p, \lambda; x_0, x_1, c)\lambda d\lambda \quad (2)$$

We illustrate this procedure for modeling a light curve in Figure 1. The fixed configuration parameters used to specify the model are listed in Table 1, while the fitted model parameters ($\boldsymbol{m}_0$, $\boldsymbol{m}_1$, $\boldsymbol{cl}$}) are listed in Table 2.

2.1. Model Definitions

The specification of the model given above is degenerate; for example, the scale of the flux surfaces may be changed by reducing M_0 and M_1 and increasing x_0 by the same factor. To remove degeneracies, we apply further model definitions, which are used as constraints during the training process. These definitions are arbitrary, and have been chosen to define a "fiducial" supernova whose SED is the M_0 flux surface at the mean of the observed distributions of light-curve parameters. The definitions are:

1. The rest-frame synthetic B-band flux of the M_0 component at peak is fixed such that $m_B^{\rm peak} = 10.5$ when $x_0 = 1$.
2. The rest-frame synthetic B-band flux of the M_1 component at peak is defined to be zero.
3. The distribution of the light-curve parameter x_1 in the training sample is defined to have zero mean.
4. The distribution of x_1 has a standard deviation of unity.
5. The distribution of the light-curve parameter c in the training sample is defined to have zero mean.
6. The color law is defined such that $CL(4300\,AA) = 0$ and $CL(5430\,\mathring{A}) = -1$, corresponding to central wavelengths for B and V bandpasses.
7. The distributions of x_1 and c have no correlation in the training sample.

The location and scale of the x_1 distribution and the location of the c distribution inferred for any cosmological sample are thus *relative* to the demographics of the training sample for a particular model. However, the scale of the c distribution has been normalized by definition 6. Our model definitions and SALT2 differ only in the last definition, intended to separate the phase-independent color from the phase-dependent M_1 component. The SALT2 training code does not constrain the correlation between x_1 and c, and instead fixes the V-band flux (and B-band flux following definition 2) of the M_1 component to be 0 at peak brightness. This restricts the SALT2 training code's ability to determine the best parameterization of the SN Ia population by requiring that SALT2 x_1 has no effect on observed $B - V$ color at peak. From a phenomenological perspective, it is easier to make inferences about the latent populations when the stretch and color parameters are uncorrelated, as correlated parameters imply that there is redundant information in the two parameters. Removing the correlation between the parameters x_1 and c also has a physically intuitive meaning; if the color is sourced by dust extinction, it should not depend on the processes associated with light-curve stretch.

The phenomenological c parameter is often referred to as a "color." However, c is not defined by the value of a particular photometric color at a particular phase, but rather by a fixed *effect* on the B–V color at all phases.[14] The choice of B–V color as the normalization does not restrict our color model to any particular wavelength dependence. As the color law is defined over the whole wavelength range of the model, the c parameter can be inferred using data from any wavelengths, subject to the model uncertainties derived from the data during the training process. These definitions are sufficient to define the SALT3 flux model without degeneracy, and are enforced by taking them as constraints on the fitted model parameters $\{\boldsymbol{m_0}, \boldsymbol{m_1}, \boldsymbol{cl}\}$ during the model training process.

2.2. Photometric Model Uncertainties

As Rubin (2020) and Rose et al. (2020) suggest that SN Ia SEDs are determined by 3–5 SED parameters and a color, a

[14] SALT3 predicts that the photometric color observed at some phase p between two narrowband filters centered at wavelengths λ_1, λ_2 for a supernova with light-curve parameters x_1, c will be distributed normally with mean $c \times (CL(\lambda_1) - CL(\lambda_2)) + (\frac{2.5}{\log 10})(\log(M_0(p, \lambda_2)) - \log(M_0(p, \lambda_1)) + x_1 \times (M_1(p, \lambda_2)/M_0(p, \lambda_2) - M_1(p, \lambda_1)/M_0(p, \lambda_1)) + O(x_1^2))$. The variance is determined by the color scatter defined in Section 2.2 as $(2.5/\log 10)^2 \times (k^2(\lambda_1) + k^2(\lambda_2))$.

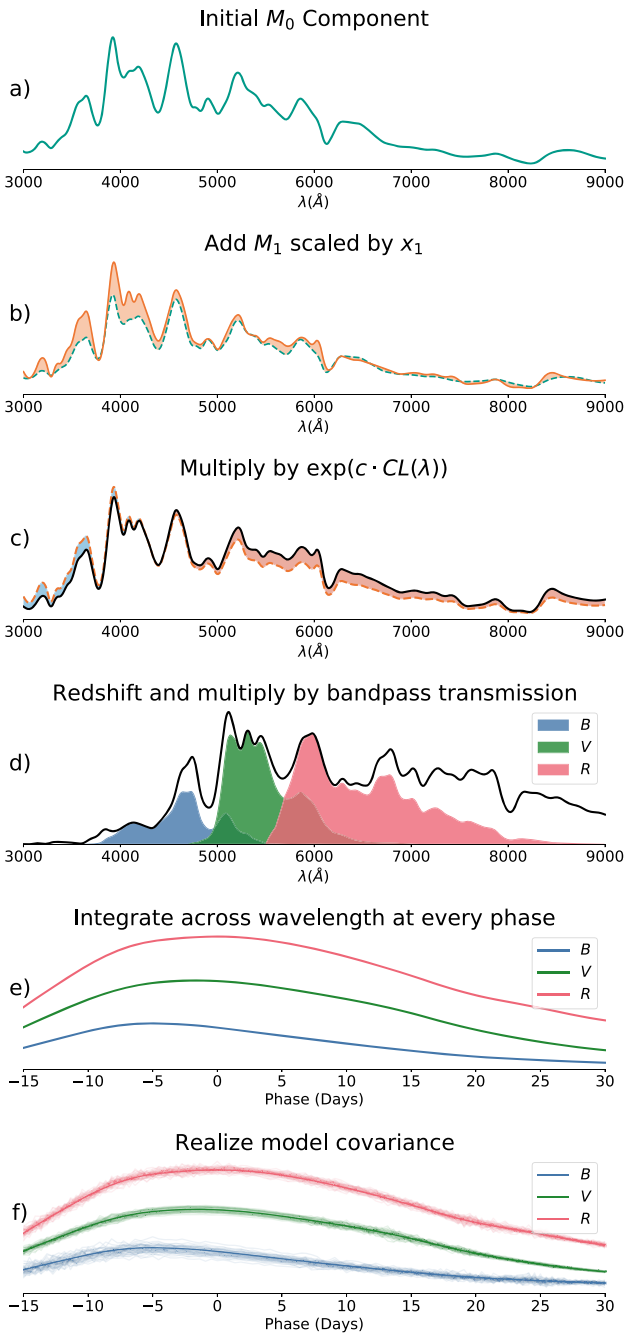

Figure 1. Construction of a SALT light curve for a given SN Ia, here shown using the SALT2.JLA model. In panels (a)–(c), we apply the M_1 component and the color law to the base M_0 spectral flux at a phase of +10 days. In panel (d), the spectrum is redshifted, and integrated across a bandpass transmission. In panel (e), this process is repeated at every phase, to produce a light curve. Panel (f) shows realizations of the model uncertainties as semitransparent lines. The effect of the color scatter is seen as as a phase-coherent offset within a given realization. This coherent uncertainty is dominant in V band around maximum light, causing smooth realizations of the light curve at this time. At later phases in V, model uncertainties uncorrelated in phase cause incoherent scatter between realizations.

model with a single SED parameter (x_1 in our case) will not capture the full diversity of the population. We refer to the variation unexplained by our model as "in-sample variance." This variation can be contrasted with the uncertainty in the model parameters due to training with a finite sample of SNe with finite signal-to-noise photometry, which we refer to as "out-of-sample variance." For a photometric observation in a filter with central wavelength λ_c and phase p, in-sample variance is addressed by a "model variance" composed of two terms. The first is a diagonal uncertainty defined as

$$\sigma_f^2(p,\lambda_c) = \left[x_0 \exp(c \cdot CL(\lambda_c)) \int T\left(\frac{\lambda}{1+z_{\rm Hel}}\right)\lambda d\lambda\right]^2 \times [\sigma_{M_0}^2(p,\lambda_c) + x_1^2\sigma_{M_1}^2(p,\lambda_c) + 2x_1 C_{M_0M_1}(p,\lambda_c)\sigma_{M_0}(p,\lambda_c)\sigma_{M_1}(p,\lambda_c)], \quad (3)$$

where $\sigma_{M_0}(p,\lambda_c;\sigma_{\rm M_0})$ and $\sigma_{M_1}(p,\lambda_c;\sigma_{\rm M_1})$ represent variability in the flux surfaces and $C_{M_0M_1}(p,\lambda_c;\boldsymbol{C_{M_0,M_1}})$ is the correlation between flux surfaces. These variance terms are described by zeroth-order B-splines (equivalent to binning the data by phase and wavelength) with eight basis functions in wavelength and 12 basis functions in phase. As detailed in Section 2.5, we use a maximum likelihood estimator to determine the B-spline coefficients $\boldsymbol{\sigma_{M_0}}$, $\boldsymbol{\sigma_{M_1}}$, $\boldsymbol{C_{M_0,M_1}}$. This is distinct from the approach of SALT2, which took the in-sample variance to have the same form as that of the out-of-sample variance, scaling the latter (evaluated by leave-one-out tests) using a smooth function or "error snake" to fix the χ^2_ν of the photometry of the training sample to unity. Because the in-sample variance is determined by variability of the underlying population of SNe light curves while the out-of-sample variance is determined by the distribution of available training data, we consider our approach to be a better account of the remaining variance of the SN Ia population.

The second term of the in-sample variance is a covariant "color scatter" that allows light curves of the same supernova in different bands to be coherently offset relative to one another. This model is similar to chromatic models of intrinsic scatter like that of Guy et al. (2010) and the diagonal terms of the covariance matrix from Chotard et al. (2011). The color scatter model describes all photometric observations made using the same bandpass as correlated, with no correlation between measurements in different bandpasses. This is quantified by a relative covariance $k^2(\lambda_c;\boldsymbol{a})$ between two fluxes, which is parameterized by the exponential of a fourth-order polynomial in wavelength, where the polynomial coefficients $\boldsymbol{a}$ are fitted parameters. Thus, the model covariance matrix for two photometric observations i and j in broadband filters X_i,X_j of a given supernova with model fluxes $\vec{f}_{\rm Model}$ is

$$k(\lambda;\boldsymbol{a}) = \exp(\sum_{i=0}^{4} \boldsymbol{a}_i \lambda^i) \quad (4)$$

$$(\Sigma_{\rm Model})_{ij} = \delta_{ij}\sigma_f^2(p_i,\lambda_{c(i)}) + \begin{cases} k^2([\lambda_c])(\vec{f}_{\rm Model})_i(\vec{f}_{\rm Model})_j & X_i = X_j \\ 0 & \text{otherwise} \end{cases}. \quad (5)$$

2.3. Modeling of Spectral Data

Although the SALT3 model is intended for use as a photometric light-curve model, spectral data are included in the training process in order to better constrain the shape of spectral features. However, given that spectral data have larger

Table 1
Fixed Model Parameters

Parameter	Description	JLA Training	SALT3 Training
$A_{\rm Spec}$	Spectral Suppression	⋯	0.75
$N_{\rm Phase}$	Number of phase basis functions for flux surfaces	20	20
N_λ	Number of basis functions on wavelength axis for flux surfaces	97	126
N_{CL}	Number of polynomial coefficients in the color law[a]	6	7
$N_{\sigma,\rm Phase}$	Number of phase basis functions for flux uncertainty surfaces	⋯	9
$N_{\sigma,\lambda}$	Number of basis functions on wavelength axis for flux uncertainty surfaces	⋯	5
$A_{\rm Phase}$	Phase gradient regularization weight	0	1000
A_λ	Wave gradient regularization weight	10	10000
$A_{\rm Dyadic}$	Dyadic regularization weight	1000	1000

Notes. Model parameters for SALT2.JLA training and SALT3.K21 training. We note that, given our adjusted regularization scheme, regularization weights are not directly comparable. We were unable to determine the exact spectral suppression in the JLA training, but from Mosher et al. (2014) we expect that this is fine-tuned for a given training sample.
[a] Note that, in the color law, there are two fewer degrees of freedom than there are polynomial coefficients, as the zeroth- and first-order coefficients are constrained to satisfy definition 6 in 2.1.

Table 2
Model and Model Uncertainty Parameters

Parameter	Category	Number	Description
$\boldsymbol{m_0}$	Flux Model	$N_{\rm Phase} \times N_\lambda$	M_0 B-spline coefficients
$\boldsymbol{m_1}$	Flux Model	$N_{\rm Phase} \times N_\lambda$	M_1 B-spline coefficients
$\boldsymbol{cl}$	Flux Model	N_{CL}	Color law
x_0	Nuisance	$N_{\rm SN}$	Overall flux normalization for each SN
x_1	Nuisance	$N_{\rm SN}$	Stretch of each SN
c	Nuisance	$N_{\rm SN}$	Color for each SN
$\boldsymbol{y}$	Nuisance	$\propto N_{\rm spec}$	Spectral recalibration
$\boldsymbol{\sigma_{M_0}}$	Uncertainty	72	Uncertainty in M_0
$\boldsymbol{\sigma_{M_1}}$	Uncertainty	72	Uncertainty in M_1
$\boldsymbol{C_{M_0,M_1}}$	Uncertainty	72	Correlation between M_0 and M_1
$\mathbf{a}$	Color Scatter	4	Color scatter

calibration uncertainties compared to broadband fluxes, we follow the SALT2 training code in "recalibrating" spectral data, modulating the model by a smooth function to match the continuum of the observed spectrum. We modify the spectral flux equation (Equation (1)) for use with spectra during the training by removing the color term and replacing it with a recalibration term of similar form,

$$F_{\rm spec}(p,\,\lambda) = x_0[M_0(p,\,\lambda) + x_1 M_1(p,\,\lambda)]\exp(\sum^{n_{\rm RC}} y_i \lambda^i/i!), \tag{6}$$

where the recalibration term is $\exp(\sum^{n_{\rm RC}} y_i \lambda^i/i!)$, and the spectral recalibration nuisance parameters y_i are fitted during the training procedure. When spectra are perfectly calibrated, the fitted recalibration term will reproduce the effect of the color parameter on the spectrum; by removing the color law entirely from the spectral flux equation, we remove any impact of miscalibrated or contaminated spectra on the color law. Any broadband or low-frequency miscalibration of the training spectra will thus have no impact on the final model. High-frequency miscalibration of the spectra or contamination from host-galaxy light may propagate biases into the final model. The quantity $n_{\rm RC}$ controls how many recalibration parameters are allowed for each spectrum, and is determined based on the wavelength extent of the spectrum and the number of filter bands available for that SN (additional filter bands better constrain the recalibration term, allowing for more free parameters). The model variances are defined similarly to the photometry, without the contribution of the color scatter:

$$\begin{aligned}\sigma_F^2(p,\,\lambda) = &\left[x_0 \exp(\textstyle\sum_i^n y_i \lambda^i/i!)\right]^2 \\ &\times [\sigma_{M_0}^2(p,\,\lambda) + x_1^2 \sigma_{M_1}^2(p,\,\lambda) \\ &+ 2x_1 C_{M_0M_1}(p,\,\lambda)\sigma_{M_0}(p,\,\lambda)\sigma_{M_1}(p,\,\lambda)].\end{aligned} \tag{7}$$

2.4. Regularization

In regions of phase-wavelength space that are poorly constrained by spectra, the M_0 and M_1 components can acquire artifacts such as high-frequency ringing, or deconvolution noise. To reduce these artifacts, we use a "regularization" procedure that penalizes large derivatives in the model where there is an absence of spectroscopic data, as parameterized by a binned spectral density function $N_{\rm eff}(p,\,\lambda)$, with every spectrum assigned equal weight. We implement three kinds of regularization: phase gradient, wave gradient, and dyadic, each of which is similar to the regularization terms of Mosher et al. (2014). These are applied to the two flux surfaces $M_0(p,\,\lambda)$ and $M_1(p,\,\lambda)$. For a flux surface $S(p,\,\lambda)$, the regularization terms are defined as

1. Phase gradient regularization penalizes large derivatives with respect to phase in less-constrained regions

$$\chi^2_{\rm Phase}[S(p,\,\lambda)] = A_{\rm Phase} \sum_i^{2N_{\rm Phase}} \sum_j^{2N_\lambda} \frac{\left(\frac{\partial S}{\partial p}\Big|_{\substack{p=p_i\\ \lambda=\lambda_j}}\right)^2}{N_{\rm eff}(p_i,\,\lambda_j)}. \tag{8}$$

2. Wave gradient regularization similarly penalizes large derivatives with respect to wavelength in less-constrained regions

$$\chi^2_{\rm Phase}[S(p,\,\lambda)] = A_{\rm Wave} \sum_i^{2N_{\rm Phase}} \sum_j^{2N_\lambda} \frac{\left(\frac{\partial S}{\partial \lambda}\Big|_{\substack{p=p_i\\ \lambda=\lambda_j}}\right)^2}{N_{\rm eff}(p_i,\,\lambda_j)}. \tag{9}$$

3. Dyadic regularization encourages the flux surfaces to be separable in phase and wavelength, and is zero when a

flux surface takes the form $S(p,\ \lambda) = g(p) \times h(\lambda)$

$$\chi^2_{\rm Dyadic}[S(p,\ \lambda)] = A_{\rm Dyadic} \sum_{i}^{2N_{\rm Phase}} \sum_{j}^{2N_\lambda} \frac{1}{N_{\rm eff}(p_i,\ \lambda_j)} \times \left[\frac{\partial S}{\partial p}\bigg|_{\substack{p=p_i \\ \lambda=\lambda_j}} \frac{\partial S}{\partial \lambda}\bigg|_{\substack{p=p_i \\ \lambda=\lambda_j}} - S(p_i,\ \lambda_j) \frac{\partial^2 S}{\partial p \partial \lambda}\bigg|_{\substack{p=p_i \\ \lambda=\lambda_j}} \right]^2. \quad (10)$$

The relative strength of the three regularization terms is determined by the weights $\{A_{\rm Phase}, A_{\rm Wave}, A_{\rm Dyadic}\}$, which are inputs to the SALTshaker code. The summations over phase and wavelength points are evenly spaced over the model phase and wavelength ranges, with twice as many points as basis functions along each axis. As regularization can bias the model surfaces by oversmoothing them (Mosher et al. 2014), we tune the weights to ensure that the regularization terms do not contribute significantly to the total χ^2. Further work to determine how this regularization scheme affects the model and to choose optimal model configurations will require applying our training and analysis framework to simulations (M. Dai et al. 2021, in preparation; Pierel et al. 2021).

2.5. Training Procedure

Based on the model definitions above, we construct the training χ^2 for the model as

$$\chi^2_{\rm Total} = \chi^2_{\rm Phot} + A_{\rm Spec} \frac{N_{\rm Phot}}{N_{\rm Spec}} \chi^2_{\rm Spec} + \chi^2_{\rm Constraints} + \chi^2_{\rm Reg}, \quad (11)$$

where $N_{\rm Phot}$, $N_{\rm Spec}$ are, respectively, the number of photometric and spectroscopic data points.

The photometric χ^2 term is

$$\chi^2_{\rm Phot} = \sum_{n}^{N_{\rm SN}} (\vec{f}^{(n)}_{\rm obs} - \vec{f}^{(n)}_{\rm Model})^T (\Sigma^{(n)}_{\rm Total})^{-1} (\vec{f}^{(n)}_{\rm obs} - \vec{f}^{(n)}_{\rm Model})$$
$$\Sigma^{(n)}_{\rm Total} = {\rm diag}((\sigma^{(n)}_{Phot})2) + \Sigma^{(n)}_{\rm Model}, \quad (12)$$

where for the nth supernova, $\vec{f}^{(n)}_{\rm obs}$ is the vector of observed photometric fluxes and $\vec{f}^{(n)}_{\rm Model}$ is the vector of the model fluxes integrated over the photometric bandpass. The covariance $\Sigma^{(n)}_{\rm Total}$ combines photometric uncertainties and the model uncertainty covariance described in Equation (5). The factor $A_{\rm Spec}$ is a constant "spectral suppression" term that down-weights the contribution of spectra to the training χ^2 in order to reduce the sensitivity of the training to unknown systematic errors in the spectral data (Guy et al. 2007). We set this term such that the spectral and photometric data have roughly equal contributions to the total training χ^2. The spectral χ^2 term is then defined

$$\chi^2_{\rm Spec} = \sum_{n}^{N_{SN}} \sum_{i}^{N_{\rm spec}} \sum_{j}^{N_{\rm points}} \frac{[F_{\rm spec}(p^{(n,i)},\ \lambda^{(n,i)}_j) - (\vec{f}^{(n,i)}_{\rm obs})_j]^2}{(\sigma_{\rm obs})^2_j + \sigma_{\rm Model}(p^{(n,i)},\ \lambda^{(n,i)}_j)^2}, \quad (13)$$

where for the nth supernova, $F_{\rm spec}(p^{(n,i)},\ \lambda^{(n,i)}_j)$ is the model spectral flux at the jth wavelength bin of the ith spectrum, $(\vec{f}^{(n,i)}_{\rm obs})_j$ is the observed spectral flux, $(\sigma_{\rm obs})_j$ is the photometric uncertainty, and $\sigma_{\rm Model}(p^{(n,i)},\ \lambda^{(n,i)}_j)$ is the model uncertainty in Equation (7). While calculating $\chi^2_{\rm phot}$ and $\chi^2_{\rm spec}$, we redden the model to account for the Milky Way $E(B-V)$ along the line of sight to each SN using the reddening law of Fitzpatrick (1999), but we do not include this term in the equations above, for simplicity. The $\chi^2_{\rm Constraints}$ is composed of penalty terms used to enforce the model definitions described in Section 2.1, and the regularization term is defined as

$$\chi^2_{\rm Reg} = \sum_{S(p,\lambda)}^{\{M_0,M_1\}} \chi^2_{\rm Phase}[S(p,\ \lambda)] + \chi^2_{\rm Wave}[S(p,\ \lambda)] + \chi^2_{\rm Dyadic}[S(p,\ \lambda)]. \quad (14)$$

The SALTshaker code is initialized with the configuration parameters shown in Table 1. SALTshaker then determines best-fit values of the model parameters. These parameters are shown in Table 2. We define flux model parameters as those that control the flux surfaces and color law, uncertainty model parameters as those that control the output model uncertainties, and nuisance parameters as those that describe individual supernovae.

SALTshaker alternates between simultaneously fitting the nuisance parameters and model flux parameters while keeping uncertainties fixed and fitting the model uncertainty parameters while keeping model fluxes fixed. The color scatter is kept fixed with the corresponding covariance term set to $k(\lambda_c) = 0$ during this stage of the training process, as we find that allowing nonzero color scatter results in biased flux surfaces due to the regularization procedure. The flux and nuisance parameters are fit using an iterative Levenberg–Marquardt algorithm. To reduce the number of (computationally expensive) Jacobian evaluations of the model residuals required, we use Schubert's method to perform a rank one update on the Jacobian after each iteration while maintaining its sparsity structure (Schubert 1970; Marwil 1979). However, this technique is unsuitable for determining the model uncertainties.

To fit the model uncertainty parameters defined in Equation (5), we define a log-likelihood ($\mathcal{L}_{\rm Phot.Errs.}$) given as

$$-2\log(\mathcal{L}_{\rm Phot.Errs.}) = \chi^2_{\rm Phot} - \sum_{n}^{N_{SN}} \log(|\Sigma^{(n)}_{\rm Total}|). \quad (15)$$

The model uncertainty parameters are chosen to maximize the log-likelihood using the optimizer iMinuit while the flux model and nuisance parameters are kept fixed (James & Roos 1975; Dembinski et al. 2020). After the training has converged with the color scatter fixed such that $k(\lambda) = 0$, we use iMinuit to fit the parameters controlling the color scatter; during this final fit, the color law is allowed to vary (having previously been fit as a flux model parameter). The out-of-sample variance is then estimated by inverting the Hessian matrix of the flux parameters obtained from the model fitting process, with the regularization terms suppressed by a factor of 100, and propagating those parameter uncertainties into the flux surfaces. The model uncertainties and out-of-sample variance are added together as the total model uncertainty surface.

The computation time required for the SALT3.JLA training sample with SALTshaker is approximately three minutes per iteration, with ~25 iterations required for convergence. On the larger SALT3.K21 sample, iterations are significantly longer, with ~25 minutes required per iteration. We note that 29-Å spectral binning, which was used in the SALT2.JLA training,

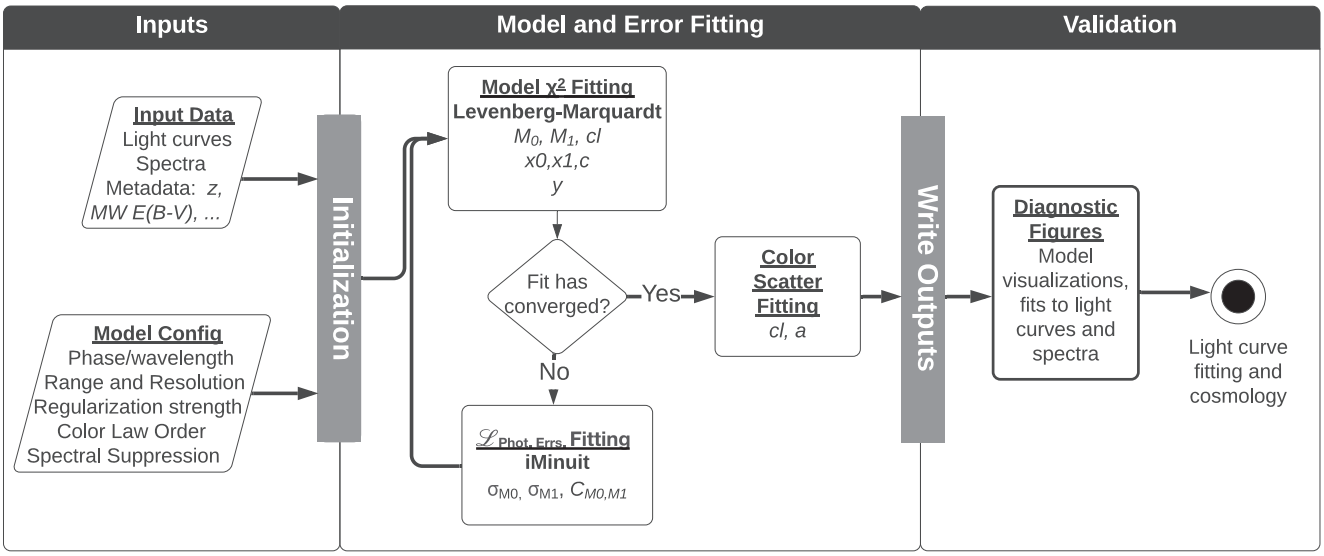

Figure 2. Activity diagram of the `SALTshaker` training process, with variable names in the model and model error fitting sections corresponding to the descriptions in Table 2.

improves speeds substantially by reducing the amount of data by up to an order of magnitude. We include this as an option in `SALTshaker`, but the models in this work use the native binning of the input spectra. Finally, the slowest component of the code is the iterative fitting of the error model, which for SALT3.K21 requires approximately 4 hours to perform 80 iterations and reach convergence through iMinuit. Error model iterations are performed every five iterations by default, but could likely be performed less often without adversely affecting the final model. Faster methods of error model fitting will be an important avenue for future improvement and should improve speeds considerably.

An overview of the `SALTshaker` training procedure is given in Figure 2.

3. `SALTshaker` Validation

Here, we show that our method is capable of producing a trained model that reproduces inferred distances from SALT2 by comparing trained SALT3 models with SALT2.JLA. We show comparisons between synthetic photometry in multiple rest-frame filters as well as the spectral models, and compare distances between the models. We describe our metrics in Section 3.1, the JLA training sample of Betoule et al. (2014) and Mosher et al. (2014) in 3.2, and a simulated sample that mimics the demographics of JLA in 3.3. We then train on these samples in Sections 3.4 and 3.5.

3.1. Validation Procedure and Metrics

To distinguish between models trained on different input data, we define "SALT3.*X*" as the SALT3 model created with `SALTshaker` using training sample *X*. We refer to the samples used to evaluate the performance of a given SALT3 model as "validation" samples, as it will be expedient to evaluate the trained model in some cases on data that were not used in the model training. We use the `SNANA` light-curve fitting program to fit validation samples with both SALT2.JLA and SALT3.*X*.

Given fitted SALT parameters m_B, x_1, c, the variances for each parameter σ_x^2, and the covariances between the parameters $\sigma_{x,y}$, the Tripp estimator for distance modulus is

$$\mu = m_B + \alpha \cdot x_1 - \beta \cdot c - \mathcal{M} \tag{16}$$

with distance uncertainties

$$\begin{aligned}\sigma_\mu = \sigma_{\rm int}^2 + \sigma_{\mu,z}^2 + \sigma_{\rm lens}^2 + \sigma_{m_B}^2 + (\alpha\sigma_{x_1})^2 + (\beta\sigma_c)^2 \\ + 2\alpha\beta\sigma_{c,x_1} + 2\alpha\sigma_{m_B,x_1} + 2\beta\sigma_{m_B,c},\end{aligned} \tag{17}$$

where $\sigma_{\rm int}$, α, β, and $\mathcal{M}$ are nuisance parameters, $\sigma_{\mu,z}$ is computed from a peculiar velocity uncertainty of 250 $\rm km\,s^{-1}$, and $\sigma_{\rm lens} = 0.055z$. We use the `SALT2mu` method (Marriner et al. 2011; Kessler & Scolnic 2017) implemented in `SNANA` to estimate the nuisance parameters as well as distances in five redshift bins. Allowing for a shift in location and scale of the light-curve parameters, the observed distributions are similar as compared to SALT2.JLA (see Section 5.2.2). We thus expect selection biases to be common between the two models, and therefore we do not use `SALT2mu` to correct for selection biases.

Given estimated distance moduli μ and Hubble residuals relative to a nominal ΛCDM cosmology ($\Delta\mu$) for both models, we define two metrics. First, RMS($\Delta\mu$) across the validation sample and the relative distance difference between models defined as

$$\begin{aligned}{\rm Diff}(\Delta_z\mu) = [&\mu(0.40 < z < 0.6|{\rm SALT3}.X) \\ &- \mu(0.40 < z < 0.6|{\rm SALT2.JLA})] \\ &- [\mu(0.01 < z < 0.2|{\rm SALT3}.X) \\ &- \mu(0.01 < z < 0.2|{\rm SALT2.JLA})],\end{aligned} \tag{18}$$

where $\mu(Z|M)$ indicates a weighted average distance across a redshift range Z given a model M. Additionally, we will show binned Hubble residuals across the redshift range. We discuss differences in the nuisance parameter β, as this has physical implications for dust properties; however, α depends on both the demographics of the training sample and our revised separation of stretch and color, and thus has no useful comparison across models (see Section 2.1). Similarly, we compare synthetic photometry, but these comparisons are most

Table 3
Data Abbreviations

Abbreviation	Size	Description	Data	Trained Model
JLA	420	Mixture of low-redshift SNe from many surveys and high-redshift SNe from SNLS and SDSS; see Mosher et al. (2014)	3.2	3.5
simJLA	420	Simulated data that emulate the JLA training sample	3.3	3.4
K21	1083	Compilation of the JLA training sample combined with new SNe from Foundation, Pan-STARRS, the Dark Energy Survey, CfA4, and CSP, with additional spectra from `kaepora`	4	5.2
K21train	541	Half of the SNe from the K21 compilation chosen at random for use as a training sample	4	5.1
K21valid	541	The other half of the K21 SNe, used as a validation sample to see how the SALT3 model performs on data that were not part of the training	4	⋯

Note. These are the abbreviations we use to refer to compilations of cosmological SNe Ia in this work, number of SNe in each, a brief description, the section in which we discuss the data itself, and the section in which we discuss a SALT3 model trained on each.

relevant for simulated data when the underlying model is known. We train on multiple simulated and real data samples to assess how our models perform; in Table 3, we summarize the abbreviations used for these training samples.

3.2. JLA Training Sample

The JLA data used to train the SALT2.JLA model consist of 420 SNe with light curves, 83 of which include spectroscopy, from a compilation of low-z samples (see Table 4 for details and citations), the Sloan Digital Sky Survey (SDSS; Holtzman et al. 2008; Kessler et al. 2009b; Sako et al. 2018), and the Supernova Legacy Survey (Astier et al. 2006, with spectra from Walker et al. 2011; Balland et al. 2018; and private communication with M. Betoule and C. Balland). These data are summarized in Table 4, along with the new training data discussed in Section 4, and are included in the data release at http://saltshaker.readthedocs.io/.

We use the "Supercal" procedure (Scolnic et al. 2015) to update the photometric calibration (including filter function central wavelengths and photometric zero points) of the JLA training sample. Supercal uses the 3π sky coverage of the PS1 system and its observations of secondary standard stars to precisely determine the offsets between different photometric systems. Scolnic et al. (2015) found that applying these calibration corrections to the sample of SN Ia used to measure the Hubble flow *without* updating the training calibration could shift w by 0.026. While we do not study the impact of the calibration on the trained model, Taylor et al. (2021) examines this with a full reanalysis of the DES-SNIa cosmology results using a retrained SALT2 model. We also use the Schlafly & Finkbeiner (2011) corrections to the Schlegel et al. (1998) Milky Way dust maps.

3.3. Simulated JLA Training Sample

To test `SALTshaker` with a known input model and cosmology, we first *simulate* the SALT2.JLA training data to produce our "simJLA" training sample. Training on the *actual* JLA training sample is also a useful test (Section 3.5), but it does not provide a known truth model to validate the training outputs. For this simJLA test, every simulated SN that goes into the training sample has SALT2.JLA as the simulated model, allowing us to test the consistency of input and output models.

We use the `SNANA` software (Kessler et al. 2009c) to generate Monte Carlo realizations of SN photometry and spectroscopy mimicking the demographics of the JLA training sample of 420 SNe. `SNANA` simulations are frequently used to simulate a random realization for a sample, in order to explore SN distance biases as a function of redshift. Our goal here is different: we aim to accurately simulate *every individual* event in the JLA sample, including cadence and signal-to-noise for both photometric and spectroscopic observations. We therefore use the cadence, redshift, and best-fit c and x_1 from each JLA light curve as input to the simulation. Each set of measured SN properties (z, t_0, x_1, c) is used to simulate a rest-frame SED with the SALT2.JLA model. Spectral variations (intrinsic scatter) are added to the rest-frame SED using the method in K13 and the covariance model from the (Chotard et al. 2011, hereafter C11) scatter model.[15] We do not use the default (Guy et al. 2010, hereafter G10), as the G10 model has nonphysical scatter at extremely red wavelengths that does not match observations. As described in Kessler et al. (2019) the simulation applies cosmological dimming, lensing, peculiar velocity, galactic extinction, and redshifting to produce a redshifted SED at the top of the atmosphere. The filter transmissions and cadence are used to determine measured fluxes and uncertainties. Finally, the noise properties of each measured spectrum are applied to the simulated SED.

This simulation does not exactly reproduce the data set, because of random fluctuations in photometric noise and intrinsic scatter—and also because the underlying SALT2 model formalism is an approximation, as discussed in Pierel et al. (2021). Nonetheless, the simulation is similar to the data and is therefore sufficient for testing `SALTshaker`, as illustrated in Figure 3 for a representative low-z SN Ia. Comparisons between the parameters of simulations and data after fitting with the SALT2.JLA model are shown in Figure 4, with only slight observed differences in average m_B uncertainty.

3.4. Training on a Simulated SALT2.JLA Training Set

We use `SALTshaker` to train a model using our simulated JLA sample, producing a model we call SALT3.simJLA. In Figure 5, we show the relative difference between the input SALT2.JLA synthetic photometry and the corresponding synthetic photometry recovered from the SALT3.simJLA model. There is significant discrepancy in the ultraviolet, where regularization strongly impacts the recovered model spectra, but at central passband wavelengths

[15] We simulate $\alpha = 0.14$ and $\beta = 3.5$. Note that the fit value of β will be lower than the simulated value by $\sim$0.6, due to the characteristics of the C11 scatter model (Scolnic & Kessler 2016).

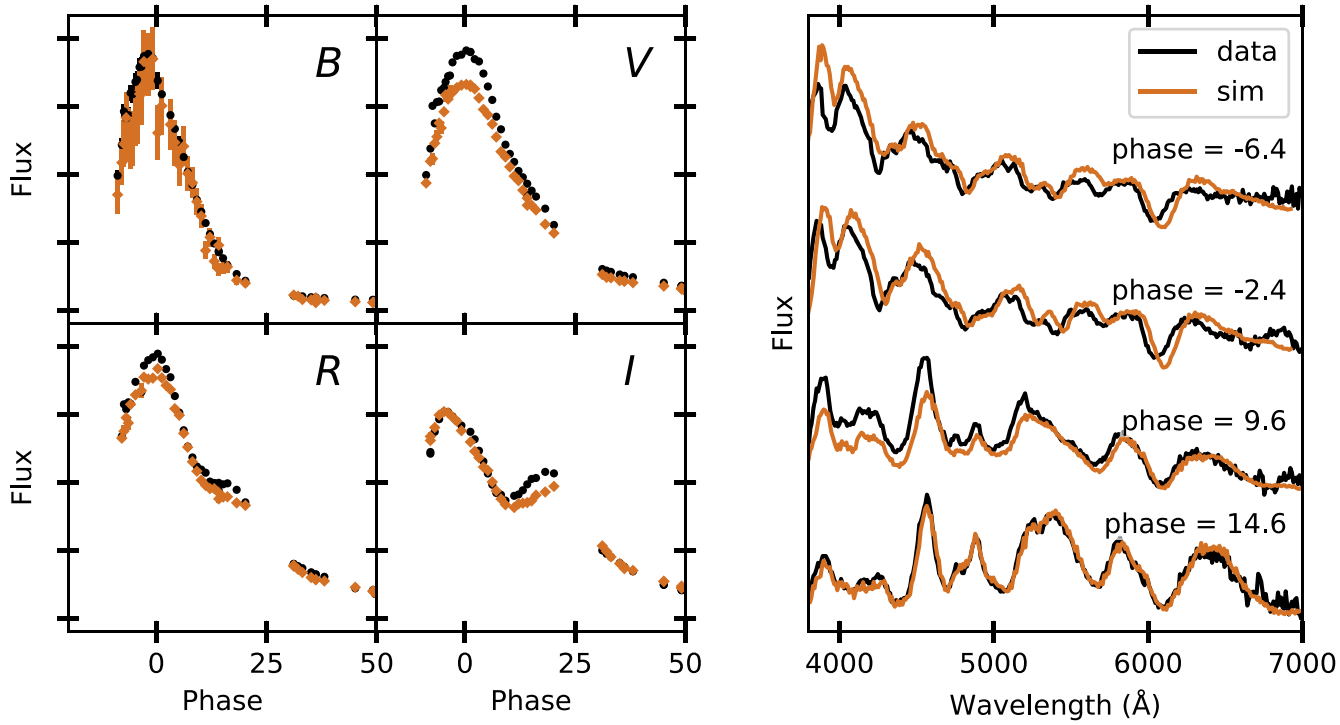

Figure 3. SN 1992A light curves (left) and spectra (right) for real data (black) and SNANA simulation (orange) generated with best-fit x_1, c, z, and time of maximum light. Random offsets between data and simulations are expected due to the C11 scatter model and spectral recalibration. The simulated spectra at red wavelengths have better S/N than the data because we model the average spectral S/N rather than its wavelength dependence.

Table 4
The K21 Compilation

Survey	N_{SN}	$N_{spectra}$[a] (JLA)	$N_{spectra}$[a] (Total)	z_{min}	z_{med}	z_{max}	Filters	References
Calan-Tololo	5	0	0	0.015	0.020	0.051	*BVRI*[b]	Hamuy et al. (1996)
CfA1	8	46	66	0.004	0.011	0.050	*UBVRI*[b]	Riess et al. (1999)
CfA2	13	108	166	0.008	0.014	0.031	*UBVRI*[b]	Jha et al. (2006)
CfA3[c]	51	31	534	0.004	0.023	0.041	*UBVRIri*	Hicken et al. (2009b)
SDSS[c]	202	0	12	0.037	0.166	0.250	ugriz	Holtzman et al. (2008)
SNLS	111	63	63	0.149	0.499	0.700	griz	Astier et al. (2006)
Misc. low-z[c]	25	152	216	0.001	0.009	0.077	*UBVRI*[b]	Jha et al. (2007)
SALT2.JLA Total	415	400	1057	0.001	0.172	0.700	...	...
CfA4[c]	30	0	0	0.009	0.029	0.070	*BVri*	Hicken et al. (2012)
CSP	13	0	36	0.011	0.029	0.058	*uBVgri*	Krisciunas et al. (2017)
Foundation	153	0	114	0.005	0.034	0.111	griz	Foley et al. (2018)
PS1 MDS[c]	266	0	0	0.026	0.297	0.630	griz	Scolnic et al. (2018)
DES-spec	206	0	0	0.078	0.362	0.850	griz	Abbott et al. (2019)
New Data Total	668	0	150	0.005	0.244	0.850	...	...
SALT3 Total	1083	400	1207	0.001	0.201	0.850	...	...

Notes. SNe Ia in the K21 compilation, after cuts: the original JLA training sample is reduced to 415 SNe (originally 420) and the new data add 668 SNe, for a total of 1083 SNe.

[a] Total number of spectra, rather than number of SNe with spectra, which is 83 for the JLA sample and an additional 297 from the new data included here. Spectra are from Filippenko et al. (1992), Wells et al. (1994), Patat et al. (1996), NAT (1997), Li et al. (2001), Salvo et al. (2001), Valentini et al. (2003), Anupama et al. (2005), Benetti et al. (2004), Kotak et al. (2005), Leonard et al. (2005), Garavini et al. (2007), Stanishev et al. (2007), Thomas et al. (2007), Foley et al. (2008), Pignata et al. (2008), Wang et al. (2009b), Foley et al. (2010), Östman et al. (2011), Walker et al. (2011), Blondin et al. (2012), Silverman et al. (2012a), Folatelli et al. (2013), and Balland et al. (2018), as well as from private communication with M. Betoule and C. Balland.

[b] We note that these filter transmission curves were not provided for these samples.

[c] Eleven SNe in these samples have additional data from the CSP (Contreras et al. 2010; Krisciunas et al. 2017), while SDSS SNe 2006oa, 2006ob, 2006on, and 2006nz also have data from CfA3 (Hicken et al. 2009b). [c]See also Rest et al. (2014) and Jones et al. (2018).

4000 Å $< \lambda <$ 7000 Å and -10 days $< p <$ 30 days, the other three light curves (B, V, and R) are consistent with the simulated model at a level better than 1%. To illustrate the impact of changes in the color law on a "typical" light curve in units of magnitudes, we use the quantity $\sigma_c \cdot \Delta CL(\lambda)$, where the standard deviation of the distribution of the SALT

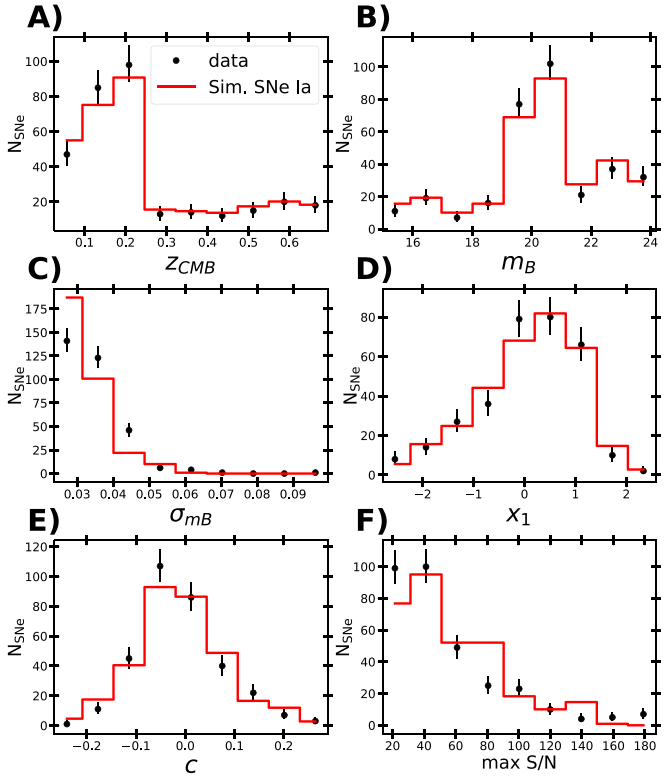

Figure 4. Distributions of parameters for the simulated (red lines) vs. real (black dots) SALT2.JLA training samples. Comparisons are shown for redshift (A), m_B (B), uncertainty on m_B (C), x_1 (D), c (E), and the maximum S/N in each SN light curve (F). All distributions are well-matched except for the uncertainty on m_B, which is slightly higher in the real data, perhaps due to the C11 model resulting in fainter SNe than the data in some bandpasses.

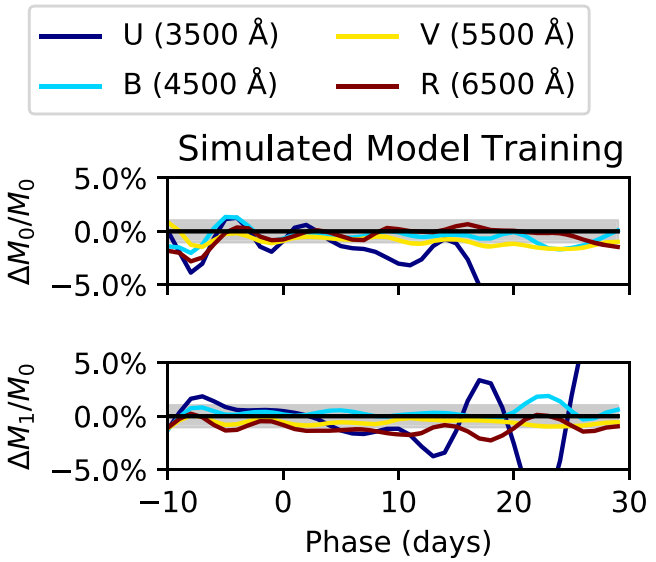

Figure 5. Relative difference between integrated model surfaces of the input SALT2.JLA simulated model and `SALTshaker` model trained on simulated data. Gray shading shows ±1%. Light curves are integrated over square *UBVR*-like bandpasses 1000 Å wide, centered at 1000 Å intervals from 3500 Å to 6500 Å. Significant discrepancies are seen in the ultraviolet, where there is sparse data and regularization schemes have largest impact. As our simulations show there is insufficient data to effectively constrain the *I*-band model, we do not show it in this figure, although the SALT2.JLA flux surfaces cover this wavelength regime.

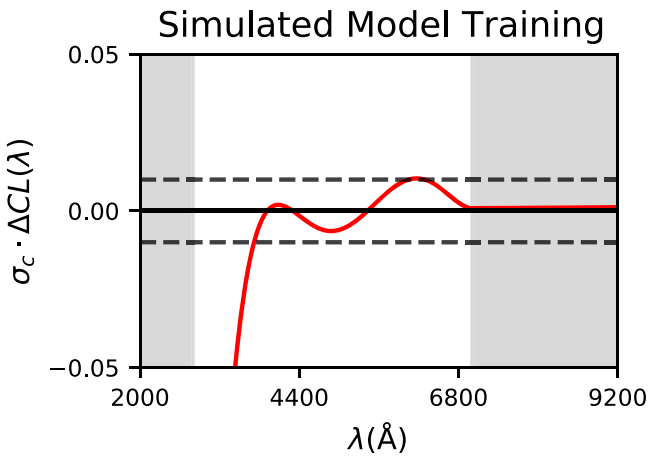

Figure 6. Change between color law of input simulated model and model trained on simulated data. We plot $\Delta CL(\lambda) = CL_{\rm trained}(\lambda) - CL_{\rm simulated}(\lambda)$, multiplied by $\sigma_c = 0.1$ to show the impact on a typical SN Ia light curve in magnitudes, as the scale of the color parameter is common between models. Dashed lines show ± 0.01mag. Gray shaded regions indicate the wavelengths for which the color law is linear. The slopes of the color law match closely at >7000 Å, leading to very small differences across the wavelength range where the color law is linear.

color parameter c is ∼0.1. As can be seen from Figure 6, the color-law difference is >0.05 mag when $\lambda < 3000\,\mathring{A}$, the regime where the color law is least constrained.

We find that the RMS of $\sigma_c \Delta CL(\lambda)$ is 0.01mag between wavelengths $3500\,\mathring{A} < \lambda < 7000\,\mathring{A}$, central filter wavelengths for which the SALT2.JLA model is considered reliable. We conclude that, because few SNe constrain the SED at $\lambda < 3000\,\mathring{A}$, the limited set of light-curve parameters c, x_1 in this wavelength region poorly constrain the color law and the spectral components. In Sections 4 and 5, we substantially expand the training data in order to address this issue.

Next, we compare distances from the two models. To avoid the statistics of the training sample limiting the precision of our validation, we simulate a "large" JLA simulation for validation, with summary statistics for the Hubble diagram fits shown in Table 5, row 1. Instead of simulating exact x_1 and c values for an "apples-to-apples" comparison with the SALT2.JLA training sample, we generate ∼3000 SNe mimicking a combination of CfA3 (Hicken et al. 2009a), SDSS, and SNLS by using x_1 and c distributions from Scolnic & Kessler (2016). We show the Hubble residuals of this simulated cosmology sample in Figure 7. Given that the Scolnic & Kessler (2016) distributions of c and x_1 are assumed uncorrelated and the training sample has demographics similar to those of the JLA sample, we can expect to recover the same α, β parameters as were measured using SALT2, which will not be true of the real data. We observe consistent measurements of β and RMS ($\Delta\mu$), with σ_{int} higher by 0.005 mag due to our treatment of uncertainties. We also observe a slightly higher value of α, which is attributed to the x_1/M_1 degeneracy. The distance difference between the models (Equation (18)) is consistent (Diff($\Delta_z\mu$) = 7 ± 11 mmag) despite the *U*-band discrepancies in the light curves seen in Figures 5 and 6. As the validation sample here has demographics similar to those of the training sample, the same sparsity of data that allows the observed ultraviolet divergences causes those same divergences to have little effect on the inferred distances. We conclude that, for a SALT model to be reliable at these wavelengths, the density of data in this region must be increased to match the density of data where we recover the input model to ∼1%, i.e., at

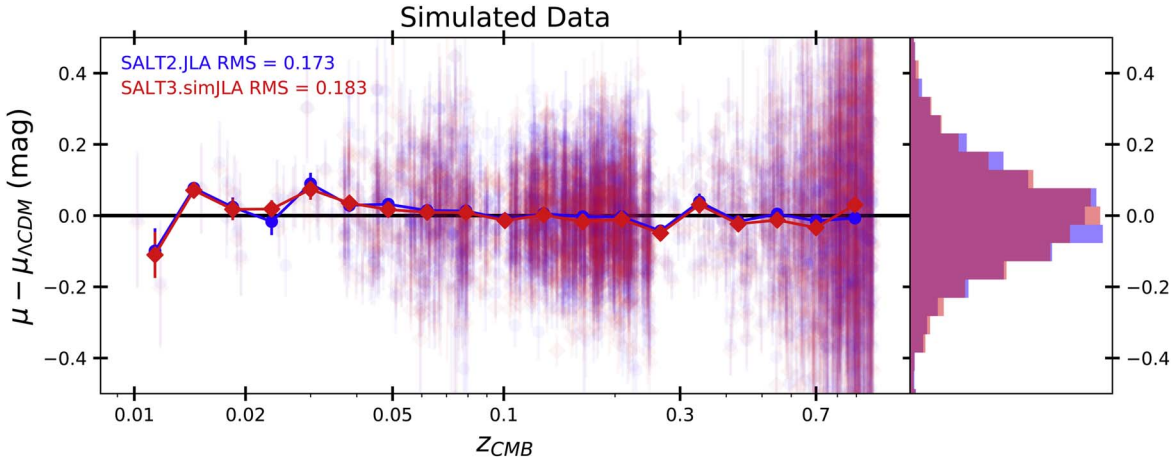

Figure 7. Hubble residual comparison with SALT2.JLA and SALT3.simJLA fits to simulated data. Each semitransparent point shows the Hubble residual (with error bar) inferred from a simulated SN light curve in the validation sample, which has been fit with either the SALT2.JLA model (blue) or the SALT3.simJLA model (red), with opaque points showing binned distances and uncertainties from each model. We do not expect SALT3 to outperform SALT2 here, because SALT2 is the simulated model, but we find consistent results.

Table 5
Summary of SALT Model Nuisance Parameters and Relative Distance Difference for Different Training Sets

	Sample			Nuisance Parameters			Hubble Flow Dist.	
	Model	$N_{\rm Training}$	Validation Sample	α	β	$\sigma_{\rm int}$	RMS ($\Delta\mu$)	Diff($\Delta_z\mu$)
	SALT2.JLA	420	large sim. JLA[b]	0.151 ± 0.003	3.101 ± 0.031	0.098	0.173	⋯
1	SALT3.simJLA	420	large sim. JLA[b]	0.165 ± 0.003	2.854 ± 0.028	0.103	0.182	-0.007 ± 0.011
	SALT2.JLA	420	JLA	0.131 ± 0.009	3.230 ± 0.114	0.128	0.153	⋯
2	SALT3.JLA	420	JLA	0.142 ± 0.011	2.952 ± 0.124	0.145	0.155	0.032 ± 0.028
	SALT2.JLA	420	K21 Valid	0.146 ± 0.007	3.016 ± 0.090	0.097	0.142	⋯
3	SALT3.K21train	541	K21 Valid	0.145 ± 0.007	2.732 ± 0.087	0.108	0.143	0.004 ± 0.021
	SALT2.JLA	420	K21 Full	0.137 ± 0.005	2.986 ± 0.060	0.103	0.142	⋯
4	SALT3.K21	1083	K21 Full	0.136 ± 0.005	2.814 ± 0.057	0.110	0.141	-0.003 ± 0.014

Notes. Cosmological results from light-curve fits using SALT3 models created with different training sets, compared to results when using the SALT2.JLA model. Training sets include: a simulated JLA training sample (simJLA, Row 1); the real JLA training sample (JLA, Row 2); an expanded training sample using the full JLA training sample but including only half the additional data we use, with the remainder used for validation (K21train, Row 3); and the full K21 compilation (highlighted in Row 4), which is the training sample used to create the SALT3 model published in this work. Our best model using the complete K21 compilation results in statistically consistent nuisance parameters and distances but slightly lower Hubble scatter, as compared to SALT2.
[a] Relative to the equivalent SALT2 fitting results, the distance between average Hubble residual at $0.01 < z < 0.2$ and the average Hubble residual at $0.4 < z < 0.6$.
[b] Large, combined simulations of CfA3, SDSS, and SNLS with a total of ∼3000 SNe to measure distance biases more precisely than is possible from a sample with the size of the JLA training sample.

least a factor of two increase (see Section 4, where we discuss the density of photometric data across the JLA wavelength range).

3.5. Training on the JLA Training Set

Next, we run `SALTshaker` on the real JLA training sample; we refer to this trained model as SALT3.JLA. We follow the validation procedure of Section 3.1 using the JLA training set for our validation sample. As shown in row 2 of Table 5, we find a similar α, a slightly lower β, and consistent RMS ($\Delta\mu$). We find that Diff($\Delta_z\mu$) = 32±28 mmag is consistent with zero. The value of $\sigma_{\rm int}$ is slightly higher for SALT3.JLA, attributable to greatly decreased model uncertainties. Given the similar RMS scatter, this $\sigma_{\rm int}$ difference does not result in reduced distance precision. The distance moduli are consistent at the 1σ level between SALT2.JLA and the SALT3.JLA, and the model surfaces are consistent within the uncertainties. We show the Hubble residuals of this model in Figure 8 (along with our extended SALT3 models discussed in Section 5).

4. Expanded Training Data: The K21 Compilation

To create a next-generation SALT model with extended wavelength range and reduced uncertainties, we add high-quality data from the Dark Energy Survey (Brout et al. 2019b), the Foundation Supernova Survey (Foley et al. 2018), the Pan-STARRS Medium Deep Survey (Scolnic et al. 2018), the Carnegie Supernova Project (Stritzinger et al. 2011), and the CfA4 Sample (Hicken et al. 2012). We discuss each of these sources of photometric data, as well as the light-curve quality cuts used to ensure that each SN in our sample is well-constrained by data, in Section 4.6. For SNe with photometric data in the training sample, we add 693 low-z spectra from the Kaepora database (Siebert et al. 2019), the majority of which originate from the Berkeley SN Ia Program (BSNIP; Silverman et al. 2012a; Stahl et al. 2020). We call this collection of data the "K21 compilation," and use our methodology to create a new SALT3 model based on this data. We show the density of photometric and spectral data in phase and wavelength space with the original JLA training sample and the additional data

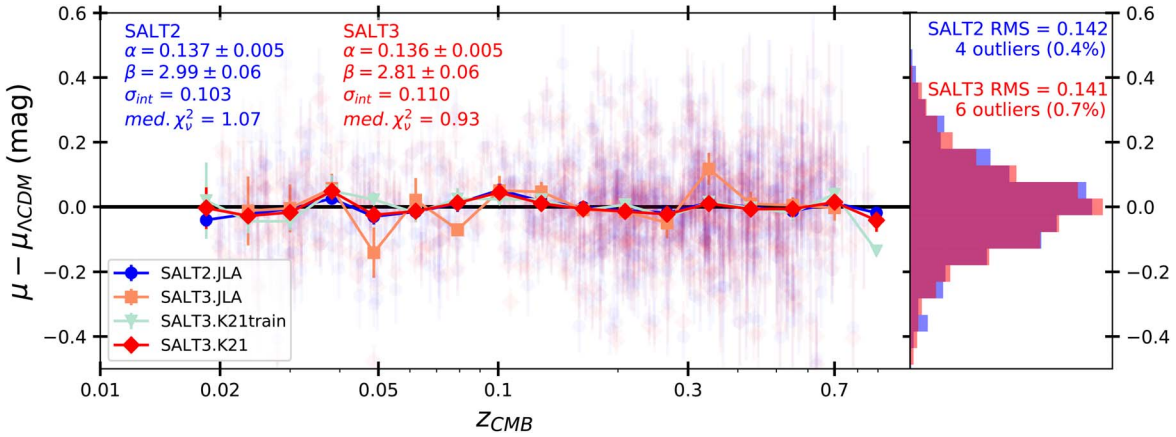

Figure 8. Comparison of the SALT2.JLA (blue) vs. SALT3.K21 (red) Hubble residuals for the full K21 compilation, with results from SALT3.JLA and SALT3.K21train shown in orange and teal, respectively. Each semitransparent point shows the Hubble residual (with error bar) inferred from a simulated SN light curve in the validation sample, which has been fit with either the SALT2.JLA model (blue) or the SALT3.simJLA model (red) with opaque points showing binned distances and uncertainties from each model. The same number of 3σ outliers are seen when using either SALT2.JLA and SALT3.K21[a]. A small change in beta can be attributed to a modified separation of color and stretch, along with improved parameter constraints on color. Additional details regarding the nuisance parameters in different training sets are shown in Table 5. (SALT3 3σ outliers are the SNe 1998ab, 05D2ci, 1995ac, 5635, 130258 and 370369. SALT2 outliers are the SNe 05D2ci, 90037, 160214, and 2002hu.)

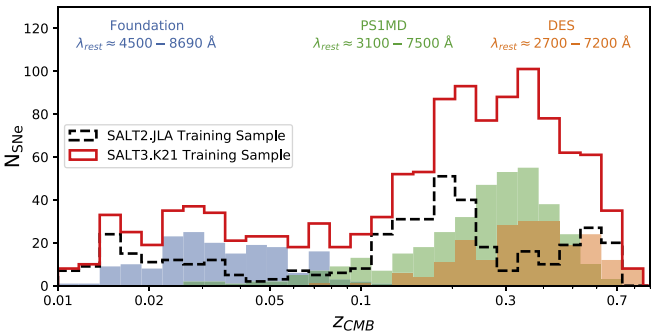

Figure 9. Redshift distribution of the JLA training data for SALT2.JLA (black dashed) and our training data for SALT3.K21 (red solid), which adds Foundation (blue), PS1 (green), and DES (orange) to the JLA training sample, as well as small samples from CSP and CfA4. At the top, we label the rest-frame central filter wavelengths of each added data set, with the full filter widths in conjunction with spectroscopy constraining the full SALT3.K21 wavelength range. The combined data set contains approximately three times as many spectra as the previous SALT2 training sample, and contains approximately 2.5 times as many SNe.

included in the K21 compilation in Figure 10. Our final training sample adds data across the phase space, but is most impactful in the red and blue regions of our wavelength range, where the JLA training data were limited. Wavelengths <3500 Å are on average covered by 1.8×more light curves, while wavelengths >7500 Å are covered by an average of 2.1× as many light curves and 5.7× as many spectra. Additional photometric data makes the distribution across phase space more uniform, where the JLA data have comparatively little data in the gap between low-redshift *B* and *V* bandpasses.

We initially characterize our performance with the extended set of data using separate samples for training and validation. We define the "K21valid" and "K21train" compilations by randomly assigning half of the supernovae to each. Finally, to produce our best model, we combine all available data to create the "K21" compilation, and summarize the data included below and in Table 4. The redshift distributions of these data are shown in Figure 9. Trained models using these compilations as training samples are discussed in Section 5.

4.1. The Carnegie Supernova Project

The Carnegie Supernova Project (Stritzinger et al. 2011) observed supernovae discovered by both amateur observers and professional surveys, with the goal of producing a stably calibrated and homogeneous set of light curves throughout the optical and NIR. Photometric observations were conducted using the Henrietta Swope Telescope and du Pont Telescope in optical (*ugriBV*) and NIR ($YJHK_s$) wavelengths.

We make use of optical observations from the second data release of the CSP (Stritzinger et al. 2011) as well as optical spectroscopic data published in Folatelli et al. (2013). Stritzinger et al. (2011) reported photometric observations of 50 objects, with median redshift of $z \sim 0.02$, of which 26 were included in the Pantheon analysis. A third data release with revised photometric calibration and more observations has been made public in Krisciunas et al. (2017); however, there is not an available cross-calibration of this data relative to other SN Ia surveys. Our compilation includes 13 SNe included in the Pantheon analysis, which additionally pass our more stringent requirements for density of photometric data. Future work plans to make use the NIR photometric observations of the CSP, as well as other available NIR data, to extend our model into the NIR (J. D. R. Pierel et al. 2021, in preparation).

4.2. The CfA4 Sample

The CfA4 sample consists of 94 SNe observed by the Center for Astrophysics Supernova Group (Hicken et al. 2012). These objects were discovered by a mixture of amateur and professional supernova searches. Photometric observations were made in the *uUBVri* bands using the F. L. Whipple Observatory and the KeplerCam instrument. Depending on photometric band, there are an average of ∼10 points in each light curve.

A total of 41 SNe from the CfA4 survey were included in the Pantheon analysis, and 30 pass our more stringent cuts for light-curve quality and are included in our compilation of training data.

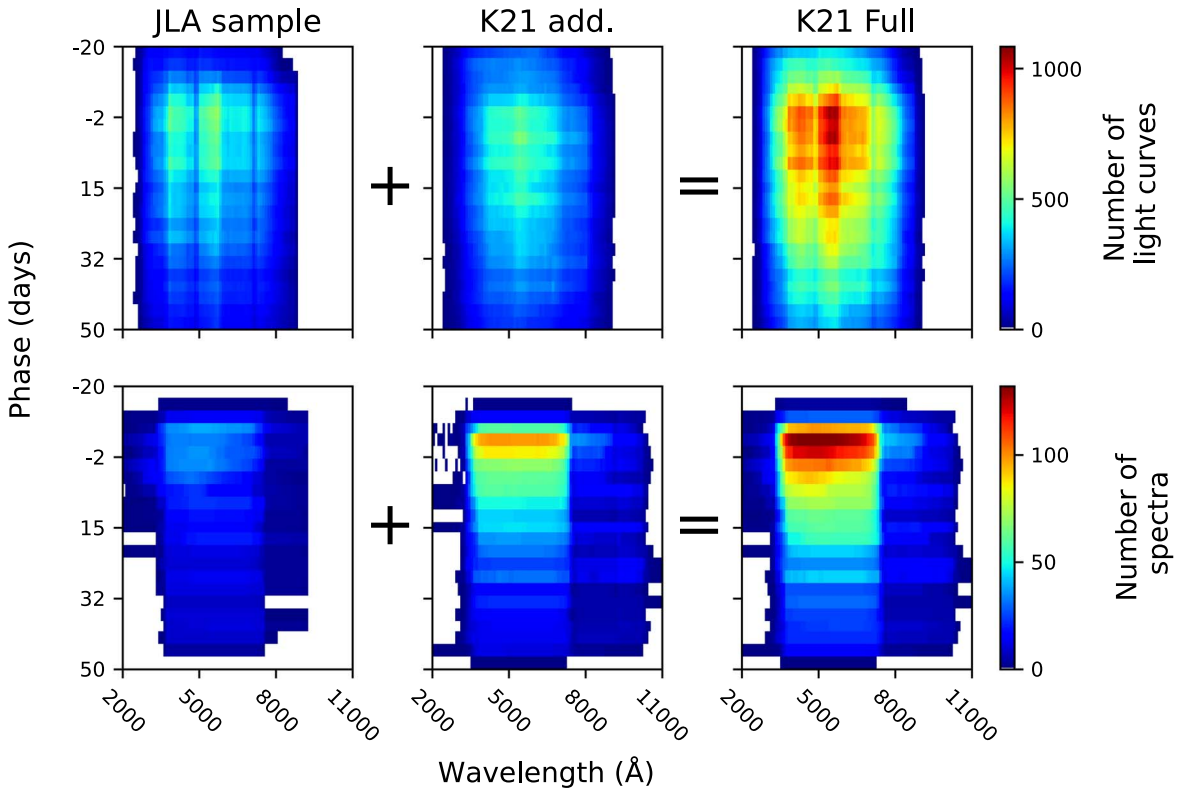

Figure 10. Number of light curves or spectra constraining each bin in our phase/wavelength space. Panels show coverage of photometric data (upper panels) and spectral data (lower panels) in the JLA training sample (left panels), the PS1, DES, and Foundation data added in the K21 compilation (middle panels), and the complete K21 compilation (right panels). A bin is considered "covered" by a given light curve if it is within the FWHM of the rest-frame bandpass. Foundation photometry provides much more extensive photometric coverage of red wavelengths, while PS1 and DES photometry provide additional rest-frame blue photometry, and the new photometric data cover the phase space more uniformly than JLA alone. Similarly, spectra from `kaepora` are immensely impactful at wavelengths $\lambda > 7500$ Å, greatly assisting the deconvolution of z-band data.

4.3. The Foundation Supernova Survey

The Foundation Supernova Survey (Foley et al. 2018) followed SNe using the Pan-STARRS1 telescope, and measured well-calibrated SN light curves in griz filters with a five-day cadence near maximum light and an approximately eight-day cadence beginning at +10 days after maximum light. To achieve reduced selection effects compared to previous surveys that targeted bright, preselected low-z galaxies, Foundation primarily followed SNe discovered by untargeted surveys such as the All-Sky Automated Survey for Supernovae (Shappee et al. 2014), the Asteroid Terrestrial-impact Last Alert System (Tonry et al. 2018), Gaia (Gaia Collaboration et al. 2016), and the Pan-STARRS Survey for Transients (Huber et al. 2015).

The Foundation first data release in Foley et al. (2018) contains 225 SNe Ia, 180 of which are cosmologically useful, and 153 of which pass our own light-curve quality cuts. These data have been used to measure cosmological parameters in Jones et al. (2019) and the correlation of host galaxy properties with SN distances in Jones et al. (2018). The *iz* band coverage of Foundation is particularly critical to creating a SALT3 model that is trained to redder wavelengths than enabled by the JLA data alone. We include spectra for 114 Foundation SNe from the Dettman et al. (2021) data release.

4.4. The Pan-STARRS Medium Deep Survey

The Pan-STARRS medium deep survey covered 70 square degrees of sky over four years, discovering approximately 5200 SNe (Jones et al. 2017; Villar et al. 2020) and spectroscopically classifying ∼10% of these at a median redshift of ∼0.35 (Rest et al. 2014). The Pantheon analysis, which combined these data with JLA, includes 279 PS1-observed SNe Ia with an average of approximately 6 observations per 10 days in griz (Scolnic et al. 2018). We use the 266 of these SNe Ia that pass our light-curve quality cuts in our training data.

4.5. The Dark Energy Survey

The Dark Energy Survey (DES) three-year spectroscopically classified SN sample contains 207 SNe Ia at a median redshift of 0.36 (Abbott et al. 2019) (of which 206 pass our light-curve quality cuts). These SNe were discovered by imaging eight 2.7 deg^2 "shallow" fields (depth ≈ 23.5 mag) and two 2.7 deg^2 "deep" fields (depth ≈ 24.5 mag) approximately once per week (Smith et al. 2020b). Transients were discovered using a difference-imaging pipeline (Kessler et al. 2015), and final photometry was performed with a "scene modeling" pipeline described in Brout et al. (2019b). See Abbott et al. (2019) and Brout et al. (2019a) for additional details regarding the DES SN Ia data and analyses. These data have a maximum redshift of ∼0.85 and complement SNLS in probing rest-frame near-UV wavelengths with well-calibrated (subpercent) photometric data (Burke et al. 2018).

4.6. Sample Selection Cuts

To ensure that SNe are suitable for inclusion in a training sample, we first require that every SN in the compilation is a

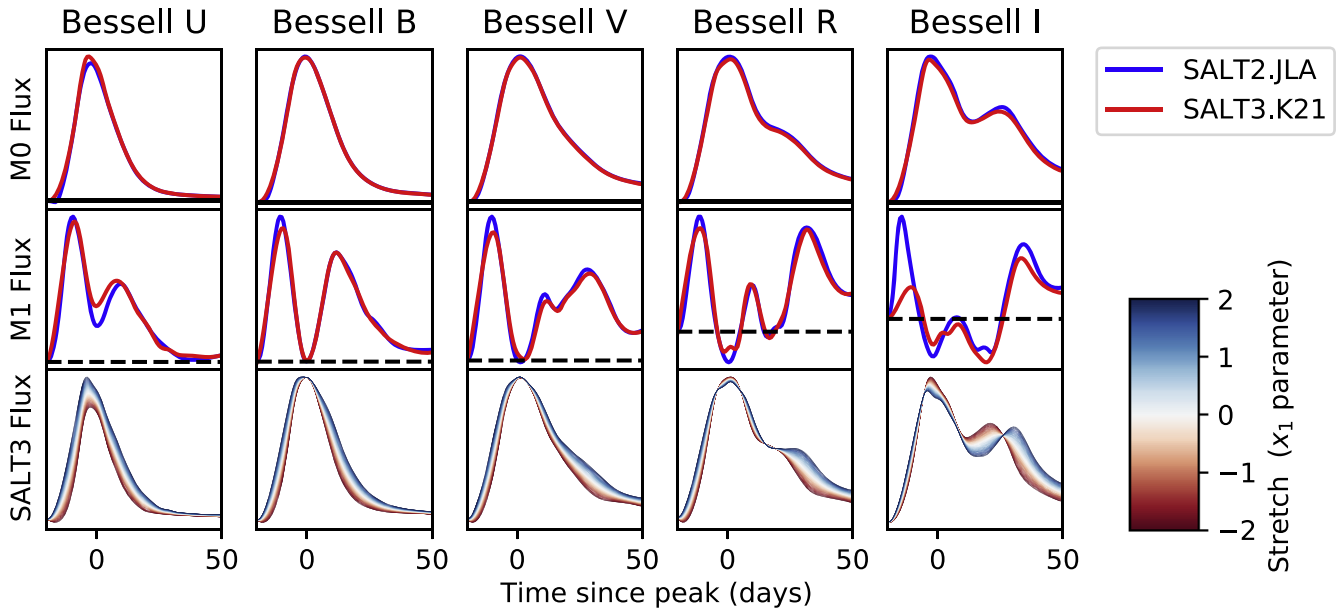

Figure 11. Comparisons of synthetic light curves. Upper two panels show light curves from both SALT2.JLA and SALT3.K21 in arbitrary units of flux. Differences are most prominent in Bessell I band. Lower panels show a family of SALT3 light curves created by varying the x_1 parameter.

spectroscopically classified, Branch-normal SN Ia (see Branch et al. 1993; we also include 1991T-like SNe Ia, following the original SALT2 training procedure of Guy et al. 2007). For the SN light curves, we make the following selection requirements (cuts):

1. At least four epochs at phases between $-10 < p < 35$ days, where p is the rest-frame phase relative to time of B-band peak.
2. At least one measurement after peak brightness ($5 < p < 20$), to constrain the shape.
3. At least one measurement in each of at least two filters at $-8 < p < 20$, to constrain the color of the SN.
4. At least one measurement prior to peak brightness ($-10 < p < -1$), to ensure a well-measured time of maximum light. This is the only cut that was not included in the original SALT2 training[16].

For the spectra, we include the original SALT2.JLA training spectra in addition to spectra taken from the `kaepora` database and spectra taken as part of the Foundation Supernova Survey. A number of these spectra were taken by the Foundation team, but most have been published on the Transient Name Server[17] as classification spectra.

To ensure minimal host-galaxy contamination in the high-z SN spectra used for model training, Guy et al. (2007) fit the spectra with a combined model including the predicted SN spectrum, the spectral recalibration parameters, and a galaxy model (elliptical, S0, Sa, Sb, and Sc templates). They removed spectra for which there was evidence for host-galaxy contamination at the 68% confidence level.

For Foundation and the additional low-z SN spectra included here, low-z SNe are much brighter relative to their host galaxies than at the redshifts probed by SNLS and SDSS. We clip host galaxy lines, mask regions with uncorrected telluric features, and remove excessively noisy or poorly calibrated regions of each spectrum, but do not attempt to subtract a host galaxy continuum. We remove a handful of spectra with poor quality from visual inspection. After cuts, there are 114 spectra from Foundation and 693 from Kaepora, a subset of which are shown in Section 5. All but five Foundation SNe have redshifts measured from host galaxy features. The complete training data are available at https://saltshaker.readthedocs.io.

5. The SALT3 Model: Extending the Wavelength Range and Training on Pantheon, Foundation, and DES Data

Having demonstrated the effectiveness of `SALTshaker` on the original JLA training data, we now train a SALT3 model using additional data, extending the free model parameters in `SALTshaker` in three ways. First, we extend the SED wavelength range to 11,000 Å so that rest frame filters centered at wavelengths up to ∼8500 angstroms, such as the Foundation z-band, can be fit with the model. Second, we extend the SED wavelength range over which the color law is fit with a polynomial (Equation (2)) to 8000 Å, just below the central wavelength of the PS1 z band. We use a sixth-order polynomial (with five effective degrees of freedom), rather than the fifth-order polynomial used for SALT2.JLA, to model the color law over the increased wavelength range with ×2.5 more training data. Third, the color scatter model is changed from a third- to a fourth-order polynomial, to allow additional flexibility over the increased wavelength range.

5.1. Validation on Extended Wavelength Range

Before presenting our model trained on all data in Section 5.2, we check the performance on separate training and validation samples. We take the "K21valid" compilation as our cosmology sample, using our "SALT3.K21train" model as described in Section 3.1, and show the resulting Hubble residuals in Figure 8 and in row 3 of Table 5. We find that nuisance parameters are similar between SALT2.JLA and the SALT3.K21train model, with $\sigma_{\rm int}$ slightly higher by 0.01 mag

[16] Five of 420 SNe from the JLA sample are removed with this cut.
[17] https://wis-tns.org/

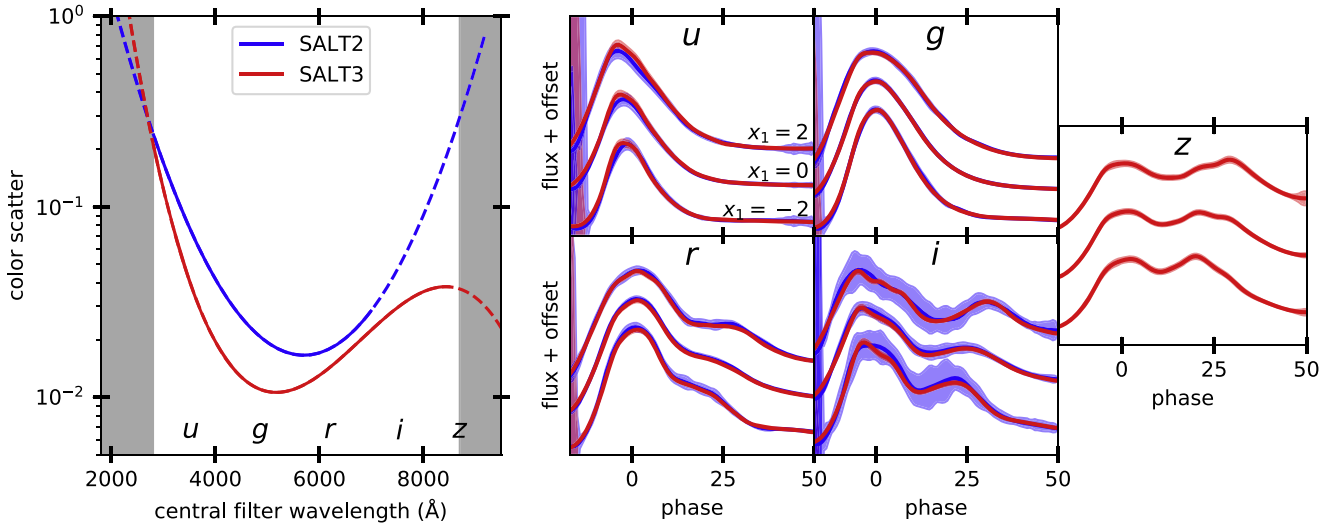

Figure 12. Color scatter as a function of central filter wavelength (left) and example light curves (right) with model uncertainties from SALT2.JLA (blue) and SALT3.K21 (red). SALT3 has comparable errors in the u and g bands, but much smaller uncertainties in the r and i bands, due to training data with much better coverage at those wavelengths, especially at large or small x_1 (with errors in both models blowing up at very early phases). Only the SALT3 model covers the z band (right-most panel). On the left, dashed lines illustrate the wavelengths where the color scatter is unconstrained by data from the JLA sample (blue) or the K21 compilation (red; gray shading).

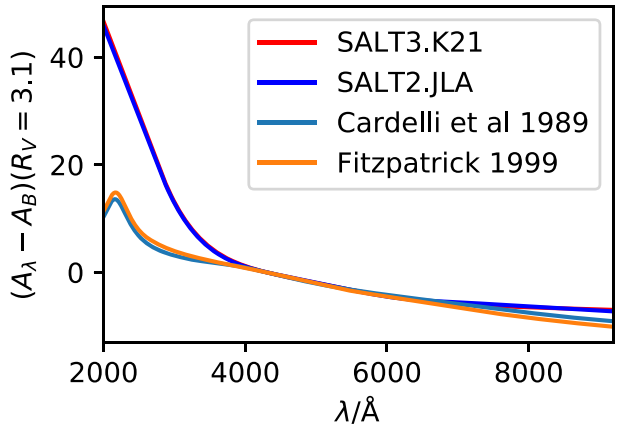

Figure 13. Comparison of SALT2.JLA vs. SALT3.K21 color law, along with the extinction curves of Cardelli et al. (1989) and Fitzpatrick (1999) for comparison. All four curves are normalized with R_V fixed to 3.1.

but with a consistent RMS($\Delta\mu$) (the SALT3 RMS is negligibly smaller). Diff($\Delta_z\mu$) is consistent with zero at the mmag level.

5.2. Training on Complete K21 Compilation

Finally, we train our best SALT3 model, which we call SALT3.K21, using all of the data described in previous sections as a training sample. The sample includes 1083 SNe, a factor of 2.5 more SNe than the JLA training sample, and 1207 spectra, a factor of three increase in the number of spectra. Synthetic light curves from SALT3.K21 and SALT2.JLA are compared in Figure 11, and the model uncertainties are compared in Figure 12. We see good consistency between SALT2.JLA and SALT3.K21, with modest differences in the u-band and some additional differences in redder bands; at both wavelength ranges, we have substantially increased the training sample (see Figure 10). Similarly, as shown in Figure 13, the color law is consistent with that of SALT2.JLA to within 1% across the entire wavelength range. As has previously been noted, the inferred color law does not resemble observed reddening laws determined from Milky Way extinction measures such as those of Fitzpatrick (1999) and Cardelli et al. (1989) in the ultraviolet and red optical wavelengths, as might be expected if SN Ia color variations are caused exclusively by host-galaxy dust similar to that of the Milky Way.

We compare Hubble residuals of the SALT3.K21 and SALT2.JLA models in Figure 8. Individual standardized distances are consistent to 0.05 mag between the two models, and the effects on the Hubble diagram are found in row 4 of Table 5. Diff($\Delta_z\mu$) is consistent with zero at 3 ± 14 mmag. Finally, row 4 of Table 5 shows nuisance parameters and Hubble diagram metrics, demonstrating that `SALTshaker` produces a new SALT3 model with slightly lower total dispersion, consistent $\sigma_{\rm int}$, and consistent distances.

We note that the β parameter is lower by 0.17 in the SALT3 model, likely due to a reconsidered separation of color and stretch. The nuisance parameter β has been physically interpreted by comparison to $R_B = R_V + 1$, the selective-to-absolute extinction ratio of a dust reddening law, finding that it is lower than expected from the typical Milky Way $R_V = 3.1$. These comparisons have been and continue to be informative, as they demonstrate that either the dust reddening laws of SN Ia host galaxies are distinct from that of the Milky Way, or that circumstellar dust may play some role in the reddening of SN Ia (see also Jha et al. 2007; Chotard et al. 2011, Burns et al. 2014; Amanullah et al. 2015; Brout et al. 2021). As we believe that the previous SALT model had confused the parameter c with some intrinsic color sourced by the parameter x_1 under the assumption that x_1 could have no effect on peak color of SN Ia, a different value of β is expected. As the β we observe is still inconsistent with expected R_B values from the mean reddening of the Milky Way, the aforementioned conclusions are not substantially affected.

5.2.1. Model Uncertainties and Hubble Scatter of SALT3.K21

The uncertainties in Figure 12 show that both the color scatter and light-curve uncertainties in redder bands are much

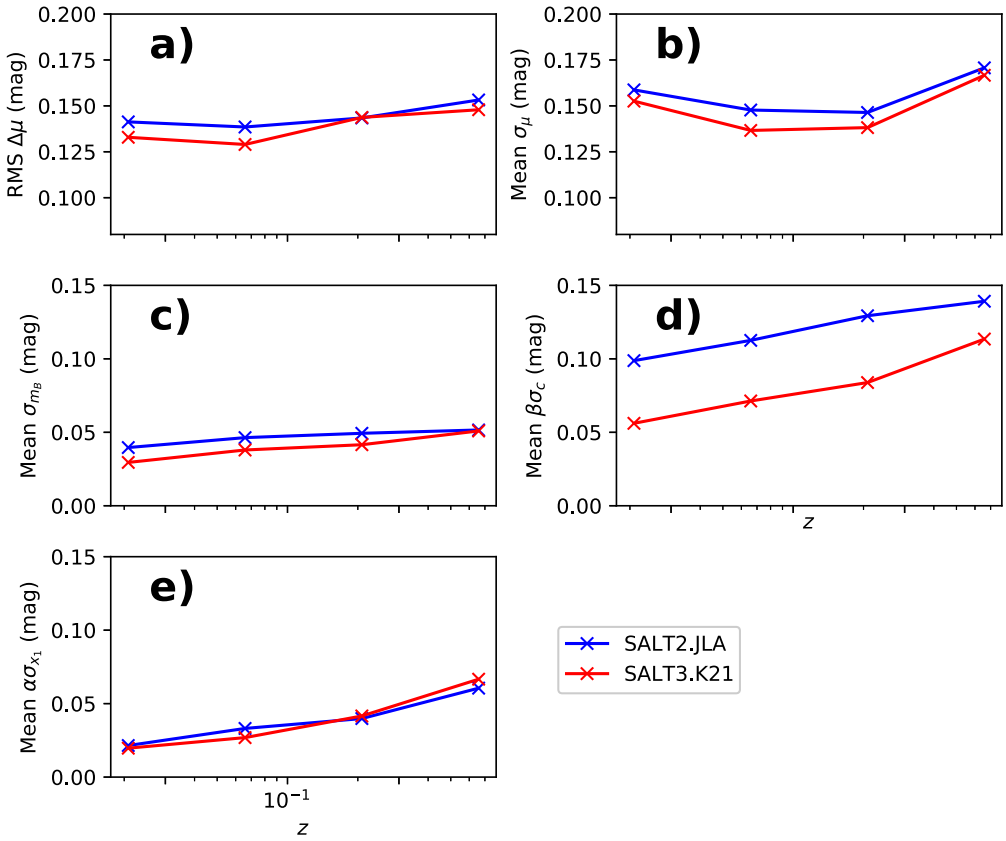

Figure 14. RMS of Hubble residuals (panel (a)) and quadratic means of uncertainties in distances and light-curve parameters (panels (b–e)) of our compilation using two SALT models in four logarithmically spaced redshift bins from $0.015 < z < 1$. SALT3.K21 improves measured RMS Hubble scatter across our entire redshift range. The smallest improvements appear to be in the intermediate redshift range $z \sim 0.2$, with larger improvements at both lower and higher redshifts where improvements to the red and blue regions of the rest-frame model are most important.

lower in SALT3.K21 than SALT2.JLA. The color scatter, representing an intrinsic variance of the SN Ia population unexplained by the model, is not in general expected to improve as the size of a training sample grows. Other light-curve uncertainties visualized in the right panels of Figure 12 include contributions from both the aforementioned "in-sample" variations, which are beyond the ability of a two-parameter model to address, as well as the "out-of-sample" uncertainties in the determination of the model, which scale with the amount of training data. However, the SALT2.JLA color scatter was itself unconstrained by the JLA data past mean passband wavelengths ∼8000 Å, and we find that, while the color scatter is significant at these wavelengths, it is much smaller than implied by the SALT2.JLA model extrapolated into this region. Our additional data constrain this region up to ∼8500 Å, and redder data will be required to see how this effect carries into the NIR. We also note lower color scatter by ∼1% in the blue, which we attribute to improved relative calibration of the training sample.

Fitting the light curves using the SALT3.K21 model, we compare the uncertainties in distances and light-curve parameters to those found using SALT2.JLA. In Figure 14, we compare performance across the redshift range. We have chosen to show binned values of $\alpha\sigma_{x_1}$ and $\beta\sigma_c$, as these are the more physically relevant quantities; any simple rescaling of the x_1, c parameters would leave these quantities unaffected. Our model shows reduced Hubble scatter and distance uncertainties over the SALT2.JLA model at nearly every redshift, with the least improvement at moderate redshifts $z \sim 0.2$, where S/N is high and the SALT2.JLA model is already performing well. Light-curve parameters from SALT3.K21train have smaller uncertainties across the redshift range, with the exception of x_1 uncertainties. Our largest improvements are found at low redshift, where the improved red wavelength coverage of our model allows the use of additional light curve bands in fitting these SNe, providing stronger constraints on light-curve parameters. As shown in panel (d) of Figure 14, the color parameter c in particular is better constrained by the use of redder data, as it provides a longer lever-arm with which to infer the underlying parameter. That these redder bands may have other underlying processes not well-represented by a single color parameter is accounted for in our construction by the color scatter determined from the data, but even in z-band these uncertainties do not exceed ∼0.04 mag. We conclude that photometric observations at these wavelengths, previously discarded in SALT2 cosmology fits, continue to provide cosmologically useful information about the SALT3 light-curve parameters.

There is a slight improvement in RMS Hubble residuals, largely attributable to decreased color uncertainties in the low-redshift sample. Breaking this down further, Figure 15 shows the Hubble scatter binned by the number of additional filters used in the SALT3.K21 light-curve fit as compared to the SALT2.JLA light-curve fit. Where two additional filters are available, RMS ($\Delta\mu$) improves by ∼10%. We conclude improvement in the SALT3 model is most noticeable when it allows us to fit SNe with existing light curves in filters out of the SALT2.JLA wavelength range. At high redshift, SALT3's

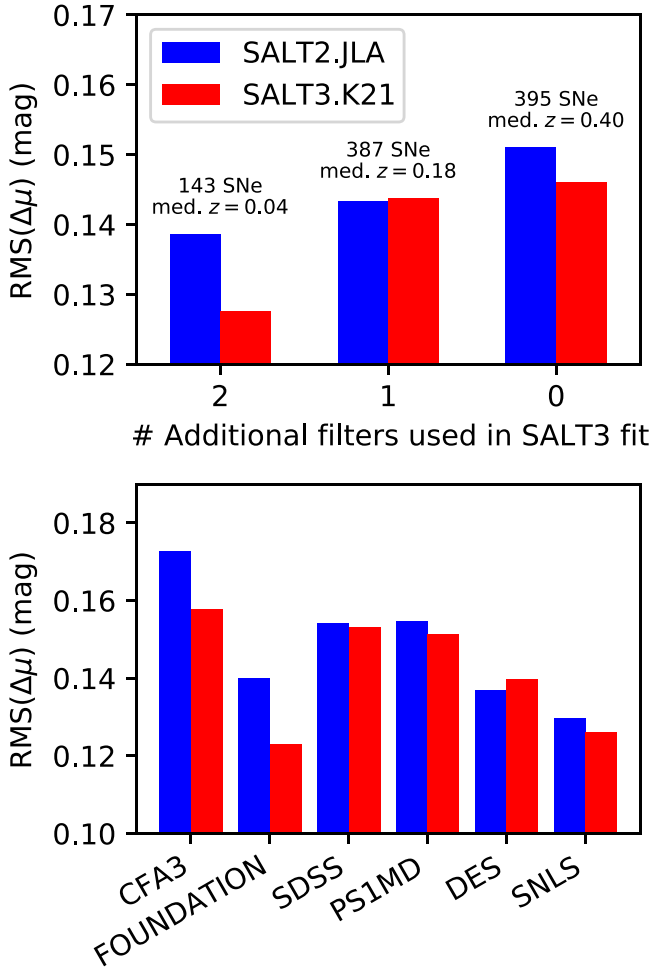

Figure 15. Upper panel: RMS ($\Delta\mu$) vs. the number of additional filters used in SALT3.K21 (red) light-curve fits as compared to fits made using SALT2.JLA (blue). The number of SNe and median redshift are printed above each bin. Lower panel: RMS ($\Delta\mu$) vs. sample for SALT3.K21 (red) and SALT2.JLA (blue), for all samples that have more than 30 SNe Ia in our compilation. As we move to surveys dedicated to searches at higher redshift, fewer SNe have data that could not be fit with SALT2, and the advantage of SALT3's wavelength range becomes less important.

improved constraints on the NUV model reduce Hubble scatter by ∼0.01 mag, although there is typically not sufficiently red data to take full advantage of the extended wavelength range at these redshifts. Similarly, when breaking down results by survey, as we show in Figure 15, the greatest improvement is in the low-z CFA3 and Foundation samples, where we can make use of I- and iz-band observations (respectively) previously unused in SALT2.JLA-based analyses at low redshift. We are able to reduce the RMS ($\Delta\mu$) of the Foundation sample from 0.144 ± 0.001 mag to 0.125 mag, an improvement of 15%.

5.2.2. Comparison of SALT3.K21 and SALT2.JLA Light-curve Parameters

The `SALTshaker` training procedure results in different distributions of light-curve parameters compared to SALT2.JLA. Some differences are due to changes in how the SALT3 model is defined, while others are from the demographics of the training sample. For example, a x_1/M_1 degeneracy is broken by setting a constraint on the standard deviation $\sigma_{x_1} = 1$, both in our procedure and in the original training procedure; therefore, including additional data with higher stretch increases the scale of the M_1 component.

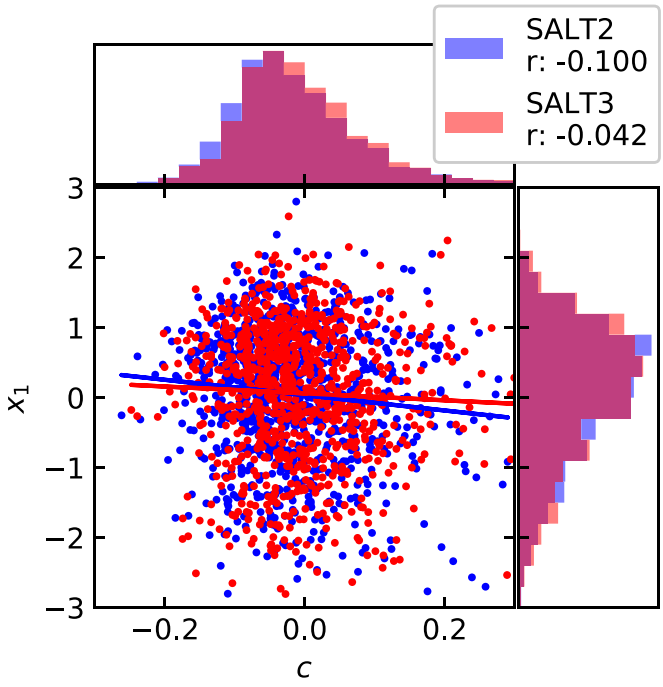

Figure 16. Scatterplot of the x_1 and c light-curve parameters measured with SALT2.JLA and SALT3.K21. We observe a linear transformation of the parameters, as expected from changes in the demographics of the training sample and our procedure for separating stretch and color. Best fit c-x_1 lines and correlation coefficients r are shown, to illustrate the rotation of the distribution as SALT3 forces the independence of x_1 and c.

In Figure 16, we compare the distributions of the x_1 and c parameters from both models. We note a slight rotation, shift, and scale change of the color/stretch distribution relative to SALT2.JLA. These linear transformations cancel in the Tripp standardization of distance, absorbed into the nuisance parameters α, β, $\mathcal{M}$. The offset in color in particular is from including bluer high-redshift SNe from PS1 and DES in the training sample. The rotation is due to the distinct procedures for separating stretch and color between the SALT2 training code and `SALTshaker`. SALT3.K21 x_1 and c values can be approximated from the SALT2.JLA values by the transformations

$$x_1^{(\mathrm{SALT3})} \approx 1.028 x_1^{(\mathrm{SALT2})} + 0.138 c^{(\mathrm{SALT2})} + 0.005 \quad (19)$$

$$c^{(\mathrm{SALT3})} \approx 0.002 x_1^{(\mathrm{SALT2})} + 0.985 c^{(\mathrm{SALT2})} + 0.013. \quad (20)$$

To test the similarity of the shapes of the distributions, which can potentially affect selection bias corrections for cosmology, we use the two-dimensional Kolmogorov–Smirnov—like test statistic of Press & Teukolsky (1988). As the means and scales of the distributions lack physical information, we modify the procedure by linearly transforming both sets of light-curve parameters to have the same mean, standard deviation, and x_1-c correlation before calculating the statistic. We find the test statistic for the SALT2.JLA and SALT3.K21 x_1 and c parameters, then bootstrap-resample the data and calculate the test statistic with the resampled data. We derive a p-value = 0.895, and conclude that we cannot distinguish the shapes of the underlying distributions.

6. Conclusions

We have presented SALTshaker, a new Python-based training code to train a phenomenologically motivated light-curve model using the SALT framework, in addition to a retrained SALT model we call SALT3. SALTshaker is publicly available at https://github.com/djones1040/SALTShaker with documentation at https://saltshaker.readthedocs.io/en/latest/. The SALTshaker documentation includes links to the training data and the SALT3.K21 model, and the SALT3.K21 model is compatible with and included in the latest versions of the popular SNANA and sncosmo codes.

The SALT3.K21 model itself features updated calibration with Supercal (Scolnic et al. 2015) and revised MW $E(B\text{-}V)$ estimates from Schlafly & Finkbeiner (2011). The training sample includes light curves and spectra from 1083 SNe—2.5 times larger than previous training samples—and extends to the rest-frame iz bands. Due to its larger wavelength range, we find that SALT3.K21 distances making use of iz data at low-z are approximately 15% more precise, equivalent to increasing the low-z sample size by 30%.

We found no significant change in distances inferred using a Tripp estimator between models, with a consistency between high- and low-z samples of 3 ± 14 mmag. As part of an upcoming cosmology analysis, we have employed SALT3.K21 within the PIPPIN (Hinton & Brout 2020) framework, generating simulations and performing bias corrections following the methodologies of Kessler et al. (2019) and Kessler & Scolnic (2017). Although this work is preliminary, we find that cosmological parameters measured using a larger light-curve data set of ∼2000 SNe Ia (to be released in D. Brout et al., in preparation) are consistent between SALT3.K21 and SALT2.JLA.

Several light-curve models have been developed for cosmological supernovae, including MLCS (Riess et al. 1996), MLCS2k2 (Jha et al. 2007), SiFTO (Conley et al. 2008), SNooPy (Burns et al. 2011, 2014), SNEMO (Saunders et al. 2018), SUGAR (Léget et al. 2020), and BayesSN (Mandel et al. 2011, 2020). In the context of other modern light-curve models, SALT3 offers an approach to model design and training process that prioritizes the use of heterogeneous spectral and photometric data to provide extensive phase and wavelength coverage and native k-corrections through cosmology-independent training. The model framework is minimally changed from SALT2, so SNANA simulations, bias corrections, and other analysis products are expected to require little revision.

Over the coming years, we expect SALTshaker will continue to be developed and improved as additional SN data becomes available and additional SN standardization parameters (e.g., host mass) are discovered and explored. Further development work could focus on the error model, which is currently based on central filter wavelengths rather than integrated quantities (see Section 2.2). This is a potential source of systematic uncertainty because observer-frame filter functions are contracted in the rest frame.

Additionally, SALTshaker enables a more rigorous evaluation of systematic uncertainties such as those arising from limited training data, photometric calibration uncertainties, or treatment of SN spectra. These can be evaluated in a straightforward and rigorous way by retraining the SALT3 model surfaces on simulated data. Although we have demonstrated that SALTshaker can faithfully recover a simulated model at the ∼1% level, future work will also be able to fully validate the model training process using an entire analysis chain that includes training, bias corrections, and cosmology fitting (M. Dai et al. in preparation). Similarly, while the SALT3 model surfaces presented in this work have been trained on data recalibrated to the level of 1% via the Supercal procedure (Scolnic et al. 2015, 2018), we have left quantifying the reduced calibration uncertainties as a topic for further work.

Previous work has investigated the impact of potential systematic uncertainties for SNIa cosmology such as perturbations to line velocity, multiple sources of reddening with distinct effects, and correlations of host galaxy properties with the SNIa SED (Pierel et al. 2021). However, future work could implement the SALTshaker method on a training sample with such underlying effects, simulated with the BYOSED code, the impact of which has yet to be explored. By propagating any biases introduced by the training code into cosmology, we may quantify the potential impact of currently unmodeled supernova phenomenology on cosmology for the first time.

Looking to future missions, we find that, for the forecast SN Ia survey of the Roman Space Telescope, assuming the ALLz strategy (Hounsell et al. 2018), the extended wavelength range in SALT3.K21 makes use of ∼20% more observations compared to SALT2.JLA; this increase is about a factor of 2 for redshifts $z < 0.5$, and falls to about a 10% increase at $z = 1.5$. Samples of SN Ia light curves will increase by orders of magnitude with the Vera Rubin Telescope's LSST (Ivezić et al. 2019) and the Roman Space Telescope (Hounsell et al. 2018). For error budgets to continue improvement, light-curve models should not be tied to outdated calibration standards, and it is essential that the model training process be regarded as a key component of an integrated cosmology analysis, as has been done in Betoule et al. (2014) and Mosher et al. (2014). SALTshaker provides a framework for consistent, robust model improvements in future years and direct inclusion into existing cosmology analysis pipelines.

We thank Julien Guy and the developers of the SALT and SALT2 models for their contributions to the field; Julien Guy, Chris Lidman, and Georgie Taylor for valuable assistance and discussion; and Julien Guy, Marc Betoule, and the SNLS team for providing the original snpca code and JLA training sample. We are grateful for all astronomers who acquired the photometry and spectroscopy used to train the SALT3.K21 model.

D.O.J. is supported by a Gordon and Betty Moore Foundation postdoctoral fellowship at the University of California, Santa Cruz, and by NASA through the NASA Hubble Fellowship grant HF2-51462.001 awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under contract NAS5–26555. Support for this work was provided by the STScI Director's Discretionary Fund. This research at Rutgers University (S.W.J., M.D.) was supported by NASA contract NNG16PJ34C and DOE award DE-SC0011636. M.D. is also supported by the Horizon Fellowship at the Johns Hopkins University. This work at the University of South Carolina (S. Rodney, J. D. R. Pierel) was supported by grant HST-AR-15808 from the Space Telescope Science Institute. This work was supported by NASA contract No. NNG17PX03C issued through the Nancy G. Roman Science Investigation Teams Program.

This work was completed in part with resources provided by the University of Chicago's Research Computing Center.

This research is based on observations made with the NASA/ESA Hubble Space Telescope obtained from the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 526555. These observations are associated with programs GO-4016, GO-4252, GO-8611, GO-9114, and GO-10182.

This research is based on observations made with the International Ultraviolet Explorer, obtained from the MAST data archive at the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 526555. These observations are associated with programs METOO, SNMRK, SNNRK, STKRK, and VILSP.

This paper is based in part on observations made with the Southern African Large Telescope (SALT) using the Robert Stobie Spectrograph (RSS), through allocations made to Rutgers University via programs 2015-1-MLT-002, 2016-1-MLT-007, and 2017-1-MLT-002 (PI: S.W. Jha).

Based in part on observations obtained at the Southern Astrophysical Research (SOAR) telescope (NOIRLab Prop. IDs 2015A-0253, 2015B-0313, 2017A-0306, 2017B-0169, 2018A-0277; PI: R. Foley), which is a joint project of the Ministério da Ciência, Tecnologia e Inovações do Brasil (MCTI/LNA), the US National Science Foundation's NOIRLab, the University of North Carolina at Chapel Hill (UNC), and Michigan State University (MSU).

Based in part on observations at Kitt Peak National Observatory at NSF's NOIRLab (NOIRLab Prop. IDs 2015A-0253, 2015B-0313, 2017A-0306, 2017B-0169, 2018A-0277; PI: R. Foley), which is managed by the Association of Universities for Research in Astronomy (AURA) under a cooperative agreement with the National Science Foundation. The authors are honored to be permitted to conduct astronomical research on Iolkam Du'ag (Kitt Peak), a mountain with particular significance to the Tohono O'odham.

A major upgrade of the Kast spectrograph on the Shane 3 m telescope at Lick Observatory was made possible through generous gifts from the Heising-Simons Foundation as well as William and Marina Kast. Research at Lick Observatory is partially supported by a generous gift from Google.

Some of the data presented herein were obtained at the W. M. Keck Observatory, which is operated as a scientific partnership among the California Institute of Technology, the University of California and the National Aeronautics and Space Administration. The Observatory was made possible by the generous financial support of the W. M. Keck Foundation. The authors wish to recognize and acknowledge the very significant cultural role and reverence that the summit of Maunakea has always had within the indigenous Hawaiian community. We are most fortunate to have the opportunity to conduct observations from this mountain.

Software: AstroPy (Astropy Collaboration et al. 2013, 2018), Astroquery (Ginsburg et al. 2019), extinction(Barbary 2016), iMinuit (James & Roos 1975; Dembinski et al. 2020), Matplotlib (Hunter 2007), NumPy (Harris et al. 2020), SciPy (Virtanen et al. 2020), sncosmo (Barbary et al. 2015, 2016), tqdm (da Costa-Luis et al. 2021).

ORCID iDs

W. D. Kenworthy https://orcid.org/0000-0002-5153-5983
D. O. Jones https://orcid.org/0000-0002-6230-0151
M. Dai https://orcid.org/0000-0002-5995-9692
R. Kessler https://orcid.org/0000-0003-3221-0419
D. Brout https://orcid.org/0000-0001-5201-8374
M. R. Siebert https://orcid.org/0000-0003-2445-3891
K. G. Dettman https://orcid.org/0000-0001-7519-133X
G. Dimitriadis https://orcid.org/0000-0001-9494-179X
R. J. Foley https://orcid.org/0000-0002-2445-5275
S. W. Jha https://orcid.org/0000-0001-8738-6011
Y.-C. Pan https://orcid.org/0000-0001-8415-6720
A. Riess https://orcid.org/0000-0002-6124-1196
S. Rodney https://orcid.org/0000-0003-1947-687X
C. Rojas-Bravo https://orcid.org/0000-0002-7559-315X

References

Abbott, T. M. C., Allam, S., Andersen, P., et al. 2019, ApJL, 872, L30
Amanullah, R., Johansson, J., Goobar, A., et al. 2015, MNRAS, 453, 3300
Anupama, G. C., Sahu, D. K., & Jose, J. 2005, A&A, 429, 667
Astier, P., Guy, J., Regnault, N., et al. 2006, A&A, 447, 31
Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, AJ, 156, 123
Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33
Avelino, A., Friedman, A. S., Mandel, K. S., et al. 2019, ApJ, 887, 106
Balland, C., Cellier-Holzem, F., Lidman, C., et al. 2018, A&A, 614, A134
Barbary, K. 2016, Extinction v0.3.0, Zenodo, doi:10.5281/zenodo.804967
Barbary, K., rbiswas4, Rodney, S., et al. 2015, Sncosmo: v1.0.0, Zenodo, doi:10.5281/zenodo.592747
Barbary, K., Barclay, T., Biswas, R., et al. 2016, SNCosmo: Python Library for Supernova Cosmology, Astrophysics Source Code Library, http://ascl.net/1611.017
Benetti, S., Meikle, P., Stehle, M., et al. 2004, MNRAS, 348, 261
Betoule, M., Kessler, R., Guy, J., et al. 2014, A&A, 568, A22
Blondin, S., Matheson, T., Kirshner, R. P., et al. 2012, AJ, 143, 126
Branch, D., Fisher, A., & Nugent, P. 1993, AJ, 106, 2383
Brout, D., Hinton, S., & Scolnic, D. 2021, ApJL, 912, L26
Brout, D., Sako, M., Scolnic, D., et al. 2019b, ApJ, 874, 106
Brout, D., Scolnic, D., Kessler, R., et al. 2019a, ApJ, 874, 150
Burke, D. L., Rykoff, E. S., Allam, S., et al. 2018, AJ, 155, 41
Burns, C. R., Stritzinger, M., Phillips, M. M., et al. 2011, AJ, 141, 19
Burns, C. R., Stritzinger, M., Phillips, M. M., et al. 2014, ApJ, 789, 32
Cardelli, J. A., Clayton, G. C., & Mathis, J. S. 1989, ApJ, 345, 245
Chotard, N., Gangler, E., Aldering, G., et al. 2011, A&A, 529, L4
Conley, A., Guy, J., Sullivan, M., et al. 2011, ApJS, 192, 1
Conley, A., Sullivan, M., Hsiao, E. Y., et al. 2008, ApJ, 681, 482
Contreras, C., Hamuy, M., Phillips, M. M., et al. 2010, AJ, 139, 519
da Costa-Luis, C., Larroque, S. K., Altendorf, K., et al. 2021, tqdm: A fast, Extensible Progress Bar for Python and CLI, v4.59.0, Zenodo, doi:10.5281/zenodo.4586769
Dembinski, H., Ongmongkolkul, P., Deil, C., et al. 2020, scikit-hep/iminuit: v1.4.9, v1.4.9, Zenodo, doi:10.5281/zenodo.3951328
Dettman, K. G., Jha, S. W., Dai, M., et al. 2021, arXiv:2102.06524
Dhawan, S., Jha, S. W., & Leibundgut, B. 2018, A&A, 609, A72
Di Valentino, E., Mena, O., Pan, S., et al. 2021, CQGra, 38, 153001
Filippenko, A. V., Richmond, M. W., Matheson, T., et al. 1992, ApJL, 384, L15
Fitzpatrick, E. L. 1999, PASP, 111, 63
Folatelli, G., Morrell, N., Phillips, M. M., et al. 2013, ApJ, 773, 53
Foley, R. J., Filippenko, A. V., & Jha, S. W. 2008, ApJ, 686, 117
Foley, R. J., & Kasen, D. 2011, ApJ, 729, 55
Foley, R. J., Narayan, G., Challis, P. J., et al. 2010, ApJ, 708, 1748
Foley, R. J., Scolnic, D., Rest, A., et al. 2018, MNRAS, 475, 193
Freedman, W. L. 2021, ApJ, 919, 16
Gaia Collaboration, Prusti, T., de Bruijne, J. H. J., et al. 2016, A&A, 595, A1
Garavini, G., Nobili, S., Taubenberger, S., et al. 2007, A&A, 471, 527
Ginsburg, A., Sipőcz, B. M., Brasseur, C. E., et al. 2019, AJ, 157, 98
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
Hamuy, M., Phillips, M. M., Suntzeff, N. B., et al. 1996, AJ, 112, 2408

Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, Natur, 585, 357
Hicken, M., Challis, P., Jha, S., et al. 2009b, ApJ, 700, 331
Hicken, M., Challis, P., Kirshner, R. P., et al. 2012, ApJS, 200, 12
Hicken, M., Wood-Vasey, W. M., Blondin, S., et al. 2009a, ApJ, 700, 1097
Hinton, S., & Brout, D. 2020, JOSS, 5, 2122
Holtzman, J. A., Marriner, J., Kessler, R., et al. 2008, AJ, 136, 2306
Hounsell, R., Scolnic, D., Foley, R. J., et al. 2018, ApJ, 867, 23
Huber, M., Chambers, K. C., Flewelling, H., et al. 2015, ATel, 7153, 1
Hunter, J. D. 2007, CSE, 9, 90
Ivezić, Ž., Kahn, S. M., Tyson, J. A., et al. 2019, ApJ, 873, 111
James, F., & Roos, M. 1975, CoPhC, 10, 343
Jha, S., Kirshner, R. P., Challis, P., et al. 2006, AJ, 131, 527
Jha, S., Riess, A. G., & Kirshner, R. P. 2007, ApJ, 659, 122
Jones, D. O., Riess, A. G., Scolnic, D. M., et al. 2018, ApJ, 867, 108
Jones, D. O., Scolnic, D. M., Foley, R. J., et al. 2019, ApJ, 881, 19
Jones, D. O., Scolnic, D. M., Riess, A. G., et al. 2017, ApJ, 843, 6
Kelly, P. L., Hicken, M., Burke, D. L., Mand el, K. S., & Kirshner, R. P. 2010, ApJ, 715, 743
Kessler, R., Becker, A. C., Cinabro, D., et al. 2009b, ApJS, 185, 32
Kessler, R., Bernstein, J. P., Cinabro, D., et al. 2009a, PASP, 121, 1028
Kessler, R., Bernstein, J. P., Cinabro, D., et al. 2009c, PASP, 121, 1028
Kessler, R., Brout, D., D'Andrea, C. B., et al. 2019, MNRAS, 485, 1171
Kessler, R., Marriner, J., Childress, M., et al. 2015, AJ, 150, 172
Kessler, R., & Scolnic, D. 2017, ApJ, 836, 56
Kotak, R., Meikle, W. P. S., Pignata, G., et al. 2005, A&A, 436, 1021
Krisciunas, K., Contreras, C., Burns, C. R., et al. 2017, AJ, 154, 211
Krisciunas, K., Garnavich, P. M., Stanishev, V., et al. 2007, AJ, 133, 58
Krisciunas, K., Suntzeff, N. B., Phillips, M. M., et al. 2004, AJ, 128, 3034
Léget, P. E., Gangler, E., Mondon, F., et al. 2020, A&A, 636, A46
Leonard, D. C., Li, W., Filippenko, A. V., Foley, R. J., & Chornock, R. 2005, ApJ, 632, 450
Li, W., Filippenko, A. V., Treffers, R. R., et al. 2001, ApJ, 546, 734
Mandel, K. S., Narayan, G., & Kirshner, R. P. 2011, ApJ, 731, 120
Mandel, K. S., Thorp, S., Narayan, G., Friedman, A. S., & Avelino, A. 2020, arXiv:2008.07538
Marriner, J., Bernstein, J. P., Kessler, R., et al. 2011, ApJ, 740, 72
Marwil, E. 1979, SJNA, 16, 588
Mosher, J., Guy, J., Kessler, R., et al. 2014, ApJ, 793, 16
NAT 1997, NATO Advanced Study Institute (ASI) Series C, Vol. 486, Thermonuclear Supernovae (Berlin: Springer)
Östman, L., Nordin, J., Goobar, A., et al. 2011, A&A, 526, A28
Patat, F., Benetti, S., Cappellaro, E., et al. 1996, MNRAS, 278, 111
Perlmutter, S., Aldering, G., Goldhaber, G., et al. 1999, ApJ, 517, 565
Phillips, M. M., & Phillips, M. M. 1993, ApJL, 413, L105
Pierel, J. D. R., Jones, D. O., Dai, M., et al. 2021, ApJ, 911, 96
Pierel, J. D. R., Rodney, S., Avelino, A., et al. 2018, PASP, 130, 114504
Pignata, G., Benetti, S., Mazzali, P. A., et al. 2008, MNRAS, 388, 971
Press, W. H., & Teukolsky, S. A. 1988, ComPh, 2, 74
Rest, A., Scolnic, D., Foley, R. J., et al. 2014, ApJ, 795, 44
Riess, A. G., Casertano, S., Yuan, W., et al. 2021, ApJL, 908, L6
Riess, A. G., Filippenko, A. V., Challis, P., et al. 1998, AJ, 116, 1009
Riess, A. G., Kirshner, R. P., Schmidt, B. P., et al. 1999, AJ, 117, 707
Riess, A. G., Macri, L. M., Hoffmann, S. L., et al. 2016, ApJ, 826, 56
Riess, A. G., Press, W. H., & Kirshner, R. P. 1996, ApJ, 473, 88
Riess, A. G., Rodney, S. A., Scolnic, D. M., et al. 2018, ApJ, 853, 126
Rigault, M., Brinnel, V., Aldering, G., et al. 2020, A&A, 644, A176
Rigault, M., Copin, Y., Aldering, G., et al. 2013, A&A, 560, A66
Rose, B. M., Dixon, S., Rubin, D., et al. 2020, ApJ, 890, 60
Rubin, D. 2020, ApJ, 897, 40
Sako, M., Bassett, B., Becker, A. C., et al. 2018, PASP, 130, 064002
Salvo, M. E., Cappellaro, E., Mazzali, P. A., et al. 2001, MNRAS, 321, 254
Saunders, C., Aldering, G., Antilogus, P., et al. 2018, ApJ, 869, 167
Schlafly, E. F., & Finkbeiner, D. P. 2011, ApJ, 737, 103
Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, ApJ, 500, 525
Schubert, L. K. 1970, MaCom, 24, 27
Scolnic, D., Casertano, S., Riess, A., et al. 2015, ApJ, 815, 117
Scolnic, D., & Kessler, R. 2016, ApJL, 822, L35
Scolnic, D. M., Jones, D. O., Rest, A., et al. 2018, ApJ, 859, 101
Shappee, B. J., Prieto, J. L., Grupe, D., et al. 2014, ApJ, 788, 48
Siebert, M. R., Foley, R. J., Jones, D. O., et al. 2019, MNRAS, 486, 5785
Siebert, M. R., Foley, R. J., Jones, D. O., & Davis, K. W. 2020, MNRAS, 493, 5713
Silverman, J. M., Foley, R. J., Filippenko, A. V., et al. 2012a, MNRAS, 425, 1789
Smith, M., D'Andrea, C. B., Sullivan, M., et al. 2020b, AJ, 160, 267
Smith, M., Sullivan, M., Wiseman, P., et al. 2020, MNRAS, 494, 4426
Stahl, B. E., Zheng, W., de Jaeger, T., et al. 2020, MNRAS, 492, 4325
Stanishev, V., Goobar, A., Benetti, S., et al. 2007, A&A, 469, 645
Stritzinger, M. D., Phillips, M. M., Boldt, L. N., et al. 2011, AJ, 142, 156
Sullivan, M., Conley, A., Howell, D. A., et al. 2010, MNRAS, 406, 782
Taylor, G., Lidman, C., Tucker, B. E., et al. 2021, MNRAS, 504, 4111
The LSST Dark Energy Science Collaboration, Mandelbaum, R., Eifler, T, et al. 2018, arXiv:1809.01669
Thomas, R. C., Aldering, G., Antilogus, P., et al. 2007, ApJL, 654, L53
Tonry, J. L., Denneau, L., Heinze, A. N., et al. 2018, PASP, 130, 064505
Tripp, R. 1998, A&A, 331, 815
Valentini, G., Di Carlo, E., Massi, F., et al. 2003, ApJ, 595, 779
Verde, L., Treu, T., & Riess, A. G. 2019, NatAs, 3, 891
Villar, V. A., Hosseinzadeh, G., Berger, E., et al. 2020, ApJ, 905, 94
Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, NatMe, 17, 261
Walker, E. S., Hook, I. M., Sullivan, M., et al. 2011, MNRAS, 410, 1262
Wang, X., Filippenko, A. V., Ganeshalingam, M., et al. 2009a, ApJL, 699, L139
Wang, X., Li, W., Filippenko, A. V., et al. 2009b, ApJ, 697, 380
Wells, L. A., Phillips, M. M., Suntzeff, B., et al. 1994, AJ, 108, 2233
Wood-Vasey, W. M., Friedman, A. S., Bloom, J. S., et al. 2008, ApJ, 689, 377