The Astrophysical Journal, 975:86 (31pp), 2024 November 1 https://doi.org/10.3847/1538-4357/ad5e6c

The Dark Energy Survey Supernova Program: Cosmological Analysis and Systematic Uncertainties

M. Vincenzi[1,2], D. Brout[3,4,5], P. Armstrong[6], B. Popovic[7], G. Taylor[6], M. Acevedo[8], R. Camilleri[9], R. Chen[1], T. M. Davis[9], J. Lee[8], C. Lidman[6,10], S. R. Hinton[9], L. Kelsey[11], R. Kessler[12,13], A. Möller[14], H. Qu[8], M. Sako[8], B. Sanchez[15], D. Scolnic[1], M. Smith[16], M. Sullivan[16], P. Wiseman[16], J. Asorey[17], B. A. Bassett[18,19], D. Carollo[20], A. Carr[9], R. J. Foley[21], C. Frohmaier[16], L. Galbany[22,23], K. Glazebrook[14], O. Graur[11], E. Kovacs[24], K. Kuehn[25,26], U. Malik[6], R. C. Nichol[27], B. Rose[1,28], B. E. Tucker[6], M. Toy[16], D. L. Tucker[29], F. Yuan[6], T. M. C. Abbott[30], M. Aguena[31], O. Alves[32], S. S. Allam[29], F. Andrade-Oliveira[32], J. Annis[29], D. Bacon[11], K. Bechtol[33], G. M. Bernstein[8], D. Brooks[34], D. L. Burke[35,36], A. Carnero Rosell[31,37,38], J. Carretero[39], F. J. Castander[22,23], C. Conselice[40,41], L. N. da Costa[31], M. E. S. Pereira[42], S. Desai[43], H. T. Diehl[29], P. Doel[34], I. Ferrero[44], B. Flaugher[29], D. Friedel[45], J. Frieman[13,29], J. García-Bellido[46], M. Gatti[8], G. Giannini[13,39], D. Gruen[47], R. A. Gruendl[45,48], D. L. Hollowood[21], K. Honscheid[49,50], D. Huterer[32], D. J. James[5], N. Kuropatkin[29], O. Lahav[34], S. Lee[51], H. Lin[29], J. L. Marshall[52], J. Mena-Fernández[53], F. Menanteau[45,48], R. Miquel[39,54], A. Palmese[55], A. Pieres[31,56], A. A. Plazas Malagón[35,36], A. Porredon[57], A. K. Romer[58], A. Roodman[35,36], E. Sanchez[59], D. Sanchez Cid[59], M. Schubnell[32], I. Sevilla-Noarbe[59], E. Suchyta[60], M. E. C. Swanson[45], G. Tarle[32], C. To[49], A. R. Walker[30], N. Weaverdyck[32,61], M. Yamamoto[1] and
(DES Collaboration)

[1] Department of Physics, Duke University Durham, NC 27708, USA; maria.vincenzi@physics.ox.ac.uk
[2] Department of Physics, University of Oxford, Denys Wilkinson Building, Keble Road, Oxford OX1 3RH, UK
[3] Department of Astronomy, Boston University, 725 Commonwealth Avenue, Boston, MA 02215, USA
[4] Department of Physics, Boston University, 590 Commonwealth Avenue, Boston, MA 02215, USA
[5] Center for Astrophysics—Harvard & Smithsonian, 60 Garden Street, Cambridge, MA 02138, USA
[6] The Research School of Astronomy and Astrophysics, Australian National University, ACT 2601, Australia
[7] Univ Lyon, Univ Claude Bernard Lyon 1, CNRS, IP2I Lyon/IN2P3, IMR 5822, F-69622, Villeurbanne, France
[8] Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, PA 19104, USA
[9] School of Mathematics and Physics, University of Queensland, Brisbane, QLD 4072, Australia
[10] Institute of Cosmology and Gravitation, University of Portsmouth, Portsmouth, PO1 3FX, UK
[11] Centre for Gravitational Astrophysics, College of Science, The Australian National University, ACT 2601, Australia
[12] Department of Astronomy and Astrophysics, University of Chicago, Chicago, IL 60637, USA
[13] Kavli Institute for Cosmological Physics, University of Chicago, Chicago, IL 60637, USA
[14] Centre for Astrophysics & Supercomputing, Swinburne University of Technology, Victoria 3122, Australia
[15] Aix Marseille Univ, CNRS/IN2P3, CPPM, Marseille, France
[16] School of Physics and Astronomy, University of Southampton, Southampton, SO17 1BJ, UK
[17] Departamento de Fíca Teóa and Instituto de Fíca de Partílas y del Cosmos (IPARCOS-UCM), Universidad Complutense de Madrid, 28040 Madrid, Spain
[18] African Institute for Mathematical Sciences, 6 Melrose Road, Muizenberg, 7945, South Africa
[19] South African Astronomical Observatory, P.O. Box 9, Observatory 7935, South Africa
[20] INAF-Osservatorio Astronomico di Trieste, via G. B. Tiepolo 11, I-34143 Trieste, Italy
[21] Santa Cruz Institute for Particle Physics, Santa Cruz, CA 95064, USA
[22] Institute of Space Sciences (ICE, CSIC), Campus UAB, Carrer de Can Magrans, s/n, 08193 Barcelona, Spain
[23] Institut d'Estudis Espacials de Catalunya (IEEC), 08034 Barcelona, Spain
[24] Argonne National Laboratory, 9700 South Cass Avenue, Lemont, IL 60439, USA
[25] Australian Astronomical Optics, Macquarie University, North Ryde, NSW 2113, Australia
[26] Lowell Observatory, 1400 Mars Hill Road, Flagstaff, AZ 86001, USA
[27] School of Mathematics and Physics, University of Surrey, Guildford, Surrey, GU2 7XH, UK
[28] Department of Physics, Baylor University, One Bear Place #97316, Waco, TX 76798-7316, USA
[29] Fermi National Accelerator Laboratory, P.O. Box 500, Batavia, IL 60510, USA
[30] Cerro Tololo Inter-American Observatory, NSF's National Optical-Infrared Astronomy Research Laboratory, Casilla 603, La Serena, Chile
[31] Laboratório Interinstitucional de e-Astronomia—LIneA, Rua Gal. José Cristino 77, Rio de Janeiro, RJ—20921-400, Brazil
[32] Department of Physics, University of Michigan, Ann Arbor, MI 48109, USA
[33] Physics Department, 2320 Chamberlin Hall, University of Wisconsin-Madison, 1150 University Avenue Madison, WI 53706-1390, USA
[34] Department of Physics & Astronomy, University College London, Gower Street, London, WC1E 6BT, UK
[35] Kavli Institute for Particle Astrophysics & Cosmology, P. O. Box 2450, Stanford University, Stanford, CA 94305, USA
[36] SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA
[37] Instituto de Astrofisica de Canarias, E-38205 La Laguna, Tenerife, Spain
[38] Universidad de La Laguna, Dpto. Astrofíca, E-38206 La Laguna, Tenerife, Spain
[39] Institut de Física d'Altes Energies (IFAE), The Barcelona Institute of Science and Technology, Campus UAB, 08193 Bellaterra (Barcelona), Spain
[40] Jodrell Bank Center for Astrophysics, School of Physics and Astronomy, University of Manchester, Oxford Road, Manchester, M13 9PL, UK
[41] University of Nottingham, School of Physics and Astronomy, Nottingham NG7 2RD, UK
[42] Hamburger Sternwarte, Universität Hamburg, Gojenbergsweg 112, 21029 Hamburg, Germany
[43] Department of Physics, IIT Hyderabad, Kandi, Telangana 502285, India
[44] Institute of Theoretical Astrophysics, University of Oslo. P.O. Box 1029 Blindern, NO-0315 Oslo, Norway
[45] Center for Astrophysical Surveys, National Center for Supercomputing Applications, 1205 West Clark Street, Urbana, IL 61801, USA
[46] Instituto de Fisica Teorica UAM/CSIC, Universidad Autonoma de Madrid, 28049 Madrid, Spain
[47] University Observatory, Faculty of Physics, Ludwig-Maximilians-Universität, Scheinerstr. 1, 81679 Munich, Germany
[48] Department of Astronomy, University of Illinois at Urbana-Champaign, 1002 W. Green Street, Urbana, IL 61801, USA
[49] Center for Cosmology and Astro-Particle Physics, The Ohio State University, Columbus, OH 43210, USA
[50] Department of Physics, The Ohio State University, Columbus, OH 43210, USA

[51] Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA 91109, USA
[52] George P. and Cynthia Woods Mitchell Institute for Fundamental Physics and Astronomy, and Department of Physics and Astronomy, Texas A&M University, College Station, TX 77843, USA
[53] LPSC Grenoble—53, Avenue des Martyrs, 38026 Grenoble, France
[54] Institució Catalana de Recerca i Estudis Avançats, E-08010 Barcelona, Spain
[55] Department of Physics, Carnegie Mellon University, Pittsburgh, PA 15312, USA
[56] Observatório Nacional, Rua Gal. José Cristino 77, Rio de Janeiro, RJ—20921-400, Brazil
[57] Ruhr University Bochum, Faculty of Physics and Astronomy, Astronomical Institute, German Centre for Cosmological Lensing, 44780 Bochum, Germany
[58] Department of Physics and Astronomy, Pevensey Building, University of Sussex, Brighton, BN1 9QH, UK
[59] Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT), Madrid, Spain
[60] Computer Science and Mathematics Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA
[61] Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA

Received 2023 November 17; revised 2024 June 18; accepted 2024 June 19; published 2024 October 25

Abstract

We present the full Hubble diagram of photometrically classified Type Ia supernovae (SNe Ia) from the Dark Energy Survey supernova program (DES-SN). DES-SN discovered more than 20,000 SN candidates and obtained spectroscopic redshifts of 7000 host galaxies. Based on the light-curve quality, we select 1635 photometrically identified SNe Ia with spectroscopic redshift $0.10 < z < 1.13$, which is the largest sample of supernovae from any single survey and increases the number of known $z > 0.5$ supernovae by a factor of 5. In a companion paper, we present cosmological results of the DES-SN sample combined with 194 spectroscopically classified SNe Ia at low redshift as an anchor for cosmological fits. Here we present extensive modeling of this combined sample and validate the entire analysis pipeline used to derive distances. We show that the statistical and systematic uncertainties on cosmological parameters are $\sigma_{\Omega_M,\mathrm{stat+sys}}^{\Lambda\mathrm{CDM}} = 0.017$ in a flat ΛCDM model, and $(\sigma_{\Omega_M}, \sigma_w)_{\mathrm{stat+sys}}^{w\mathrm{CDM}} = (0.082, 0.152)$ in a flat wCDM model. Combining the DES SN data with the highly complementary cosmic microwave background measurements by Planck Collaboration reduces by a factor of 4 uncertainties on cosmological parameters. In all cases, statistical uncertainties dominate over systematics. We show that uncertainties due to photometric classification make up less than 10% of the total systematic uncertainty budget. This result sets the stage for the next generation of SN cosmology surveys such as the Vera C. Rubin Observatory's Legacy Survey of Space and Time.

Unified Astronomy Thesaurus concepts: Type Ia supernovae (1728); Cosmological constant experiments (335); Dark energy (351); Cosmological parameters (339)

1. Introduction

The modern understanding of the physical evolution of our Universe comes from a number of cosmological probes that can constrain the Universe's expansion history and growth of structure. The Dark Energy Survey (DES) employs multiple probes (Type Ia supernovae, weak lensing, large-scale structure, galaxy clusters) to accurately measure a generation-defining picture of the components of the Universe. In this paper, we present the analysis and distance constraints of the full five years of Type Ia (Ia) supernovae (SNe) discovered and measured by the DES Supernova program (DES-SN).

SNe Ia are used to make some of the most precise constraints on the nature of dark energy and the expansion history of the Universe from a large span in cosmic history $0 \lesssim z \lesssim 2$. The latest compilations (e.g., Pantheon+; Brout et al. 2022a; Scolnic et al. 2022) include ∼1500 distinct SNe; however, they have relied on real-time spectroscopic confirmations of the SNe themselves to be verified as type Ia. There is already an equally large number of SNe discovered for which spectroscopic confirmation was not possible, and recent cosmological analyses have begun to show that the contamination from other types of SNe in the analyses is not debilitating, and not even the largest systematic uncertainty (Sako et al. 2011; Campbell et al. 2013; Jones et al. 2018, 2019, and Popovic et al. 2024). This is, in part, due to the advancement of photometric classification algorithms (e.g., Möller & de Boissière 2020; Qu et al. 2021) that incorporate an improved set of non-Ia spectrophotometric templates and modeling (Vincenzi et al. 2019). Despite the success of recent photometric analyses (Jones et al. 2018, 2019), these samples have not received the same level of use in the broader cosmological community of combined-probe analyses. The onus has remained on the SN community to demonstrate that the accuracy of photometric SN analyses is not a limitation. This work presents an opportunity to set the stage for future analyses of orders of magnitude larger SN samples, such as the Rubin Observatory Legacy Survey of Space and Time (LSST; Ivezić et al. 2019) and the Nancy Grace Roman Space Telescope (Roman; Hounsell et al. 2018; Rose et al. 2021), where photometric classification represents the only viable path to fully exploit the statistical power of these surveys.

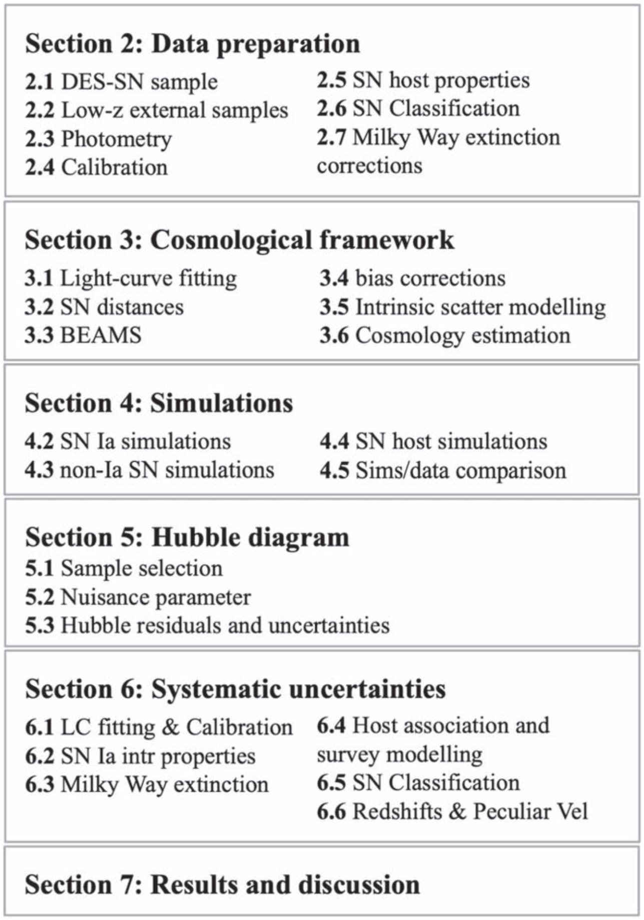
The analysis presented here is of the full 5 yr photometrically classified set of SNe from DES and additional external samples of spectroscopically confirmed low-z SNe (the DES-SN5YR analysis). This work was preceded by the analysis of the first 3 yr of spectroscopically classified DES-SNe Ia (DES-SN3YR: Abbott et al. 2019, also including external low-z SN samples). The DES-SN3YR sample included 207 SNe Ia from DES and was critical in the development and motivation of analyses leading up to the work presented here. This includes photometry and calibration (Burke et al. 2018; Brout et al. 2019a; Lasker et al. 2019), survey and SN Ia population modeling (Kessler et al. 2019a; Popovic et al. 2021), understanding and modeling of the “mass/dust step” (Lampeitl et al. 2010; Sullivan et al. 2010; Smith et al. 2018; Scolnic et al. 2020; Brout & Scolnic 2021; Popovic et al. 2021a; Chen et al. 2022; Dixon et al. 2022; Duarte et al. 2023; Wiseman et al. 2022; Meldorf et al. 2023), estimates and treatment of systematic uncertainties (Brout et al. 2019b, 2020), and the automation of the analysis pipeline (Hinton & Brout 2020).

In this work, we also introduce a number of new supporting analyses that are part of the full DES-SN5YR suite of papers. Besides photometric classification (Vincenzi et al. 2021, 2021b; Möller et al. 2022), the biggest differences and improvements in the methodology of DES-SN5YR compared to the DES-SN3YR analysis are due to:

1. Upgrading to the SALT3 light-curve model from the SALT2 model (Kenworthy et al. 2021; Taylor et al. 2023).
2. Improved modeling of SN Ia intrinsic scatter using host dust-based models (Popovic et al. 2021a), focusing on modeling correlations between SN Ia properties and both host mass and host color (rest-frame $u - r$, following Kelsey et al. 2023) and modeling the SN Ia progenitor age distribution (Wiseman et al. 2022), thus significantly improving previous analyses that neglect modeling of host properties (Scolnic & Kessler 2016).
3. Modeling of the host spectroscopic redshift efficiency (Vincenzi et al. 2021) instead of the efficiency of spectroscopic typing as in the DES-SN3YR spectroscopic sample (Smith et al. 2020a) and modeling the fraction of host-mismatched SNe (Qu et al. 2024).
4. Updated DES internal calibration (Sevilla-Noarbe et al. 2021) and cross-calibration (Brout et al. 2022b).
5. The inclusion of wavelength-dependent atmospheric effects on photometry (Lee et al. 2023) in addition to the chromatic corrections used in the DES-SN-3YR analysis (Lasker et al. 2019).
6. Using an unbinned systematic covariance matrix for cosmological constraints (Brout et al. 2020; Kessler et al. 2023).
7. Improved statistical validation of the methodology (Armstrong et al. 2023; Camilleri et al. 2024).

This analysis is the culmination of these works. In this paper, we provide the derivation of the DES-SN5YR distances and the assessment of the impact of systematic uncertainties on distances and cosmological fits. The unblinded cosmological constraints from the techniques established in this analysis are presented by DES Collaboration (2024).

The structure of this paper is presented in Figure 1. In Section 2, we present the DES-SN5YR data set. In Section 3, we briefly review the cosmological framework implemented in the analysis, with particular attention to how potential contamination from non-Ia SNe is accounted for in the cosmological fitting. In Section 4, we describe how simulations of DES-SN5YR are built. In Section 5, we present the inferred SN distances and our final Hubble diagram. In Section 6, we discuss the various sources of systematic uncertainties included in our analysis and present our systematic error budget. Finally, in Sections 7 and 8, we discuss our results and present our conclusions.



Section 2: Data preparation
2.1 DES-SN sample
2.2 Low-z external samples
2.3 Photometry
2.4 Calibration
2.5 SN host properties
2.6 SN Classification
2.7 Milky Way extinction corrections

Section 3: Cosmological framework
3.1 Light-curve fitting
3.2 SN distances
3.3 BEAMS
3.4 bias corrections
3.5 Intrinsic scatter modelling
3.6 Cosmology estimation

Section 4: Simulations
4.2 SN Ia simulations
4.3 non-Ia SN simulations
4.4 SN host simulations
4.5 Sims/data comparison

Section 5: Hubble diagram
5.1 Sample selection
5.2 Nuisance parameter
5.3 Hubble residuals and uncertainties

Section 6: Systematic uncertainties
6.1 LC fitting & Calibration
6.2 SN Ia intr properties
6.3 Milky Way extinction
6.4 Host association and survey modelling
6.5 SN Classification
6.6 Redshifts & Peculiar Vel

Section 7: Results and discussion

Figure 1. Overview of the paper.

2. Data

2.1. Dark Energy Survey SN Sample

The DES-SN program is a 5 yr survey using the Dark Energy Camera (DECam; Flaugher et al. 2015) on the Victor M. Blanco telescope (Cerro Tololo, Chile), covering 10 $\sim$3 deg^2 fields distributed across the DES footprint (two “E” fields, two “S” fields, three “X” and “C” fields; see Smith et al. 2020a, Figure 1 and Table 2). Two out of ten fields (“X3” and “C3”) have been observed to reach a single-visit depth of 24.5 mag in the r band (deep fields), while the remaining eight fields reach a single-visit depth of 23.5 mag. Only a small fraction of the DES-SN candidates have been spectroscopically followed up using several spectroscopic facilities. For the majority of the transients, host galaxy redshifts have been collected using the auxiliary Australian DES survey (OzDES), which used the 2dF fiber positioner and AAOmega spectrograph (Smith et al. 2004) on the Anglo-Australian Telescope (AAT) to collect host galaxy redshifts (Lidman et al. 2020). The SN sample collected by the DES-SN program is the largest and deepest cosmological SN sample from a single telescope to date (see Figure 2). Kessler et al. (2015) and Smith et al. (2020a) describe in detail the SN search strategy and spectroscopic follow-up associated with the DES-SN program.

2.2. Low-redshift Samples

We combine the DES-SN sample with various external low-redshift ($z < 0.1$) SN surveys. These include CfA3 (Hicken

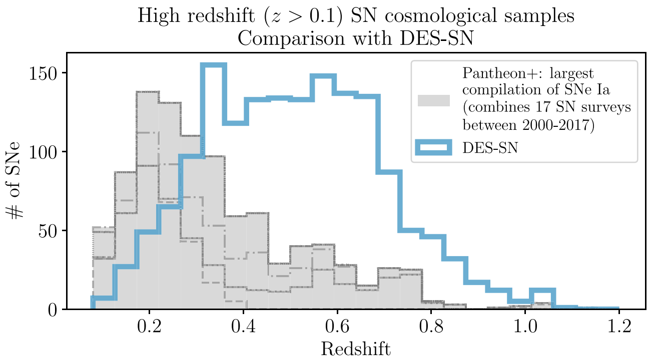


Figure 2. Redshift distribution of the DES-SN sample compared to the Pantheon+ compilation. DES-SN is the largest and deepest sample of SN Ia to date. The dashed lines represent the different surveys' redshift distributions contributing to the total Pantheon+ compilation. The Pantheon+ and the DES-SN samples presented in this paper have a small overlap of 146 SNe (previously published by Brout et al. 2019a).

et al. 2009), CfA4 (Hicken et al. 2012), Carnegie Supernova Project DR3 (CSP DR3; Krisciunas et al. 2017), and the Foundation SN sample (Foley et al. 2018). These external surveys span a redshift range of $0.01 < z < 0.1$ and provide a lever arm to improve constraints on the dark energy equation of state. For this analysis, we include only low-z SNe above redshift 0.025 to mitigate the effects of peculiar velocities. Finally, we add a 1% error floor in quadrature to the low-z SN photometry (2% for the Foundation z band), following Scolnic et al. (2018) and Jones et al. (2019).

We do not include other historical low-z SN samples, e.g., LOSS (Ganeshalingam et al. 2013), SOUSA,[62] or intermediate-redshift SN samples, e.g., the Sloan Digital Sky Survey (SDSS) SN sample (Sako et al. 2018). This choice is made to avoid including a larger number of systematic uncertainties in our analysis (for every survey, we need to account for additional systematics related to survey calibration and survey-specific selection effects) and emphasize the contribution of the DES-SN program at redshift $z > 0.1$.

2.3. SN Photometry

We measure the DES-SN photometry using the scene modeling photometry (SMP. Astier et al. 2013) pipeline presented by Brout et al. (2019a), which simultaneously models the time-varying SN flux and the time-independent background host galaxy flux. In comparison to faster difference imaging pipelines, this technique provides more accurate flux and flux uncertainty measurements. Sanchez et al. (2024) present a detailed comparison between the DES SMP photometry and photometry from difference imaging and demonstrate that the implementation of SMP significantly reduces (i) flux uncertainties and (ii) effects attributed to the so-called surface-brightness anomaly (i.e., unexplained flux scatter increasing with the host galaxy surface brightness at the SN location; Kessler et al. 2015, 2019a).

In addition, DES-SN photometry is corrected for wavelength-dependent atmospheric effects such as Differential Chromatic Refraction (DCR; Filippenko 1982) and wavelength-dependent (λ-dependent) seeing, which affect ground-based observations. DCR occurs because the index of refraction of our atmosphere is wavelength-dependent, while λ-dependent seeing is caused by variations in the atmospheric refractive index due to atmospheric turbulence. These two effects cause a color-dependent mis-modeling of the shape of the point-spread function (which is reconstructed using stars that are generally redder than the average SN at $z = 0$) and of the position of the SN. Lee et al. (2023) describe the methods used to correct DES-SN photometry for such wavelength-dependent atmospheric effects and assess their impact on DES-SN distance estimation and cosmological results. We do not include wavelength-dependent atmospheric corrections for external low-z samples.

2.4. Calibration

Accurate photometric calibration of DECam filters and intersurvey calibration is essential in SN cosmology to estimate SN brightnesses at different redshifts and when combining SNe from different surveys. For the DES-SN sample, calibration is performed in two stages.

First, DES images are internally calibrated using a catalog of 17 million tertiary standard stars within the DES footprint built using the Forward Global Calibration Method (FGCM) as conceived by Stubbs & Tonry (2006) and as implemented in DES by Burke et al. (2018). Not only does this method provide accurate (∼1%) absolute calibration, but it also provides excellent all-sky uniformity of <3 mmag for DES (Sevilla-Noarbe et al. 2021; Rykoff et al. 2023).

The FGCM tertiary standard star catalog provided in Burke et al. (2018) was utilized in the preliminary DES-SN3YR cosmological analysis. The FGCM catalog was updated in the period between DES-SN3YR and DES-SN5YR and here we use the `Y3GOLD` stellar catalogs as presented in Appendix 3 of Sevilla-Noarbe et al. (2021). The improvements are summarized as follows: (i) improved aperture photometry corrections, (ii) an update to the publicly released DES Y3A2 standard bandpass (see Sevilla-Noarbe et al. 2021), (iii) improved uniformity in years following the bad weather of year 3, (iv) improved astrometric solution utilizing longer temporal baseline, and (v) other technical and practical improvements.

Second, the tertiary standard stars are calibrated to primary standard stars to place them on the AB system. Within DES, AB offsets were calculated to the HST Caslpec standard star C26202 in Rykoff et al. (2023), given in Table 1. However, because SN Ia cosmology analyses combine multiple surveys to cover both low redshift and high redshift to obtain competitive cosmological constraints, here we utilize the calibration of Brout et al. (2022b; Supercal-Fragilistic) which is an improvement on the Scolnic et al. (2015; Supercal) method. This method consists of simultaneously cross-calibrating the FGCM catalog with the AB-calibrated stellar catalogs from numerous other wide-field surveys (e.g., PS1, SDSS, SNLS). Supercal-Fragilistic uses priors on each modern survey's published AB calibrations and refits for a new solution that minimizes the differences between surveys and mitigates potential systematic errors. Supercal-Fragilistic finds similar offsets as those found in Rykoff et al. (2023), but of larger magnitude (see Table 1); however, these offsets are consistent with each other given that the external data used to perform the calibration is largely independent. In this work, we have chosen to adopt the offsets from Supercal-Fragilistic because: (i) the low-z samples also analyzed in this work have been calibrated in Supercal-Fragilistic, (ii) this includes covariance between DECam filters and low-z filters (utilized in our distance

---

[62] Light curves are available at https://pbrown801.github.io/SOUSA/.

Table 1
DECam AB Offsets and Uncertainties

References	g	r	i	z
Rykoff et al. (2023)	$+0.001 \pm 0.011$	-0.003 ± 0.011	-0.001 ± 0.011	$+0.002 \pm 0.012$
Brout et al. (2022b)	$+0.002 \pm 0.006$	-0.009 ± 0.006	-0.007 ± 0.006	$+0.006 \pm 0.006$

likelihood in Equation (1)), (iii) Supercal-Fragilistic provides the mechanism to create multiple realizations of interfilter correlated calibrations from the Supercal-Fragilistic covariance matrix with the other low-z samples, (iv) Supercal-Fragilistic obtains smaller uncertainties due to the utilization of more external data. This change results in a ∼5 mmag color correction for $g - z$ (∼3 mmag for $g - i$) relative to what was used in DES-SN3YR.

The AB offset uncertainty reported in the C26202-based analysis of Rykoff et al. (2023) is ∼0.011 mag. The reported DES5YR uncertainties (stat+syst) in Supercal-Fragilistic covariance are roughly half of the size on the diagonal (6 mmag), which is the result of leveraging the cross-calibration of multiple surveys utilizing multiple primary standard stars. The full Supercal-Fragilistic covariance[63] is used to determine the effects of correlated systematic uncertainties in both light-curve fitting and in SALT3 model training. Systematic uncertainties due to absolute calibration of the DECam and low-z filters are discussed in Section 6.

2.5. Host Galaxy Association, Redshifts, and Host Properties

For each SN, we identify the host galaxy using the directional light radius (DLR) method presented by Sullivan et al. (2006) and Gupta et al. (2016). We define as "hostless" SNe for which no galaxy is detected with DLR < 4. The galaxies identified as likely hosts of DES transients are targeted using the AAOmega spectrograph on the 3.9 m AAT as part of the OzDES program (Yuan et al. 2015; Childress et al. 2017; Lidman et al. 2020). A full description of the different sources of redshifts used in our sample and the host spectroscopic redshift efficiency for the DES-SN sample are presented by Vincenzi et al. (2021) and Sanchez et al. (2024).

To characterize SN host galaxies, we mainly focus on two global host galaxy properties: the stellar mass ($M_\star$) and rest-frame $u - r$ color. These are the two properties we can most reliably estimate given the limited broadband photometry available for our SN hosts. For DES-SN hosts, these galaxy properties are measured using DES broadband photometry and, when available, u-band and JHK photometry from external surveys (Wiseman et al. 2020; Hartley et al. 2022). We use the galaxy spectral energy distribution (SED) fitting code by Sullivan et al. (2010) and the PÉGASE2 galaxy spectral templates (Fioc & Rocca-Volmerange 1997; Le Borgne & Rocca-Volmerange 2002), assuming a Kroupa (2001) initial mass function. In the DES-SN cosmological sample used in this work (see Section 5), we find that 68% of the DES-SN hosts are assigned to high stellar mass ($>10^{10} M_\odot$).

For consistency, we remeasure stellar masses and rest-frame $u - r$ colors of the low-z SN host galaxies using the same code and initial mass function used for the DES-SN hosts. We use optical and UV photometry[64] to ensure the same rest-frame wavelength coverage used for the DES-SN hosts. We find that a significant fraction of low-z hosts previously assigned a stellar mass lower than $10^{10} M_\odot$ are reassigned a larger stellar mass ($>10^{10} M_\odot$). As a result, the fraction of low-z SNe in high-mass hosts is 69%, compared to 59% in the previous analysis (we only consider SNe included in the cosmological sample presented in Section 5). The details of the host property measurements for DES-SN and for the external low-z samples are presented in Sanchez et al. (2024) and Kelsey et al. (2023).

Spectroscopic redshifts for the low-z sample are incorporated following the revisions presented by Carr et al. (2022). Low-z SN redshifts require additional corrections for peculiar velocities (that are negligible for high-redshift DES-SNe). The nominal peculiar velocities used for this analysis were determined by Peterson et al. (2022) and are based on 2M++ density fields (Carrick et al. 2015) with global parameters found in Said et al. (2020), combined with group velocities estimated by Tully (2015). We consider uncertainties on peculiar velocity estimates to be 250 km s^{-1} (Scolnic et al. 2018).

2.5.1. "Hostless" SNe and SNe in Faint Hosts

Given the requirement of a spectroscopic redshift from the identified host, we exclude from our sample SNe considered "hostless" (i.e., no host detected within four DLRs). Using the deep coadded images from Wiseman et al. (2020) and Qu et al. (2024), we estimate the fraction of likely SN Ia (probability of being an SN Ia larger than 50% according to different classification methods; see Section 2.6) for which no host is detected and find the fraction to be of 1.4% and 1.8% in the two catalogs, respectively (see also Qu et al. 2024, Section 5.1).

Given the requirement of a spectroscopic redshift from the identified SN host, in our sample, we not only exclude hostless SNe but also SNe found in galaxies fainter than ∼23.5 mag in the r band (the OzDES limit; see Vincenzi et al. 2021). This requirement introduces strong redshift-dependent selection effects in our sample and biases our sample toward SNe found in high-mass bright galaxies. In Section 4.4, we show that these host mass- and brightness-dependent selection effects are accurately modeled in our simulations and corrected for in our analysis.

2.6. Non-Ia Classification

In our baseline approach, we classify the DES-SN sample using the open-source algorithm SuperNNova (Möller & de Boissière 2020),[65] a photometric SN classifier based on recurrent neural networks. SuperNNova is trained to classify different types of transients using photometric data only (i.e., fluxes and flux uncertainties in different filters) and, optionally, redshift information. It does not rely on feature extraction or light-curve fitting and it uses machine learning techniques, e.g., recurrent neural networks. This is the first SN cosmological

---

[63] https://github.com/PantheonPlusSH0ES/DataRelease

[64] We use UV photometry from the Galaxy Evolution Explorer (Bianchi et al. 2017) and SDSS (u band).

[65] https://github.com/supernnova/SuperNNova

Table 2
BBC Input and Outputs Chart

BBC	Variants
Inputs	S3 fit parameters[a], $\sigma_{\rm S3fit}$, $P_{\rm Ia}$, biasCor simulations
Intermediate outputs	$\mu_{\rm bias}$ (Equation (7)), $\mathcal{P}_{\rm BEAMS(Ia)}$ (Equation (6)), M_ς (Equation (5)), $\sigma_{\rm floor}$ (Equation (9))
Outputs (for cosmology)	$\mu_{\rm obs}$, $\sigma_{\mu,\rm unbin}$ (Equations (8) and (10)), α, β, γ, $\sigma_{\rm gray}$

Note.
[a] SALT3 fitted parameters m_x, x_1, c.

analysis that exploits machine learning techniques for classification.

The training of SuperNNova (and most classification algorithms based on machine learning) requires large (>100,000) and representative samples of SN light curves (Möller & de Boissière 2020). For this reason, the subset of spectroscopically confirmed DES-SNe is not suitable as a training sample, and simulations are used instead. We train SuperNNova on realistic simulations of DES-like light curves, built following the approach described in Section 4 and by Vincenzi et al. (2021). Simulations include SNe Ia, peculiar SNe Ia, and core-collapse SNe generated using the core-collapse template library by Kessler et al. (2019b) and Vincenzi et al. (2019). A detailed analysis of training methods and performances of SuperNNova in the context of the DES-SN5YR analysis is presented by Vincenzi et al. (2021b) and Möller et al. (2022).

Using SuperNNova, we estimate for each SN event its probability of being an SN Ia, $P_{\rm Ia}$. These probabilities are then incorporated into the cosmological analysis as described in Section 3.3.

2.7. Milky Way Extinction Corrections

Milky Way (MW) extinction corrections are applied to the light-curve fitting model. In the analysis, we exclude SNe with Milky Way reddening larger than 0.25. The 10 DES-SN fields have been specifically chosen in Milky Way regions with low dust extinction (median reddening $E(B-V)_{\rm MW} < 0.02$); however, significant differences are observed from field to field (average $E(B-V)_{\rm MW}$ is <0.01 in the E,C fields, ∼0.02 in the X fields, and ∼0.03 in the S fields). SNe in the low-z SN samples have a median Milky Way extinction of 0.04 (low-redshift SN surveys generally require large sky coverage; therefore Milky Way regions with higher dust extinction cannot be avoided).

For each SN, we estimate $E(B-V)_{\rm MW}$ using the Milky Way reddening maps presented by Schlafly et al. (2010) and use the reddening law by Fitzpatrick (1999) with $R_V = 3.1$.

3. Cosmological Analysis Framework

In this section, we give an overview of the cosmological framework used to go from observed light curves to SN distances and cosmological fitting. SN distances are obtained after light-curve fitting (Section 3.1) using the Bayesian Estimation Applied to Multiple Species (BEAMS) with Bias Corrections (BBC) framework (Sections 3.3, 3.4, and 3.5). In Table 2, we present a schematic overview of the inputs and (intermediate and final) outputs of the BBC framework.

3.1. Light-curve Fitting

The first step to measure SN Ia distances is to perform light-curve fitting of the multiband photometry observed for each SN. This step is necessary to standardize SN brightnesses. In this analysis, we perform light-curve fitting using the SALT3 model framework (Kenworthy et al. 2021). The SALT3 model is defined by five SN-dependent parameters: redshift z, day of peak brightness ($t_{\rm peak}$), an amplitude term m_x or x_0 (where $m_x = -2.5\,\log_{10}(x_0)$ and it is approximately the B-band peak brightness of the SN), a color c parameter (that affects the overall amplitude of the model by a factor $\exp(c\cdot {\rm CL}(\lambda))$, where ${\rm CL}(\lambda)$ is the SALT3 model color law), and a stretch x_1 parameter. In the SALT3 fitting process, the best-fit values and uncertainties of each parameter are determined in order to measure SN distances (see Section 3.2). In our analysis, SN redshifts are known with high accuracy (spectroscopic redshifts included in our sample have an uncertainty of $\sim 10^{-4}$; see Section 2.5); therefore, parameter z is fixed in the light-curve fitting.

The SALT3 model is based on the widely used SALT2 model (Guy et al. 2007); however, it provides an improved estimation of uncertainties as a function of the phase and color and an extended central passband wavelength range of 2800–8000 Å (compared to a range of 2800–7000 Å in SALT2). The SALT3 model is trained on a compilation of 1083 SNe with 1207 spectra presented by Kenworthy et al. (2021). Taylor et al. (2023) showed that there is negligible difference in the SN cosmology results from the choice of SALT2 or SALT3 model, when the models are trained on the same input data. In this analysis, we train our own SALT3 model `SALT3.DES5YR`.

Following Taylor et al. (2023), we train the `SALT3.DES5YR` model using the sample from Kenworthy et al. (2021), based on the calibration values presented by Brout et al. (2022b). We choose to remove any observer-frame U-band training data from the training set, as calibration of the near-UV ground-based data is challenging (the atmospheric extinction is larger and variable from site to site and with airmass, the filter's characterization historically poorer, and the cross-calibration approach presented by Brout et al. 2022b is not applicable). This affects 97 SNe in the training sample, from CfA and miscellaneous low-z samples.[66]

Given the difficulty in training SN Ia in the UV, where the SN flux is low and good-quality SN data are limited, we avoid the far-UV and use passbands whose central wavelength (λ_b) satisfies 3500 Å$< (\lambda_b)/(1+z) <$ 8000 Å. The lack of good rest-frame UV-band modeling is an important limitation for our analysis because the DES-SN survey aims to probe the high-redshift SNe (where observer-frame optical is rest-frame UV).

3.2. Measuring SN Ia Distances

SN Ia distance moduli, $\mu_{\rm obs}$, are defined as (e.g., Tripp 1998; Astier et al. 2006)

$$\mu_{\rm obs} = m_x + \alpha x_1 - \beta c + \gamma G_{\rm host} - \mathcal{M} - \Delta\mu_{\rm bias}, \tag{1}$$

---

[66] Our training sample still includes u-band data from the more recent SDSS and CSP SN surveys, for which the filter transmissions have been measured and well characterized. For the older CfA U-band data, we do not have measured filter transmissions. This has caused several calibration issues, as highlighted by various cosmological analyses (Sullivan et al. 2011; Brout et al. 2022a).

where m_x, x_1, and c are the SALT3 light-curve parameters discussed in Section 3.1. The global nuisance parameters α and β set the amplitude of the stretch–luminosity and color–luminosity corrections, and $\mathcal{M}$ is the absolute magnitude of an SN Ia with $x_1 = 0$ and $c = 0$. The fourth term of the equation, $\gamma G_{\rm host}$, encapsulates any residual dependency between SN Ia luminosities and their host galaxy properties. This dependency is modeled as a step function, $\gamma G_{\rm host}$, defined as

$$\gamma G_{\rm host} = \begin{cases} +\gamma/2 & \text{if } P > P_{\rm step}, \\ -\gamma/2 & \text{otherwise}, \end{cases} \tag{2}$$

where P is an SN host galaxy property, usually stellar mass $M_\star$, γ is the size of the residual "step", and $P_{\rm step}$ is the threshold at which the step is measured, usually fixed to $10^{10} M_\odot$ for stellar mass.

Many cosmological analyses have shown that SNe Ia observed in high-mass galaxies ($M_* > 10^{10} M_\odot$) are approximately 0.07 mag brighter than SNe in lower-mass galaxies after color and light-curve stretch corrections (Lampeitl et al. 2010; Sullivan et al. 2010; Smith et al. 2020b; Kelsey et al. 2021). Recently, Brout & Scolnic (2021) and Kelsey et al. (2023) highlighted that this so-called "mass step" is highly color dependent: smaller (or negligible) for blue SNe and significant (>0.1 mag) for redder SNe.

Brout & Scolnic (2021, hereafter BS21), supported by the work of Salim & Narayanan (2020), propose that dust is the underlying cause of the SN mass step and show how different R_V values in high- and low-mass host galaxies could explain the observed brightness step.

In addition, Kelsey et al. (2023) analyzed the DES-SN sample and measured the SN brightness step between SNe found in intrinsically blue and intrinsically red host galaxies (a "color step," rather than a mass step). Kelsey et al. (2023) found a significant color step even after corrections for the dust-driven mass step and suggested that either stellar mass is not the optimal proxy to describe a dust-driven brightness step, or that other astrophysical factors (e.g., SN progenitor physics) might play an important role in explaining the dependence of SN Ia luminosities on their host galaxies. For this reason, in our analysis, we implement Equation (2) *either* assuming $P = M_\star$, $P_{\rm step} = 10^{10} M_\odot$ (the mass step, $\gamma_{M_\star}$) *or* $P = (u - r)$, $P_{\rm step} = 1$ (the color step, γ_{u-r}; see also Section 4.2.2).

Parameters α, β, and γ are determined prior to, and independently of, the cosmological parameters, using the approach presented by Marriner et al. (2011; see Section 3.3). The term β is treated as the observed, effective β, i.e., we do not fit separately for an intrinsic β and a "dust β" (or R_V) parameter (see further discussion in Section 3.4). We note that without a calibrated absolute distance scale, $\mathcal{M}_B$ is degenerate with the cosmological parameter H_0 and therefore is not addressed in this work. Finally, corrections for biases resulting from selection effects and analysis choices are applied in the $\Delta\mu_{\rm bias}$ term of Equation (1). These selection effects are determined from accurate simulations of the survey and using models of the residual scatter. We discuss the modeling and implementation of bias corrections in Section 3.4.

3.3. The BEAMS Approach

Photometric SN samples require the application of photometric classification algorithms to determine the SN types. We incorporate type Ia classification *probabilities* in the cosmological fits using the BEAMS framework (presented by Kunz et al. 2007, 2012; Newling et al. 2012).

The BEAMS approach was developed to incorporate SN probabilities and marginalize the overcontamination from non-Ia SNe while performing a cosmological fit. Kessler & Scolnic (2017) extended the BEAMS framework to include modeling and correction of selection effects, and to incorporate the Marriner et al. (2011) method of measuring nuisance parameters independent of cosmological parameters. This extended framework is referred to as BBC. In the BEAMS approach, the analytical form of the likelihood includes two terms, one that models the SN Ia population, $\mathcal{L}_{\rm Ia}$, and the other that models a population of contaminants, $\mathcal{L}_{\rm CC}$,

$$\prod_{i=1}^{N_{\rm SNe}} (\mathcal{L}_{\rm Ia}^i + \mathcal{L}_{\rm CC}^i). \tag{3}$$

The two terms of the likelihood, $\mathcal{L}_{\rm Ia}^i$ and $\mathcal{L}_{\rm CC}^i$, are defined as

$$\begin{aligned} \mathcal{L}_{\rm Ia}^i &= P_{\rm Ia}^i \times D_{\rm Ia}(z_i, \mu_{{\rm obs},i}, \mu_{\rm ref}) \\ \mathcal{L}_{\rm CC}^i &= (1 - P_{\rm Ia}^i) \times D_{\rm CC}(z_i, \mu_{{\rm obs},i}, \mu_{\rm ref}). \end{aligned} \tag{4}$$

The term $D_{\rm Ia}$ is defined as:

$$D_{\rm Ia} = \exp\left(-\frac{(\mu_{{\rm obs},i} + M_\zeta - \mu_{\rm ref}(z_i))^2}{\sigma_{\mu,i}^2}\right), \tag{5}$$

where $\mu_{{\rm obs},i}$ is defined in Equation (1), $\mu_{\rm ref}(z_i)$ is the distance modulus of the ith SN as predicted assuming a fixed reference cosmology (e.g., $\Omega_M = 0.3$, $w = -1$), and M_ζ is the offset quantifying by how much observations deviate from the reference cosmology. They *absorb* the cosmological information, enabling a cosmology-independent fit of the SN nuisance parameters (as shown by Marriner et al. 2011).[67] The M_ζ offsets are calculated for 20 redshift bins (ζ), equally spaced on a logarithmic scale, and they effectively constitute a "binned" version of the SN Hubble diagram (z_ζ, $M_\zeta + \mu_{\rm ref}(z_\zeta)$). However, we emphasize that we only use this binning to determine the nuisance parameters α, β; we do not use this binned Hubble diagram to fit cosmology as it has been shown to lead to an overestimate of systematic uncertainties (see Brout et al. 2020). We follow instead the unbinned approach described and validated by Kessler et al. (2023). The distance modulus uncertainties $\sigma_{\mu,i}$ are discussed in Section 3.5.

Finally, $D_{\rm CC}$ in Equation (4) is the contaminant likelihood term, and it models the non-Ia SN distance moduli distribution on the Hubble diagram. Core-collapse SNe are not standardized by the SALT3 framework; therefore it is not trivial to model $D_{\rm CC}$ analytically. In our baseline analysis, we empirically model the term $D_{\rm CC}$ using the core-collapse simulations described in Section 4 and test alternative approaches in the systematics analysis.

The two likelihood terms in Equation (4) can be used to estimate the BEAMS probability:

$$\mathcal{P}_{{\rm BEAMS(Ia)},i} = \frac{P_{\rm Ia}^i D_{\rm Ia}^i}{P_{\rm Ia}^i D_{\rm Ia}^i + (1 - P_{\rm Ia}^i) D_{\rm CC}^i}, \tag{6}$$

---

[67] The estimated SN distances are not dependent on the choice of the reference cosmology (see Kessler et al. 2023).

which effectively quantifies the likelihood of an SN belonging to the Ia population or the contaminants population, given not only its classification probability ($P^i_{\rm Ia}$) but also its inferred distance modulus and distance modulus uncertainty. A more detailed description of the BEAMS framework is given by Hlozek et al. (2012), Kunz et al. (2012), the BBC "binned" approach is described by Kessler & Scolnic (2017) and Vincenzi et al. (2021b), and the BBC "unbinned" approach used in this analysis is presented by Kessler et al. (2023).

3.4. Bias Corrections

All SN surveys are affected by selection effects introduced by their flux-limited nature. These selection effects can introduce systematic biases in cosmological analyses of SN Ia samples, and thus SN Ia distances are usually corrected for such expected biases (Equation (1)). The corrections are generally estimated using large SN Ia Monte Carlo simulations that accurately model the survey detection efficiency and other potential selection effects (Hamuy & Pinto 1999; Perrett et al. 2010; Betoule et al. 2014; Kessler et al. 2019a; Popovic et al. 2021).

In our analysis, bias corrections ($\Delta\mu_{\rm bias}$) are estimated using the BBC framework and large SN Ia simulations that model the different surveys considered in the analysis. We follow the approach presented in Popovic et al. (2021) referred to as "BBC4D." In BBC4D, the term $\Delta\mu_{\rm bias}$ is modeled as a function of the four observables z, x_1, c, $\log M_\star$, and it is defined as:

$$\Delta\mu_{\rm bias} = m_x + \alpha^{\rm true}x_1 - \beta^{\rm true}c + \gamma^{\rm true}G_{\rm host} + \mathcal{M}_B^{\rm true} - \mu^{\rm true}, \quad (7)$$

where $\alpha^{\rm true}$, $\gamma^{\rm true}$, $\mathcal{M}_B^{\rm true}$, and $\mu^{\rm true}$ are the true simulated values of nuisance parameters, the intrinsic SN brightness and distance modulus. The parameter $\beta^{\rm true}$ technically is not defined when simulating SNe following the dust-based model by BS21, in comparison to the historical approach of simulating a single β. In the BS21 model, a forward-modeled distribution of intrinsic color–luminosity relations, $\beta_{\rm int}$, and a distribution of dust R_V are combined. For this reason, in the calculation of bias corrections and uncertainties, an effective β must be assumed. While the choice of $\beta^{\rm true}$ in bias corrections has a negligible effect on the inferred cosmology (see discussion in Section 7.3), we set $\beta^{\rm true} = 2.87$. In the Popovic et al. (2021a) forward-modeling process discussed in Section 4.2, this value of β is determined to be the effective observed β.

In more simple terms, $\Delta\mu_{\rm bias}$ in Equation (7) is a correction applied to each measured SN distance given the observed SN z, x_1, c, and $\log M_\star$. These corrections are based on the distance biases estimated using *simulated* SNe that have the same *observed* distributions z, x_1, c, and $\log M_\star$. Since these distance corrections are estimated from very large simulations, we do not propagate negligible statistical uncertainties from the simulations. Instead, we test how these corrections vary based on different simulated SN Ia model parameters (e.g., vary $\beta^{\rm true}$ or other dust parameters) and simulation assumptions and we include these tests in our systematic budget (Section 6). Using realistic simulations of SN samples, Popovic et al. (2021) tested the ability of the BBC4D approach described in this section to estimate unbiased SN distances and recover the input α, β, γ, and SN intrinsic scatter $\sigma_{\rm int}$.

Throughout the analysis, we will also refer to the "BBC0D" approach, i.e., the approach of fixing $\Delta\mu_{\rm bias} = 0$ and ignoring bias corrections. This approach is not used for cosmology, but it is useful to estimate *raw* SN distances, removing any assumption on SN Ia intrinsic properties and removing the modeling of selection effects.

3.5. SN Distance Uncertainties and Intrinsic Scatter

Within BBC, the distance modulus uncertainties $\sigma_{\mu,i}$ in Equation (5) are described as

$$\sigma^2_{\mu,i} = f(z_i, c_i, M_{*,i})\sigma^2_{\rm S3fit,\mathit{i}} + \sigma^2_{\rm floor}(z_i, c_i, M_{*,i}) + \sigma^2_{z,i} + \sigma^2_{{\rm vpec},i} + \sigma^2_{{\rm lens},i}, \quad (8)$$

where $\sigma_{\rm S3fit,\mathit{i}}$ is computed from the SALT3 light-curve fit parameters,[68] $\sigma_{{\rm vpec},i}$ and $\sigma_{z,i}$ are uncertainties associated with estimates of peculiar velocities and spectroscopic redshifts, respectively, and $\sigma_{{\rm lens},i}$ denotes the uncertainties associated with weak lensing effects due to the large-scale structure the SN photons are traveling through. The terms $f(z_i, c_i, M_{*,i})$ and $\sigma_{\rm floor}(z_i, c_i, M_{*,i})$ are survey-specific scaling and additive factors that are estimated from the same simulations used for bias corrections to ensure that the reduced χ^2 in each cell of a $\{z_i, c_i, M_{*,i}\}$ grid is close to unity. The scaling term is introduced to account for Malmquist bias that suppresses fainter SNe. This bias results in naively computed uncertainties that are overestimated, and thus $f < 1$ by construction. Conversely, the additive term accounts for any additional scatter beyond the naively computed $\sigma_{\rm S3fit}$. It is the sum in quadrature of a gray term and a term that depends on the redshift, color, and host mass:

$$\sigma^2_{\rm floor}(z_i, c_i, M_{*,i}) = \sigma^2_{\rm scat}(z_i, c_i, M_{*,i}) + \sigma^2_{\rm gray}. \quad (9)$$

The two terms $f(\ldots)$ and $\sigma_{\rm scat(\ldots)}$ are computed from large simulations (also used to estimate bias corrections). If $f < 1$, we set $\sigma_{\rm floor} = 0$ to avoid negative covariances; otherwise, $f = 1$ and the $\sigma_{\rm floor}$ term is added. The term $\sigma_{\rm gray}$ is global (not survey specific), it is fitted within BBC and it enforces the reduced χ^2 of the BBC fit to be equal to one. This term is expected to be zero if the scatter model used in the simulations is accurate. This novel approach of modeling intrinsic scatter has been first introduced by Brout et al. (2022a).

In the previous DES-SN3YR analysis, we used only the scaling term, f, as described in Kessler & Scolnic (2017). With the introduction of the Brout & Scolnic (2021) intrinsic scatter model, however, we found that f can take on values much larger than unity, leading to pathologically large distance uncertainties. The $\sigma_{\rm floor}$ alternative avoids overly large uncertainties. When tested on simulations, both methods provide unbiased cosmological results.

Although $\sigma_{\mu,i}$ are used in the BBC fit, they are not suitable for a cosmology fit with an unbinned Hubble diagram containing photometrically classified events. Following Kessler et al. (2023), an unbinned Hubble diagram requires redefining the distance uncertainties,

$$\sigma_{\mu,i,{\rm final}} = \sigma_{\mu,i} \times S_\zeta \Big/ \sqrt{\mathcal{P}_{{\rm BEAMS(Ia)},i}}\,, \quad (10)$$

---

[68] The term $\sigma_{\rm S3fit,\mathit{i}}$ is defined as

$$\sigma_{\rm S3fit,\mathit{i}} = (\sigma_{m_x})^2 + (\alpha\sigma_{x_1})^2 + (\beta\sigma_c)^2 + 2\alpha C_{m_x,x_1} - 2\beta C_{m_x,c} + 2 - \alpha\beta C_{x_1,c}.$$

where $\mathcal{P}_{\mathrm{BEAMS(Ia)},i}$ is the BEAMS probability (see Equation (6) and Kessler et al. 2023) and S_ζ is a scale such that the weighted average uncertainty in each BBC-fitted redshift bin ζ is equal to the BBC-fitted offset uncertainty, σ_{M_ζ}. The average S_ζ value is 1.01. Kessler et al. (2023) used this prescription on 50 data-sized DES-SN5YR samples and showed that the w bias is consistent with zero at a level below 0.01 (including a cosmic microwave background (CMB)-like prior).

Finally, the SN distance uncertainties calculated as described in Equation (8) are renormalized for photometric SN samples. The renormalization is applied to all SNe and it is necessary to ensure that (i) likely SN contaminants have inflated distance uncertainties and are downweighted in the cosmological fit and (ii) uncertainties on the $\Delta\mu_\zeta$ offsets estimated with BBC for each redshift bin are equal to the weighted average of the distance uncertainties of the SNe in the bin. The formalism related to the renormalization of the SN distance uncertainties is described in detail in Kessler et al. (2023).

3.6. Covariance Matrix and Cosmological Parameter Estimation

The output of BBC is a set of SN distances (as well as their uncertainties) corrected for biases from selection effects and contamination. These SN distances are estimated for the nominal analysis and for a set of analysis variants, implemented to quantify systematic uncertainties (see Section 6) and to build the uncertainty covariance matrix.

The $N_{\mathrm{SNe}} \times N_{\mathrm{SNe}}$ uncertainty covariance matrix, $\boldsymbol{C}$ is defined as the sum of a (diagonal) statistical term ($\boldsymbol{C}_{\mathrm{stat}}$), and a systematic term ($\boldsymbol{C}_{\mathrm{syst}}$). Following Conley et al. (2011) and Brout et al. (2019b, Section 3.8.2), we compute the systematic covariance matrix, $\boldsymbol{C}_{\mathrm{syst}}^{ij}$, defined as

$$C_{\mathrm{syst}}^{ij} = \sum_{S=1}^{N_{\mathrm{syst}}} (\Delta\mu_{\mathrm{obs,S}}^{i})(\Delta\mu_{\mathrm{obs,S}}^{j}) W_S^2, \tag{11}$$

where $\Delta\mu_{\mathrm{Ia,S}}$ are the differences in SN distances after changing the systematic parameter S, W_S is the scale of the systematic S, and the indices i and j are iterated over the N_{SNe} in the analysis ($i, j = 1,\ldots,N_{\mathrm{SNe}}$).

SN distances and the uncertainty covariance matrix C are used in the final cosmological fit, and the χ^2 of the SN likelihood is defined as

$$\chi^2 = \Delta\mu^T \cdot C^{-1} \cdot \Delta\mu, \tag{12}$$

where $\Delta\mu$ is the N_{SNe}-dimensional vector $\{\mu_{\mathrm{obs},i} - \mu_{\mathrm{theory},i}(\Omega_M, w)\}_{i\,=\,1,\ldots,N_{\mathrm{SNe}}}$.

4. Simulations and Data Comparison

4.1. Overview of the Simulations

In this section, we present the set of simulations used in our analysis. These simulations are generated (i) to predict the distance biases affecting our SN samples (DES-SN and external low-redshift samples), (ii) to train photometric classification algorithms, and (iii) to model the core-collapse likelihood in the BBC fit (see Equation (4)). We compare our simulated DES-SN5YR samples with the observed DES-SN5YR sample. The selection criteria used to compile both the observed and simulated DES-SN5YR sample will be discussed in detail in Section 5.1.

Simulations are generated and analyzed using the SuperNova ANAlysis software (`SNANA`; Kessler et al. 2009),[69] integrated in the PIPPIN pipeline framework (Hinton & Brout 2020).[70]

Simulations are built upon the work of Kessler et al. (2019a) and Vincenzi et al. (2021). Kessler et al. (2019a) describe in detail the modeling and simulation of the DES-SN photometry and associated uncertainties, DES cadence and observing strategy, and DES detection efficiency and trigger logic to define candidates. Vincenzi et al. (2021) focus on the modeling of contamination from non-Ia SNe, simulations of SN host galaxies, and the characterization of selection effects introduced by the requirement of a spectroscopic redshift from SN host galaxies.

An `SNANA` simulation can generate realistic transient light curves from various spectrophotometric models of transients. In the DES-SN simulations, we include SNe Ia (Section 4.2) and various classes of SN contaminants (Section 4.3).

4.2. Simulation of SNe Ia

SNe Ia are simulated using the SALT3 framework. The SALT3 parameters (redshift, day of peak B-band brightness, stretch, and color) are simulated as follows. Redshifts are generated using SN Ia volumetric rates from Frohmaier et al. (2019) and simulated t_{peak} are uniformly distributed between 2013 August and 2018 March. The distribution of x_1 and the dependency of x_1 with host galaxy properties is empirically determined using the method presented by Popovic et al. (2021). In Figures 3 and 4, we compare the observed and simulated distributions of various SALT3 parameters, as well as the correlations between stretch and environmental variables. In general, our simulations model the trends and properties of the DES-SN sample very well.

4.2.1. SN Ia Intrinsic Scatter and Dust

While SALT3 light-curve fitting includes a single SN color parameter, in our simulations we introduce additional parameters to model intrinsic SN color and luminosity scatter separately from extrinsic SN dust.

We assume that the SN Ia intrinsic scatter, color distribution, and color–luminosity correlations are well described by the formalism presented by BS21, but with updated parameters following Popovic et al. (2021a). In the formalism introduced by BS21, the distribution of SN colors is modeled as the sum of an *intrinsic color* Gaussian component (described by mean c_{int} and standard deviation $\sigma_{c_{\mathrm{int}}}$) and a reddening tail due to *dust* (described as an exponentially decreasing function with the exponent scaled by τ_E). SN luminosity color corrections are modeled as $\beta_{\mathrm{int}} c_{\mathrm{int}} + R_V E(B-V)_{\mathrm{dust}}$. Assuming that the average R_V values in high and low-mass galaxies differ by approximately 1.25 reproduces the mass step across different SN colors.

The distributions of intrinsic color c_{int}, intrinsic β_{int}, R_V, and $E(B-V)_{\mathrm{dust}}$ in high- and low-mass galaxies are determined using the "Dust2Dust" fitting code presented by Popovic et al. (2021a). For different combinations of color/dust parameters, "Dust2Dust" generates synthetic SNANA SN simulations and fits them with BBC. The best-fit color/dust parameters are determined by iteratively comparing the SNANA-simulated

---

[69] https://github.com/RickKessler/SNANA

[70] https://github.com/dessn/Pippin

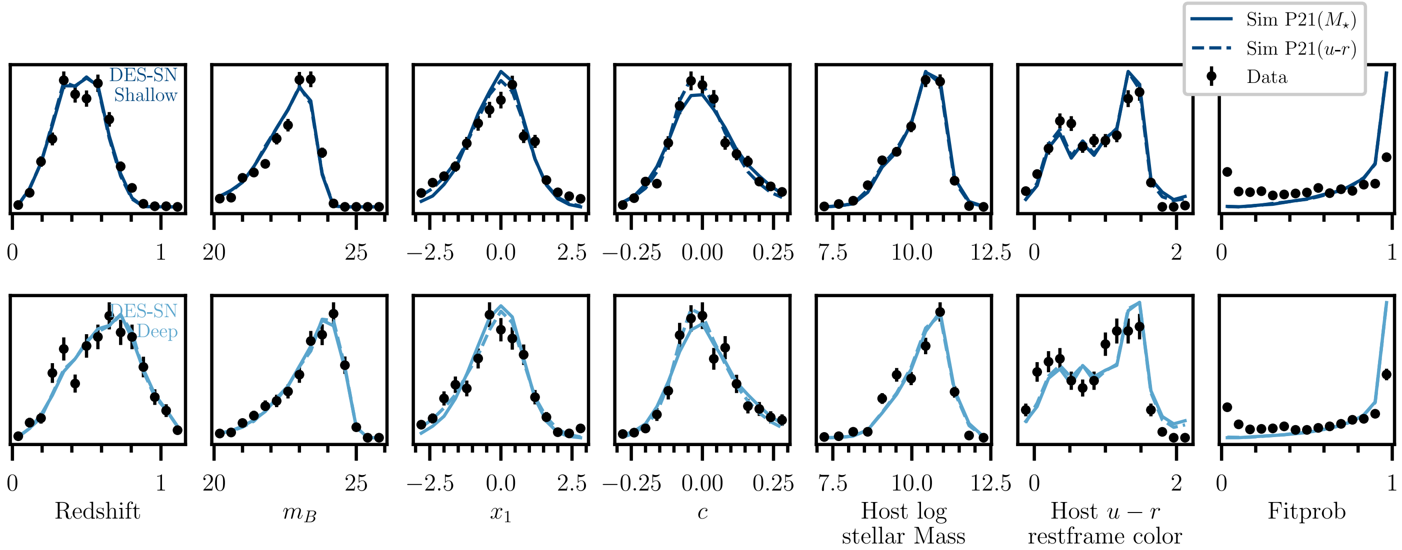


Figure 3. Comparison between observed and simulated SN and host galaxy properties. We present the comparison both for our baseline simulation (where dust and SN–host correlations are modeled as a function of the host galaxy stellar mass, P21 ($M_\star$)) and for an alternative simulation where dust and SN–host correlations are modeled as a function of the host galaxy rest-frame $u-r$ color (P21 ($u-r$); see Section 4). We present results for SNe in DES-SN Shallow fields (upper panels) and Deep fields (lower panels) separately.

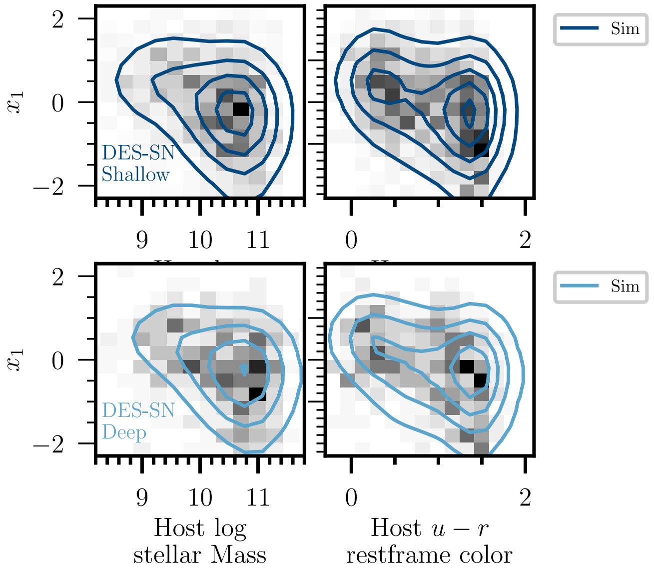


Figure 4. Comparison between observed and simulated correlations between SN stretch (x_1), and host galaxy properties (stellar mass on the left, and $u-r$ rest-frame color on the right). We present results for SNe in DES-SN Shallow (upper panels) and Deep (lower panels) separately. Data (gray 2D histogram) are compared to our baseline simulation (modeling SN–host correlations using host galaxy stellar mass, P21 ($M_\star$), solid contour lines). In general, we find excellent agreement between the observed and simulated distributions of the stretch as a function of different environmental variables and, when slicing the sample into bins in stellar mass or galaxy color, we find excellent agreement between the mean, standard deviation, and skewness of the simulated and observed x_1 distributions of the bin.

Hubble diagrams and the observed Hubble diagram. “Dust2-Dust” in particular uses the simulated and observed Hubble residuals calculated *without* applying bias corrections (i.e., the so-called “BBC0D” approach) as bias corrections are estimated making strong assumptions on SN color/dust distribution. In Figure 5, we present the comparison for our “Nominal” (also referred to as “P21($M_\star$)”) simulation, built using the BS21 formalism but with the “Dust2Dust” best-fit parameters, which are summarized in Table 3 and Appendix A. In our baseline dust-modeling approach, we do not include a mass step or color step (we set $\gamma = 0$ in Equation (2)); however, in Figure 5 we notice a difference in the residuals for $c < 0$ SNe that is not captured by simulations (but it is captured by fitting for γ; see Section 5.2).

4.2.2. SN Ia and Host Galaxy Color

Following the findings of Kelsey et al. (2023) and analogous studies (e.g., Briday et al. 2022; Wiseman et al. 2023), we develop an alternative set of SN Ia simulations that use host galaxy rest-frame $u-r$ color (instead of host galaxy *stellar mass*) as the galaxy proxy to model SN–host correlations.

First, we adapt the method presented by Popovic et al. (2021) to reproduce the (steeper) correlation between x_1 and host $u-r$ color (see Figure 4). Second, we run the “Dust2Dust” fitting code splitting SNe by host galaxy $u-r$ rest-frame color instead of host stellar mass and model dust parameters for intrinsically red and intrinsically blue galaxies (“P21($u-r$)” simulation; see Table 3). For this model, we do not simulate any additional “color step” (see Equation (2)). In Section 6.2.1, we discuss how the two simulations are implemented in the analysis.

4.3. Simulation of SN Contaminants

In our simulations, we include four classes of SN contaminants: two types of peculiar SN Ia (SN Iax and 91bg-like SNe) and two types of core-collapse SNe (stripped-envelope and hydrogen-rich SNe). SN Iax and 91bg are simulated using the templates and assumptions presented by Kessler et al. (2019b), with the revisions presented by Vincenzi et al. (2021). Core-collapse SNe are generated using templates by Vincenzi et al. (2019), using the rates by Strolger et al. (2015) and Shivvers et al. (2017) and luminosity functions in Li et al. (2011; with revisions by Vincenzi et al. 2021). A detailed description of the core-collapse simulations used for the DES analysis is presented by Vincenzi et al. (2021).

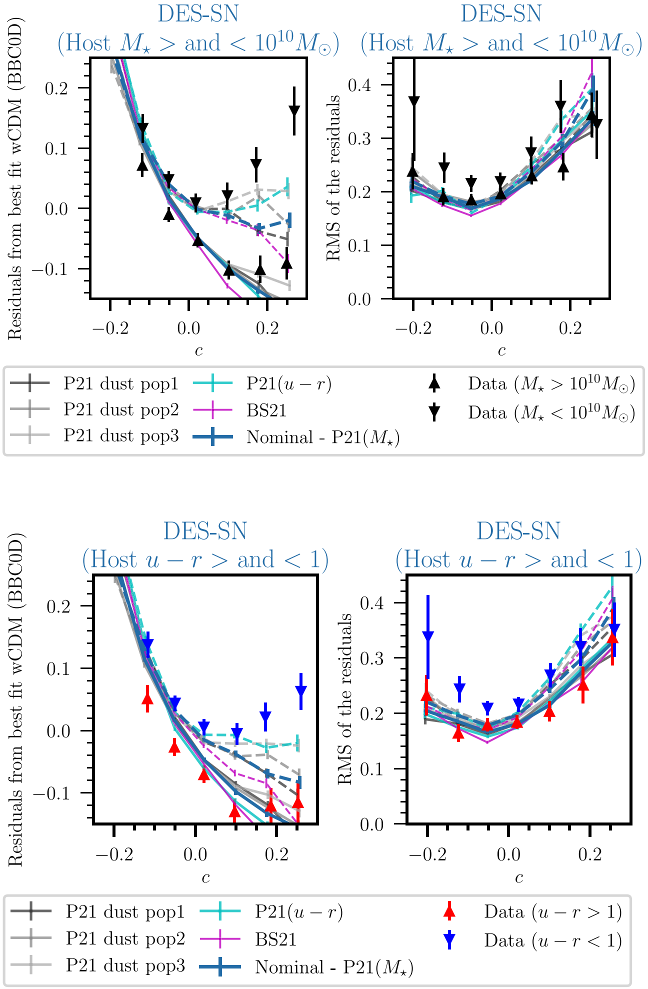


Figure 5. Average Hubble residuals (left) and rms error of the Hubble residuals (right) of the DES-SN sample as a function of the SALT3 color c (residuals are measured without applying any bias correction, i.e., in the "BBC0D" approach). The sample is split between high and low host mass galaxies (upper panels) and between intrinsically redder (rest-frame $u-r>1$) and intrinsically bluer (rest-frame $u-r<1$) galaxies (lower panels). Data are presented with upper-pointing and down-pointing triangles for high- and low-mass subsamples, respectively. Simulations are presented with solid and dashed lines for high-mass/red and low-mass/blue subsamples, respectively. The models used to generate simulations are described in Section 6.2.1. Equivalent plots for low-z are shown in Figure 19.

4.4. Simulation of Host Galaxies and Modeling Survey Selection Effects

We simulate host galaxies from the galaxy catalog presented by Qu et al. (2024). This catalog includes all galaxies detected in the coadded images of the DES-SN fields. For each galaxy, photometric redshifts (when spectroscopic redshifts are not available) are measured using the self-organizing map algorithm described in Qu et al. (2024), and galaxy properties are estimated from *griz* DES photometry using the same galaxy SED fitting code used for the DES-SN hosts (Sullivan et al. 2010). SNe Ia are assigned to galaxies following the SN rates presented by Wiseman et al. (2021). For peculiar SNe Ia and core-collapse SNe, we follow the same approach presented by Vincenzi et al. (2021).

One of the most important selection effects in the DES-SN sample is the requirement of a spectroscopic redshift from the SN host galaxy. The spectroscopic redshift efficiency for the DES-SN sample is presented by Vincenzi et al. (2021) and modeled as a function of the host galaxies' total brightness (`MAG_AUTO` in Source Extractor). In Figure 6, we compare the simulated and observed distributions of host galaxy stellar masses and brightness. At higher redshifts, the SN host galaxy mass distribution is progressively skewed toward higher values. This is primarily due to sample selection effects introduced by the requirement of a galaxy spectroscopic redshift. Our simulations very well reproduce the trends observed in the data. We highlight that the DES-SN spectroscopic redshift efficiency by Vincenzi et al. (2021) has not been adjusted to produce this agreement but was purely derived from DES-SN host galaxy data.

Table 3
Dust Parameters Used to Model the SNe Ia Population and Estimate Systematic Uncertainties

Parameter	P21($M_\star$)	BS21 [a]	P21($u-r$)
$c_{\rm int}$	−0.07	−0.084	−0.07
$\sigma_{c_{\rm int}}$	0.053	0.042	0.035
$\beta_{\rm int}$	2.07	1.98	1.86
$\sigma_{\beta_{\rm int}}$	0.22	0.35	0.21
R_V highM/red[b] hosts	1.66	1.25	1.5
σ_{R_V} highM/red hosts	0.95	1.3	1.0
R_V lowM/blue hosts	3.25	2.75	3.05
σ_{R_V} lowM/blue hosts	0.93	1.3	1.0
DES τ_E highM/red hosts	0.15	0.15	0.13
DES τ_E lowM/blue hosts	0.12	0.12	0.10
Low-z τ_E highM/red hosts	0.11	0.19	0.13
Low-z τ_E lowM/blue hosts	0.14	0.10	0.10

Notes. We present here three sets of dust parameters: "P21($M_\star$)" (best-fit parameters found using the algorithms by P21 dust and splitting the SN sample on stellar mass), "BS20" (using the original *fudged* dust parameters presented in the BS21 paper), "P21($u-r$)" (best-fit parameters found using the algorithms by P21 dust and splitting the SN sample on $u-r$ rest-frame color).

[a] *Original* BS21 dust parameters. BS21 did not use the dust parameter optimization code by Popovic et al. (2021a).

[b] "highM" refers to high-mass galaxies ($>10^{10}M_\odot$), "lowM" refers to low-mass galaxies ($<10^{10}M_\odot$), "red" refers to intrinsically red galaxies ($u-r>1$), and "blue" refers to intrinsically blue galaxies ($u-r<1$).

4.5. Simulations of the Low-redshift Samples

To simulate external low-z SN samples, we use the same inputs and modeling assumptions presented by Jones et al. (2017, 2019) and Scolnic et al. (2018), with minor adjustments mainly related to the modeling of host galaxy properties, as host galaxy properties have been remeasured for this analysis (see Section 2.5).

Intrinsic color distributions and dust properties for the low-z SN samples are the same as for the DES-SN sample (the dust-fitting code by Popovic et al. (2021a) is simultaneously run on the DES and low-z samples combined), and only the τ_E dust parameters are differentiated between high and low-z samples (see Table 3).

5. The Hubble Diagram

In Figure 7, we present the Hubble diagram of the DES-SN5YR analysis. This includes 1635 SNe from DES and 194 SNe from external low-z samples. The selection cuts applied are summarized in Table 4 and discussed below.

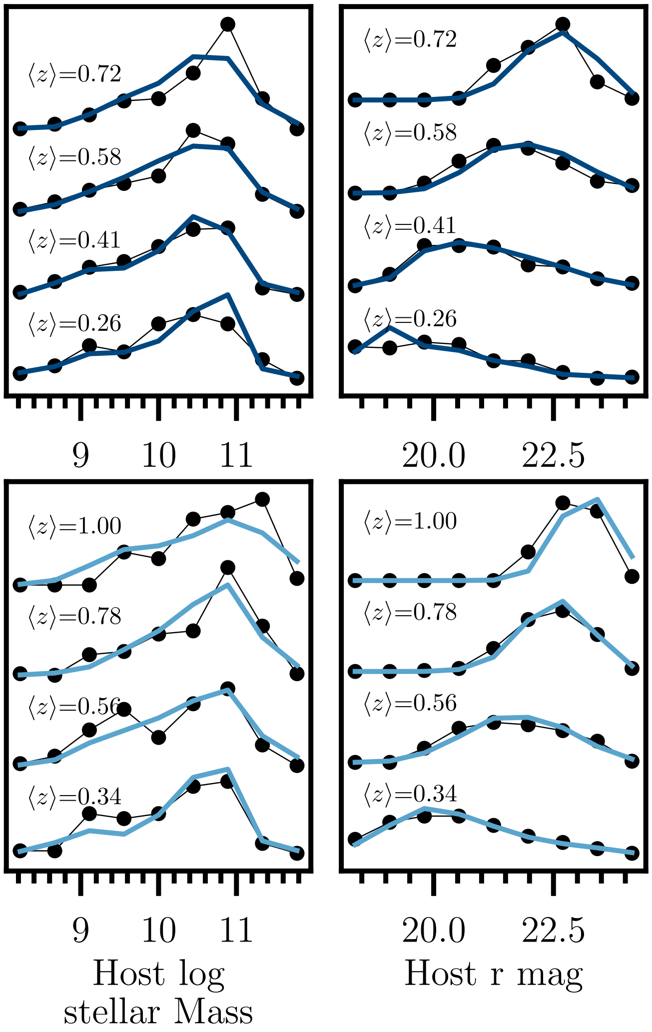


Figure 6. Comparison between observed (black symbols) and simulated (solid lines) distributions of host galaxy stellar mass (left) and host *r*-band brightness (right) for different redshift bins. We present the results for DES shallow and deep fields separately (upper and lower panels, respectively).

5.1. Sample Selection

First, we consider only DES-SNe with a host galaxy spectroscopic redshift. We do not include DES-SNe for which the host galaxy was not detected and the redshift information could only be inferred from the SN spectrum (the sample selection function of this sample is different; see Vincenzi et al. 2021, for more details). For this reason, the DES-SN sample presented in this analysis does not include all the DES-SNe Ia presented in the DES-SN3YR analysis (in this analysis, spectroscopically followed-up DES-SNe were selected, regardless of host spectroscopic redshift information; see Abbott et al. 2019; Brout et al. 2019a). To ensure good convergence of the SALT3 fit, we apply several cuts based on the quality of the light curve. We select DES-SNe with two bands that each have at least one detection with SNR > 5. In line with previous SN cosmological analyses, we require at least one observation before phase +5 days after the *B*-band peak (we do not require observations before the light-curve peak). In Appendix B, we discuss how the size of our sample and the final cosmological results change when requiring at least one detection before SN peak brightness.

We also apply SALT3-based selection cuts in both stretch and color ($-3 < x_1 < +3$ and $-0.3 < c < +0.3$) and $\sigma_{x_1} < 1$ and $\sigma_{t_{\rm peak}} < 2$ days. These cuts are commonly applied in SN analyses to select "normal" SNe Ia; they also significantly reduce contamination from peculiar SNe Ia and core-collapse SNe (Vincenzi et al. 2021; see Table 4).

In addition, applying bias corrections (as described in Section 3.4) constitutes a sample selection cut in itself. In BBC, for a small fraction of SNe, it is not possible to robustly determine bias corrections because the simulations of SN Ia used to calculate bias corrections do not have enough events in some regions of the SALT3 parameter space. As discussed in Vincenzi et al. (2021), this "valid bias correction" requirement implicitly reduces contamination because SN contaminants generally populate regions of the SALT3 parameter space that are atypical for SNe Ia. Moreover, applying Chauvenet's criterion, we iteratively apply a 4σ cut on the Hubble diagram residuals.

Finally, in order to build the unbinned systematics covariance matrix (see Section 3.6), we require each analysis variant to have the same set of events as the nominal model. This results in an additional 3.6% loss of SNe. In total, there are 1635 DES-SNe in the Cosmological Sample (see Table 4) for which 1499 (91%) are classified as likely Type Ia ($P_{\rm Ia} > 0.5$). When combined with the external low-z samples, the total number of SNe on our Hubble diagram is 1829. This sample is smaller than the photometric DES-SN sample presented by Möller et al. (2022), as it is built applying more stringent selection criteria.

In Figure 7, we present the weighted mean of Hubble residuals as a function of the redshift and do not observe any significant residual trend in our data. In Figure 8, we present the weighted mean of bias corrections $\mu_{\rm bias}$ (see Equation (7)) as a function of the redshift ($\mu_{\rm bias}$ as a function of the stretch and color is shown for the DES-SN3YR sample by Kessler et al. 2019a, Figures 9 and 10). These corrections become increasingly significant (~0.1 mag) at higher redshifts (redshift 0.5 for the SNe in the shallow DES-SN fields, 0.8 for the deep DES-SN fields), where selection effects have a more significant effect. Bias corrections also have a strong dependency on the SN color and SN stretch.

5.2. Nuisance Parameters

In our baseline analysis, we fit for the nuisance parameters α, β, γ, and $\sigma_{\rm gray}$, and we present the fitted values in Table 5. We do not fix γ to zero because we want to test for any residual brightness step that is not explained by our dust model and might be related to intrinsic SN astrophysics.

When combining the DES sample with the external low-z samples, we find $\alpha = 0.161 \pm 0.001$, $\beta = 3.12 \pm 0.03$. After accounting for dust law variation in the reported distances and uncertainties, we find $\gamma = 0.038 \pm 0.007$ (>5σ residual mass step) and a residual intrinsic scatter, $\sigma_{\rm gray}$, of 0.04. From the DES sample alone, we find consistent β ($\beta = 3.14 \pm 0.03$), slightly higher α ($\alpha = 0.170 \pm 0.004$), and a still significant mass step ($\gamma = 0.046 \pm 0.009$), with a residual intrinsic scatter of 0.04. We also compare nuisance parameters estimated from the DES-SN combined with the different low-z samples. We find that most nuisance parameters are consistent between the different sample combinations considered. The most significant discrepancy is in the fitted α for low-z samples alone (we find a significantly lower α compared to the value found when

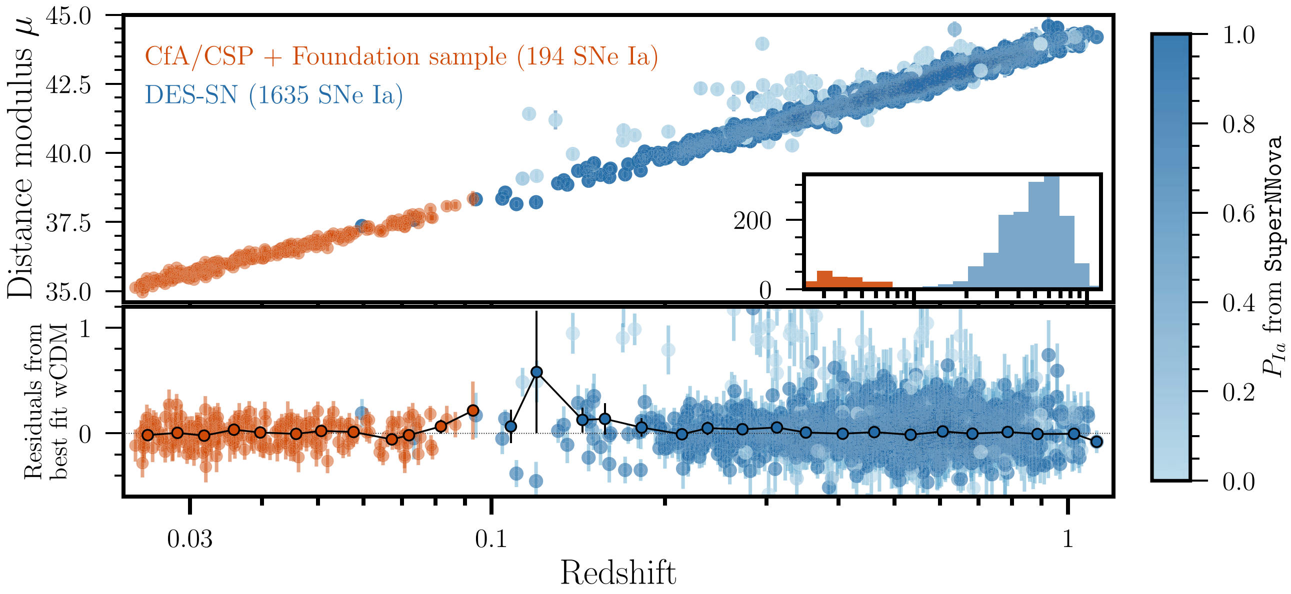


Figure 7. Hubble diagram (upper panel) and Hubble residuals (lower panel) combining the DES sample (blue) and external low-z SN samples (orange; see Section 2.2). For each SN event in the DES, we present classification probabilities, $P_{\rm Ia}$, estimated using the SuperNNova algorithm (see Section 2.6). Distance uncertainties are calculated using Equation (8) and do not include the renormalization term included in Equation (10).

Table 4
SN after Iteratively Applied Cuts

Requirement	Low-z	DES-SN		Total
		# all	$P_{\rm Ia} > 0.5$[a]	
Spec-z available, SALT3 fit converged and $z > 0.025$	247	3621	2200 [60%]	3868
"Normal SN Ia" ($\|x_1\| < 3$ and $\|c\| < 0.3$)	238	2449	2052 [83%]	2687
"Well constrained" (σ_{x1}<1, $\sigma_{t_{\rm peak}} < 2$)	238	1917	1639 [85%]	2155
Fit probability (fitprob >0.001)	221	1835	1627 [88%]	2056
Detected host galaxy	211	1806	1602 [88%]	2017
Spec-z from the host galaxy emission lines (not SN spectrum)[b]	211	1765	1563 [88%]	1976
Chauvenets criterion	209	1757	1557 [88%]	1966
Valid bias correction	204	1694	1541 [90%]	1898
Subsample of common CIDs across all systematic variants[c]	194	1635	1499 [91%]	1829
Cosmological Sample	194	1635	1499 [91%]	1829

Notes.
[a] Probabilities are from the SuperNNova (SNN) classifier trained on V19. In parenthesis, we report the percentage of likely SN Ia for each given cut.
[b] We exclude SNe for which a host was not detected and/or redshift information is from SN spectroscopic data, not from host galaxy emission lines.
[c] In order to build the systematic covariance matrix, we require having the same SNe across all systematic variants.

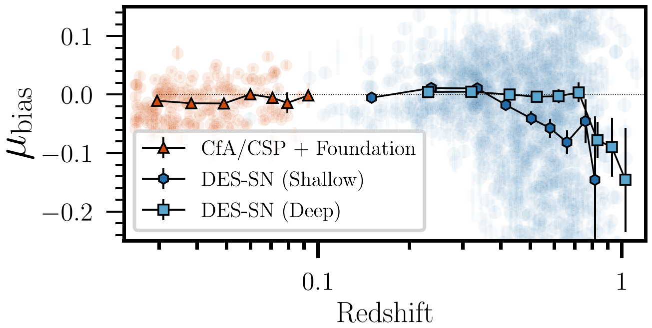


Figure 8. Bias corrections on μ for the DES sample (squares for deep fields, hexagon for shallow fields) and external low-z SN samples (triangles). Depending on the depth of the survey, corrections for selection effects start increasing at different redshifts.

Table 5
Nuisance Parameters and Cosmological Fit When Combining DES with Different Low-z External Samples and When Using DES Alone

Sample	$N_{\rm SNe}$	α	β	γ	$\sigma_{\rm gray}$	rms[a]
DES-SN + low-z	1829	0.161(1)	3.12(3)	0.038(7)	0.04	0.168
DES-SN only	1678	0.170(4)	3.14(3)	0.046(9)	0.04	0.177
DES-SN + Foundation	1796	0.166(3)	3.13(0)	0.042(8)	0.04	0.173
DES-SN + CfA/CSP	1760	0.167(3)	3.12(4)	0.043(9)	0.04	0.175
Foundation+ CfA/ CSP only	204	0.137(8)	2.90(10)	0.019(19)	0.06	0.118

Note.
[a] The rms error is measured by applying a cut of $P_{\rm Ia} > 0.5$ on the DES-SN sample.

including the DES-SN sample). We discuss these discrepancies in Section 7.1.3.

Regarding the estimation of the nuisance parameter γ (the mass step), we highlight that host mass measurement noise can introduce an underestimation of the step (and other nuisance parameters) if not properly propagated. In our analysis, we do calculate host mass uncertainties but we do not include them in the estimation of the host mass step because we consider the effect negligible (only 3% of the SN hosts are within 1σ of the mass threshold chosen to measure the mass step). In addition, we include variations in the host mass threshold in the systematics (see Section 6.2.3). Moreover, we note that Smith

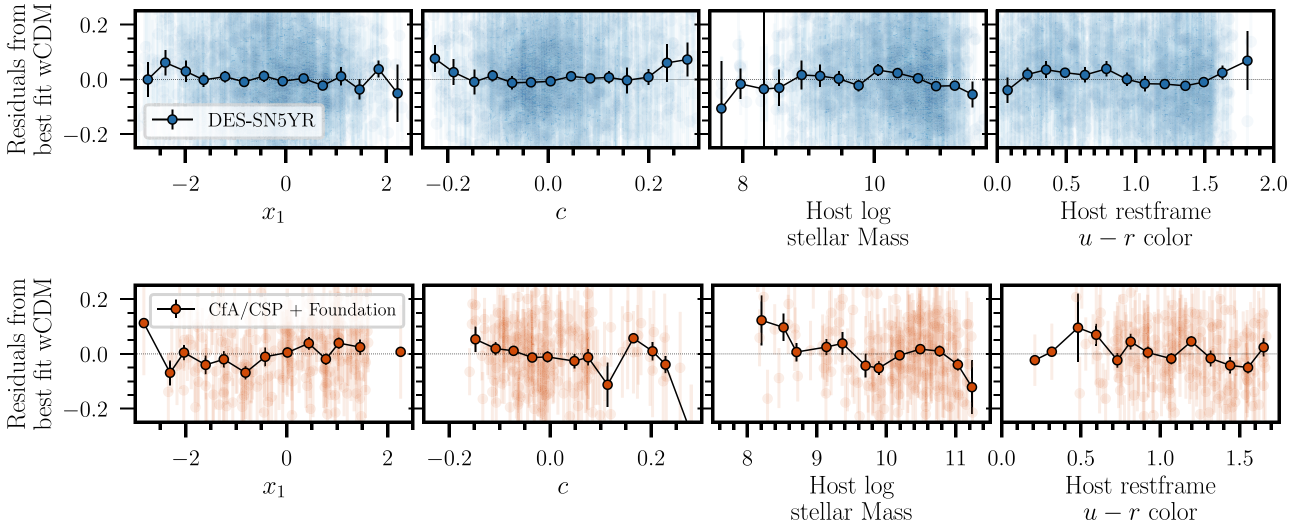


Figure 9. Hubble residuals from the DES sample (top row) and external low-z SN samples (bottom row) as a function of (from left to right) the SALT3 stretch x_1, SALT3 color c, host galaxy stellar mass, and host galaxy rest-frame $u-r$ color. Note that the residuals shown include bias corrections and the "mass step correction" following Equation (1).

et al. (2020b) showed that the mass steps recovered from the DES-SN3YR sample are consistent despite using different host mass estimators and assumptions.

5.3. Hubble Residuals

In Figures 7 and 9, we present Hubble residuals of DES-SNe and low-z SNe as a function of the redshift and other relevant SN and host parameters. These are the residuals estimated after applying the BBC4D approach described in Section 3.4.

In Figures 7 and 9, we do not observe any significant residual trend in our data. In particular, we highlight that the BBC4D approach (with the additional gray residual step γ) fully corrects the color trends highlighted in Figure 5. In Figure 5, the positive tail in the residuals of *observed* blue SNe is due to population migration (i.e., $c_{\rm obs} \neq c_{\rm true}$, due to measurement noise) described by Scolnic & Kessler (2016), and to the increasing color law at lower wavelengths (see Kenworthy et al. 2021, Figure 14). The BBC4D approach is designed to model these trends from simulations (for which both the true and observed colors, stretches, and brightnesses are known) and correct for them in the data.

Finally, Hubble residuals as a function of the host stellar mass and host rest-frame $u-r$ color present discontinuities at $10^{10} M_{\odot}$ and rest-frame $u-r$ color ~0.75, respectively, which might suggest that our modeling of discontinuous dust properties in high/low-mass host galaxies might be too simplistic.

6. Systematic Uncertainties

In this section, we describe the various sources of systematic uncertainties considered in the analysis. These are also summarized in Table 6.

6.1. Calibration and Light-curve Modeling

In this section, we discuss all the sources of systematics uncertainties related to the calibration of the DES-SN Ia fluxes and of the samples of SNe Ia that are used in the training of the SALT3 light-curve model.

The photometric systems of DES, Foundation, and the other low-z SN samples are cross-calibrated using the large and uniform sky coverage of the public Pan-STARRS stellar photometry catalog (Brout et al. 2022b). In this cross-calibration approach, the filter zero-point and mean wavelength in all systems are fitted simultaneously in order to produce a calibration uncertainty covariance matrix[71] that can be used in cosmological model constraints. The calibration uncertainty covariance matrix is used to randomly draw 10 mock realizations of zero-point calibration offsets and effective mean wavelength shifts. These correlated shifts are applied to recalibrate the SALT3 training sample and produce 10 perturbations of the SALT3 model. Following this approach, calibration uncertainties and light-curve modeling uncertainties are propagated *simultaneously* to the light-curve fitting (and not decoupled as in most previous SN cosmological analyses). The calibration uncertainty covariance matrix implemented in our analysis is presented by Brout et al. (2022b) and the relative set of SALT3 surfaces is presented by Taylor et al. (2023). In Tables 6 and 7, we label as "SALT3+Calibration" the different SALT3 surfaces implemented to propagate cross-calibration uncertainties.

In addition, we consider uncertainties associated with the fundamental flux calibration of the HST CALSPEC standards. These uncertainties are estimated to be 5 mmag/7000 Å (Bohlin et al. 2014).

6.2. SN Ia Properties and Astrophysics

In this section, we discuss sources of systematics uncertainties related to the astrophysics of SN Ia and their host galaxies. Assumptions on SN Ia intrinsic properties, their correlations with host galaxy properties, and their evolution with redshift primarily affect the bias corrections.

6.2.1. Intrinsic Scatter Model

As discussed in Section 4.2, the SNe Ia intrinsic scatter is modeled using dust-based formalism introduced by BS21. For our nominal analysis, we use the best-fit dust parameters

---

[71] https://github.com/PantheonPlusSH0ES/DataRelease/tree/main/Pantheon+_Data/2_CALIBRATION/FRAGILISTIC_COVARIANCE.npz

Table 6
Sources of Uncertainty

Baseline	Size[a]	Systematic	Label
Calibration and Light-curve Modeling (Section 6.1)			
SALT3 surfaces and ZP	1/10	10 covariance realizations	"SALT3+Calibration"
HST Calspec 2020 Update	1	5 mmag/7000 Å	"HST Calspec"
SN Ia properties and astrophysics (Section 6.2)			
Dust-based model Popovic et al. (2021a; "P21($M_\star$)")	1/3	3 realizations from MCMC dust model fitting code	"P21 dust pop 1/2/3"
	1	Original BS21 dust parameters	"BS21"
	1	Splitting on $u-r$	"P21($u-r$)"
Empirical modeling of $x_1-M_\star$ correlations	1	Modeling SN age following Wiseman et al. (2022)	"Model SN age"
No α evolution	1	$\alpha(z)=\alpha_0+\alpha_1\times z$	"α Evolution"
No β evolution	1	$\beta(z)=\beta_0+\beta_1\times z$	"β Evolution"
No γ evolution	1	$\gamma(z)=\gamma_0+\gamma_1\times z$	"γ Evolution"
Mass step location at $10^{10}M_\odot$	1	$10^{10.3}M_\odot$	"Mass Location"
$\sigma_{\rm int}$ modeling with scaling+additive scatter terms (Equation (9))	1	Scaling term only	"$\sigma_{\rm int}$ modeling"
Milky Way extinction (Section 6.3)			
MW scaling Schlafly & Finkbeiner (2011)	1	5% scaling	"MW scaling"
MW color law $R_V=3.1$ and F99	1/3	$R_V=3.0$ and CCM	"MW color law"
Host and survey modeling (Section 6.4)			
SN host catalog by Qu et al. (2024)	1	SN host catalog using DES-SVA galaxy catalog	"DES SV catalog"
Efficiency $\epsilon_z^{\rm spec}$ presented by V21	1	Shift of ± 0.2 mag in the efficiency curves	"Shift in host spec eff"
Contamination and photometric classifiers (Section 6.5)			
Classification using SuperNNova	1	SCONE, SNIRF	
Classifier training sample simulated using V19 templates	1	J17 templates, DES-CC templates ("`SuperNNova` training")	
Core-collapse SN prior to using V19 simulation	1	Polynomial fit as in Hlozek et al. (2012)	"CC SN prior"
Redshift (Section 6.6)			
Peculiar velocities using 2M ++	1	2M ++(Line-of-sight integration) or 2MRS	"Pec Velocities"
No redshift shift	1/6	$\Delta z=4\times10^{-5}$	"Redshift shift"

Note.
[a] Weighting is adopted for each source of systematic uncertainty when building the systematic covariance matrix (see also Equation (11)). In Section 6, we provide an explanation for the weights that are different from 1.

determined using the Markov Chain Monte Carlo (MCMC) fitting code "Dust2Dust" by P21 dust and splitting SNe by their host galaxy stellar mass (see values summarized in Table 3). In addition to the baseline approach (also referred to as "P21($M_\star$)"), we consider the following variations of dust-based intrinsic scatter models:

1. We randomly draw three sets of dust parameters from the MCMC chains produced by "Dust2Dust". The three realizations are presented in Figure 5. We refer to these models as "P21 population 1", "P21 population 2," and "P21 population 3".
2. We use the dust parameters *originally* presented in BS21 (see Table 3).
3. The "Dust2Dust" best fits when splitting SNe by host galaxy $u-r$ rest-frame color instead of host stellar mass. We refer to this model as "P21($u-r$)". For this model, we measure the color step γ_{u-r} (see Section 3.2).

In Figure 5, we show how the different models listed above reproduce the observed correlations between Hubble residuals and SN color, both for SNe in high- and low-mass galaxies and for SNe in red and blue galaxies.

Historically, most SN cosmological analyses have included the two following intrinsic scatter models:

1. The model presented by Guy et al. (2010; generally referred to as "G10"), according to which the SN luminosity dispersion is mostly (70%) wavelength-independent and 25% chromatic.
2. The model presented by Chotard et al. (2011; referred to as "C11") according to which the SN luminosity dispersion is mostly chromatic dependent (70%).

We do not include these models in our analysis as they are highly disfavored by both publicly released (Brout et al. 2022a) and the DES-SN5YR data in this work (see discussion in Section 7.1.1).

6.2.2. Modeling of Residual Intrinsic Scatter and Distance Uncertainties

In Equations (8) and (9), we present how uncertainties and residual intrinsic scatter floor ($\sigma_{\rm gray}$) are modeled in our analysis. The scaling and additive terms in Equations (8) and (9) are used to deflate and inflate SN distance uncertainties so that the reduced χ^2 of the cosmological fitting is close to unity across different regions of the redshift/color/host stellar mass parameter space. The additive term encapsulates any unaccounted-for SN intrinsic scatter.

We test the alternative approach of fitting the unexplained intrinsic scatter as a constant floor and using the scaling term only to inflate/deflate uncertainties when necessary. The two approaches should be fundamentally identical (and, in fact, when testing the two methods on 25 simulations, we recover the input cosmology in both cases). However, the approach used in our nominal analysis allows us to directly test whether our simulations reproduce SN Ia intrinsic scatter by testing whether $\sigma_{\rm gray}$ is 0.

Table 7
Size of Systematic Uncertainty (SN-only, No CMB Prior)

Systematic	$\sigma_{w,\mathrm{syst}}$ [a]	$\%\sigma_{\mathrm{tot}}$	δw_{syst} [b]
Total Stat+Syst	0.152	100	−0.032
Total Statistical	0.132	87	0.000
Total Systematic ($\mathcal{C}_{\mathrm{unbin}}$)	0.076	50	−0.032
Calibration and LC model	0.057	15	...
SALT3+Calibration	0.052	34	−0.036
HST Calspec	0.006	4	0.002
SN Ia astrophysics	0.133	35	...
P21 dust pop 1	0.019	12	−0.010
P21 dust pop 2	0.024	16	0.003
P21 dust pop 3	0.020	13	−0.004
P21$(u-r)$	0.000	0	0.048
Dust model as in BS21	0.027	18	−0.006
Model SN age (Section 6.2.3)	0.000	0	0.048
Change $\alpha\beta$ initial estimate	0.002	1	0.000
α Evolution	0.020	13	−0.008
β Evolution	0.000	0	−0.007
γ Evolution	0.011	7	−0.001
Mass step location	0.000	0	−0.002
σ_{int} modeling	0.013	8	−0.002
Milky Way extinction	0.034	9	...
MW 5% scaling	0.020	13	−0.011
MW color law CCM	0.014	9	−0.003
Survey modeling	0.015	4	...
DES SV catalog	0.009	6	0.002
Shift $\epsilon_z^{\mathrm{spec}}$	0.005	4	0.002
Contamination	0.028	7	...
Classifier SCONE	0.006	4	−0.000
Classifier SNIRF	0.013	9	−0.003
SuperNNova different training	0.006	4	−0.000
Core-collapse SN prior	0.003	2	−0.000
Redshift	0.037	10	...
Redshift shift	0.012	8	0.002
Peculiar velocities	0.025	16	−0.012

Notes. A detailed description of the different sources of systematics and the labeling conventions are presented in Section 6 and Table 6.

[a] The quadrature sum of systematic uncertainties is larger than the total systematic uncertainty. Internal correlations in the sample cause the effects of some systematics to partially cancel out when considering the full covariance matrix.

[b] Shift in w when including ONLY this systematic.

6.2.3. Modeling Host Galaxies and SN–Host Galaxy Correlations

In our nominal analysis, we model correlations between the SN stretch x_1 and SN host stellar mass $M_\star$ following the empirical approach presented by Popovic et al. (2021). However, we incorporate in our systematic error budget an alternative "galaxy-driven" approach that models x_1–$M_\star$ correlations starting from our current knowledge of the underlying astrophysics causing this correlation.

The galaxy-driven model used in this analysis is presented by Wiseman et al. (2022, hereafter W22). This model is based on the SN rates and SN delay time distributions presented by Wiseman et al. (2021).

The model uses galaxy evolution models to generate mock catalogs of galaxies and their properties, e.g., the stellar population age, stellar mass, star formation rate, and observed optical photometry. For each galaxy, the distribution of SN progenitor ages is determined by convolving the SN delay time distribution by Wiseman et al. (2021) with the galaxies' star formation histories and stellar populations.

Given a galaxy and its associated SN age estimate, the stretch parameter x_1 is assigned following the prescription presented by Nicolas et al. (2021; old and young SN progenitors are associated with two separate distributions determined from external nearby SN sample and represented in Figure 10). In this galaxy-driven approach, x_1–$M_\star$ correlations are the result of a physically motivated modeling of correlations between SN age and SN host galaxy mass.

In Figure 10, we present the distribution of SN stretch and SN host galaxy properties simulated using this galaxy-driven approach and compare it to the observed properties for both DES-SN and external low-z samples. We find good agreement between simulated and measured SN stretch and SN host stellar mass distributions, with the exception of low-mass galaxies. However, we find that the simulations built using this alternative model significantly underestimate the number of intrinsically red host galaxies (host rest-frame $u-r>1$). This is likely to be caused by an oversimplified approach to modeling galaxy quenching in the initial galaxy catalog. Moreover, correlations between x_1 and host properties observed in the data are steeper than what is reproduced by this alternative simulation, and this suggests that a revision of the modeling proposed by Nicolas et al. (2021) is required.

Despite the discrepancies between observations and simulations, we include this model in the systematic error budget as it provides an astrophysically motivated method to model SN–host galaxy correlations.

Finally, we include as an additional systematic uncertainty shifting the splitting point to measure the mass step from $10^{10}M_\odot$ to $10^{10.3}M_\odot$, since the typical uncertainty on our stellar mass estimates is 0.3 dex.

6.2.4. Standardization Parameter Evolution

Given the wide redshift range probed by the DES-SN sample (see Figure 2), we test for the evolution of the standardization parameters α, β, and γ, as well as the inferred σ_{floor} (see Equations (8) and (9)). We define $\alpha(z)=\alpha_0+\alpha_1\times z$, and similarly for $\beta(z)$ and $\gamma(z)$. We present and discuss our results in Section 7.1.3.

6.3. Milky Way Extinction Corrections

Inaccurate Milky Way extinction corrections can introduce biases in cosmology, especially because Milky Way extinction affects low- and high-redshift SNe differently (the average Milky Way reddening in low-redshift SNe is twice the average Milky Way reddening in DES-SNe; see Section 2.7).

For our systematic analysis, we test the effect of a global 5% scaling on Milky Way corrections, following the reanalysis and uncertainties presented by Schlafly & Finkbeiner (2011). As discussed in Schlafly et al. (2010), the Milky Way reddening law favored by the data is a Fitzpatrick (1999) reddening law with $R_V=3.1$. However, we conservatively include a systematic uncertainty in the Milky Way reddening law and analyze the data using the Cardelli et al. (1989) color law. The Cardelli et al. (1989) color law has the second lowest χ^2 when compared to the extinction derived from star photometry (twice the χ^2 associated with the Fitzpatrick 1999 reddening law; see Schlafly et al. 2010, Section 5.1.3); therefore, we weight the Milky Way color law systematic by a factor of $1/3$ ($W_S^2=1/3$; see Equation (11)).

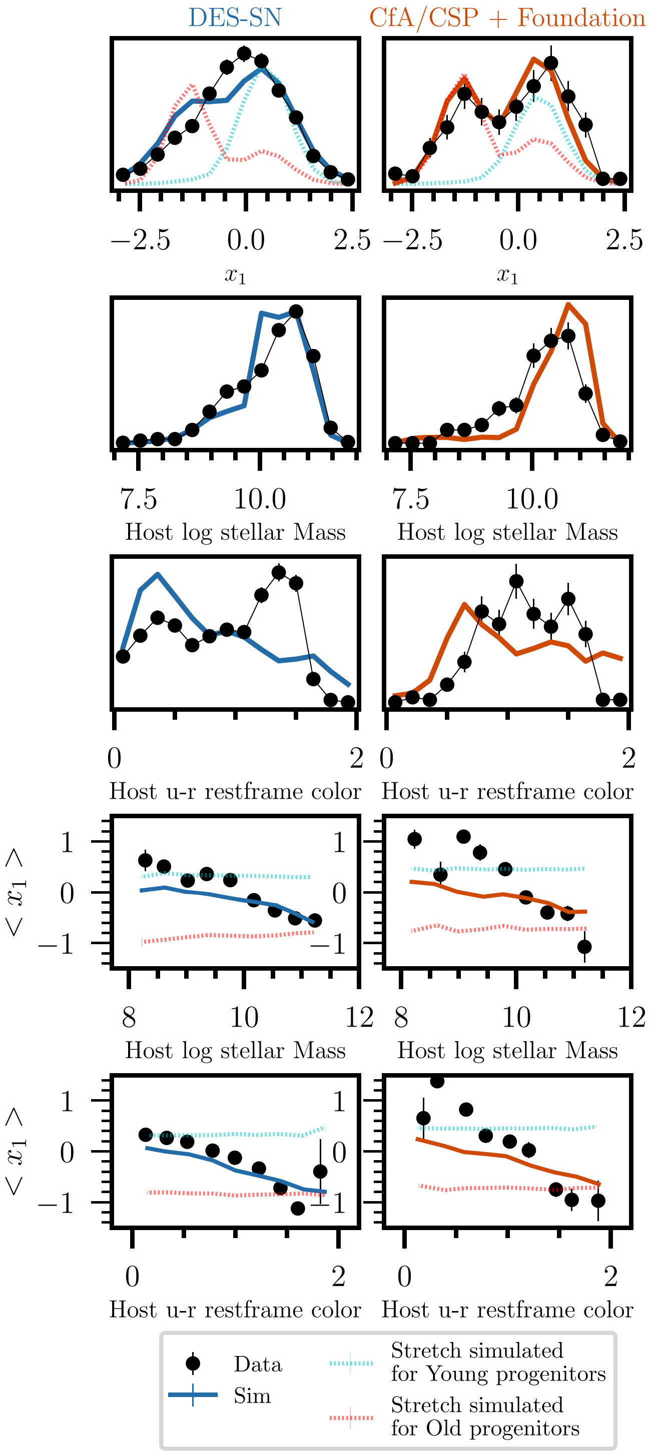


Figure 10. Comparison between data and alternative simulations generated using the galaxy-driven model by Wiseman et al. (2022). We compare the distributions of (from top to bottom) SN x_1, SN host stellar mass, SN host $u - r$ rest-frame color, and correlations between x_1 and host $M_\star$ and host $u - r$ rest-frame color. We present our results for DES-SN (left) and low-z samples (right). We also present the underlying distributions of young and old progenitors from which SN x_1 values are inferred. A similar comparison using our baseline simulations is presented in Figure 3.

6.4. Host Association and Survey Modeling

In this section, we discuss systematic uncertainties related to the modeling of host association and host-related survey selection effects.

6.4.1. Host Galaxy Mismatch

Host galaxy mis-association can occur for various reasons: the true host and/or the true host outskirts are too faint to be detected and a brighter, apparently closer (in terms of DLR) galaxy is identified as the likely host instead; or the SN is far from the true host, and fails our DLR cut. In the first case, deeper images can reduce host mis-association.

After upgrading from the shallower DES science verification (SV) images to the deeper coadded images, Wiseman et al. (2020) found that less than 1.1% of the DES-SN candidates change host galaxies, thus providing a preliminary estimate of the fraction of host mis-associations expected in the DES sample.

Qu et al. (2024) provide a more robust assessment of host mis-association in DES-SN and estimated the percentage of misidentified hosts to be 1.7%. This prediction is based on high-quality simulations, built from a galaxy catalog that accurately models galaxy light profiles. These simulations reproduce the observed distributions of SN–galaxy separations and DLRs for DES-SN (Figures 5 and 6 of Qu et al. 2024).

When comparing the two deep SN catalogs by Wiseman et al. (2020) and Qu et al. (2024), we find that 7 of 1635 DES-SNe (0.5%) have uncertain assigned hosts (i.e., they are assigned to different hosts depending on the catalog used). Therefore, we assume this source of systematic is negligible for our analysis.

6.4.2. Galaxy Catalog for SN Host Simulations

In our analysis, we test two different galaxy catalogs for simulations of DES-SN hosts. The first is generated using deep DES coadded images and is described in Section 4.4 and presented in detail by Qu et al. (2024). The second is compiled using galaxies detected in the (shallower) DES SV images and is presented by Smith et al. (2020b). The DES SV catalog includes photo-z from template fitting techniques and galaxy properties measured using the galaxy SED fitting code by Sullivan et al. (2006). This second galaxy catalog is inferior in depth compared to our nominal catalog; however, we decided to include it in our systematic error budget because it has been used in many previous DES-SN studies (Kessler et al. 2019a; Smith et al. 2020b; Vincenzi et al. 2021, 2021b; Wiseman et al. 2021). A comparison between simulations generated using the two galaxy catalogs is presented in Figure 11.

6.4.3. OzDES Selection Effects

We apply a shift of 0.2 mag to the efficiency curves presented by Vincenzi et al. (2021) and include this in the systematic error budget. A comparison between simulations generated using the alternative spectroscopic redshift efficiency is presented in Figure 11.

6.5. Contamination and Photometric Classification

To correct for core-collapse SN contamination, we test different photometric classification algorithms and different training methods. We include the different classification variants in the systematic error budget.

For the baseline analysis, we use the algorithm SuperNNova by Möller & de Boissière (2020). Vincenzi et al. (2021b) present a detailed analysis of the training and performances of SuperNNova in the context of the DES-SN cosmological

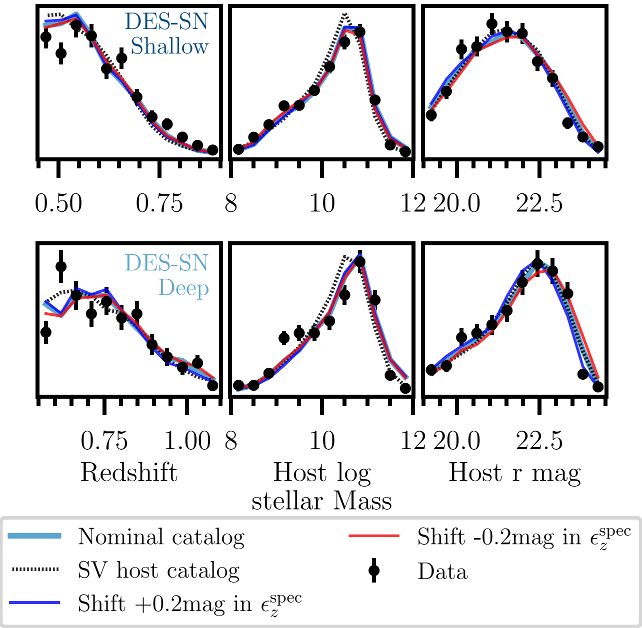


Figure 11. Comparison between observed and simulated distributions of SN redshifts (zoomed at higher redshifts; left panel), host galaxy mass (center panel), and host galaxy r-band magnitudes (the Kron-like `MAG_AUTO` magnitudes measured with SEXTRACTOR (Bertin & Arnouts 1996) from the r-band deep coadds; right panel) for our nominal simulation and two variants: using the galaxy catalog from DES Science Verification images (instead of catalog by Qu et al. 2024) and using a shifted spectroscopic redshift efficiency instead of the nominal efficiency by Vincenzi et al. (2021).

analysis. For our baseline analysis, we train SuperNNova using the simulations presented in Section 4. For our systematic analysis, we train SuperNNova using two alternative and independent libraries of core-collapse SN templates: the one presented by Jones et al. (2017, hereafter J17) and the one built from core-collapse SNe observed in DES (R. Hounsell et al. 2023, in preparation, hereafter DES-CC).

As an alternative to SuperNNova we consider two additional classification algorithms: the classifier SCONE by Qu et al. (2021) and the Supernova Identification with Random Forest (SNIRF) algorithm.[72] We train SCONE and SNIRF on the same set of simulations used to train the SuperNNova baseline model.

Finally, we test different approaches to modeling the contamination likelihood term in Equation (4). While the baseline approach uses simulations based on Vincenzi et al. (2021), we test the approach of using a polynomial fitting as in Hlozek et al. (2012; see Vincenzi et al. 2021b, for a detailed comparison of the two methods).

6.6. Redshift and Peculiar Velocity Corrections

All SN redshifts are corrected for peculiar velocities and converted to the CMB frame. For the nominal analysis, we measure peculiar velocity corrections from 2M++. For systematics, we test two alternative approaches for the correction of peculiar velocities, both discussed in Peterson et al. (2022). The first approach uses the 2M++ corrections integrating over the line-of-sight relation between distance and the measured redshift. The second approach is to use the 2MRS (Lilow & Nusser 2021) peculiar velocity map. The two approaches both have $W_S = 0.7$ (see Equation (11)) so that their sum in quadrature results in an effective contribution of 1. The details of how peculiar velocity uncertainties are incorporated into the systematic covariance matrix $\mathcal{C}_{\rm syst}$ are presented in Brout et al. (2022a, Section 3.1.3).

Additionally, we account for potential biases due to a local void or other systematic redshift error and apply a systematic redshift shift of 4×10^{-5} (Calcino & Davis 2017).

7. Discussion

7.1. Systematic Uncertainty Budget on w

The sources of systematic uncertainties described in Section 6 are used to construct the covariance matrix $\mathcal{C}$ (Section 3.6). The unbinned Hubble diagram and $\mathcal{C}$ are used in a cosmological fit to determine constraints on cosmological parameters w and Ω_M using a flat wCDM model. Here we present our sensitivity to cosmological parameters; the final unblinded results are presented in the DES key paper (DES Collaboration 2024). The total uncertainty budget on w is presented in Figure 12 and Table 7.

In Figure 12, we compare statistical-only and systematic-only uncertainties on w. We present uncertainties both when measuring cosmological constraints from SN only (DES and low-z external SNe), and when combining SNe with CMB (see inset in Figure 12). The contribution from systematic uncertainties only is evaluated as $\sqrt{\sigma^2_{w,\rm tot} - \sigma^2_{w,\rm stat}}$, where $\sigma_{w,\rm tot}$ is the total uncertainty on w measured including the covariance matrix $\mathcal{C}_{\rm syst}$ described in Equation (11).

For a flat wCDM model considering SNe alone, we find that statistical uncertainties on w dominate at $\sigma^{w\rm CDM}_{w,\rm stat} = 0.132$ which increases slightly to $\sigma^{w\rm CDM}_{w,\rm stat+syst} = 0.152$ when including both systematic and statistical uncertainties. When adding a CMB prior to the cosmological fit, we find $\sigma^{w\rm CDM}_{w,\rm stat} = 0.032$ and $\sigma^{w\rm CDM}_{w,\rm stat+syst} = 0.036$. Both with and without the prior the statistical uncertainties dominate over the systematic ones. CMB measurements are highly complementary to SN measurements and have an orthogonal direction of degeneracy on the Ω_M–w plane for a flat wCDM model. For this reason, combining SNe and CMB significantly reduces uncertainties on w, and also reduces the impact of systematics for those that primarily move the SN contours along the direction of SN degeneracy.

To separately evaluate the contribution from each source of systematic uncertainty, we evaluate the w uncertainty using $\mathcal{C}_{\rm stat+syst}$ with a single systematic and compare it to the stat-only uncertainty. In Figure 12, we show each systematic uncertainty contribution with a separate color/pattern.

In Table 7, we summarize the size of systematic uncertainties visually presented in Figure 12 (in the table, we only focus on results determined using SN only, without a CMB prior) and additionally, we present the observed shifts on best-fit w when including statistical+systematic covariance matrix compared to statistical only.

For the purpose of interpreting Table 7, we make three important caveats. First, the quadrature sum of all systematic uncertainties presented in Table 7 is larger than the total systematic uncertainty on w. The difference is due to internal correlations in the sample that cause the effects of some systematics to partly cancel out when considering the full covariance matrix.[73] For the same reason, the δw values

---

[72] https://github.com/evevkovacs/ML-SN-Classifier

[73] For the purpose of Figure 13, every systematic contribution is rescaled so that their sum is the total systematic uncertainty, 0.076 (or 0.019 when including the CMB prior).

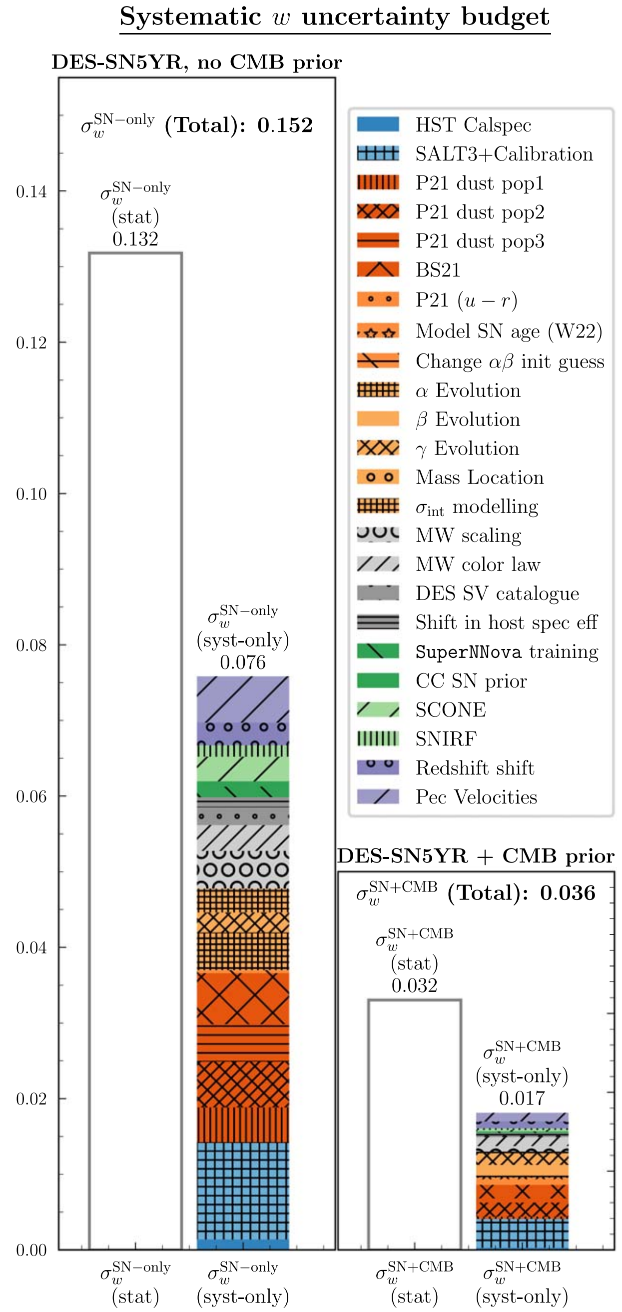


Figure 12. Systematic error budget on w and comparison with statistical uncertainty on w. We present results both with and without including the Planck Collaboration (2020) CMB prior (left and right plots, respectively; note the different y-axis scale in the two plots). The different sources of systematic uncertainty considered in this analysis are presented in Tables 6 and 7 and described in detail in Section 6.

measured for each systematic separately do not necessarily sum to the overall δw.

Second, the size of a single systematic, $\sigma_{w,\rm syst}$, and the size of the related δw, are not necessarily correlated and some systematics might have a small impact on $\sigma_{w,\rm syst}$ but cause a large δw (i.e., model SN age using W22 or using the P21$(u-r)$ intrinsic scatter model), or vice versa (i.e., including γ evolution). This difference between $\sigma_{w,\rm syst}$ and δw can arise if a systematic results in significantly less/more scatter in the Hubble residuals or in a significantly better/worse maximum likelihood.

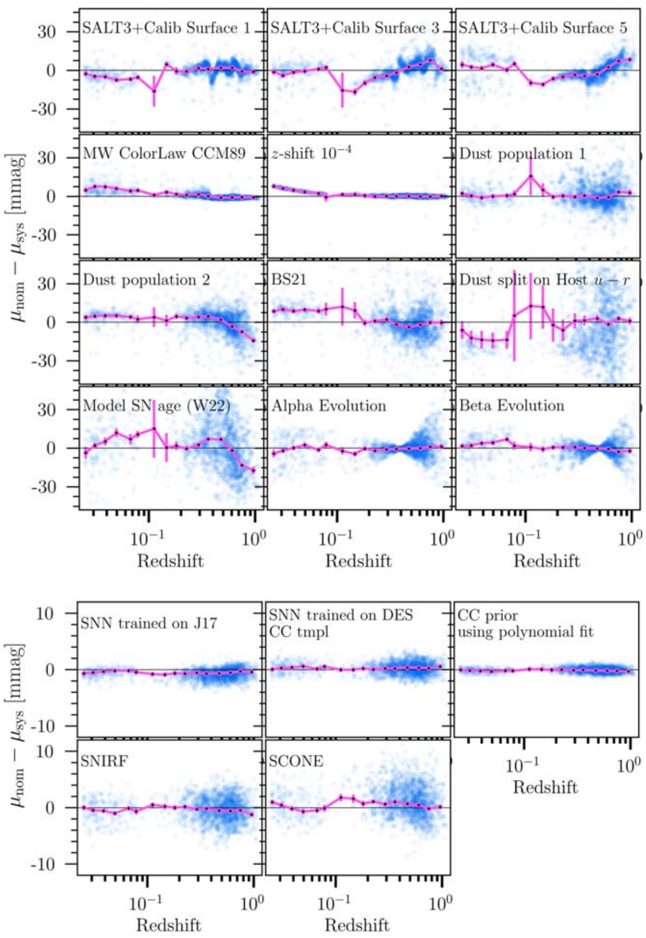


Figure 13. Effects of different sources of systematic uncertainties on SN distances. Upper panel: systematic uncertainties presented in Table 6, excluding systematics related to contamination and classification. Lower panel: contamination-related systematics. We note that the range of the y-axis is a fifth compared to the upper panel.

Third, the complementary test of *removing* one systematic at a time (instead of *adding* one source of systematic at a time) can be done (thus taking into account any internal cancellation between different sources of systematics). We performed this test and found that the systematic contributions are similar (with the exception of calibration systematics, which, when excluded, do not cause a significant reduction of the overall systematics error budget as expected).

In order to understand why some sources of systematics have larger $\sigma_{w,\rm syst}$ and/or δw, it is useful to look at changes in SN distances and fitting χ^2 when running the nominal analysis and the analysis run changing the systematic S. For a subsample of systematics, we present differences in SN distances in Figure 13, while in Table 8 we report differences in fitting chi-squared, fitted nuisance parameters and cosmological best fits. We compute the fitting chi-squared for each systematic variant as

$$\chi^2_S = \sum_{i=1}^{N_{\rm likely-Ia}} \frac{(\mu^i_{{\rm obs},S} - \mu_{{\rm best-fit},S}(z_i))^2}{\sigma^2_{\mu,\,{\rm Nominal}}}, \tag{13}$$

where $\mu_{{\rm obs},S}$ denotes the SN distances determined when changing the systematic S, $\mu_{{\rm best\text{-}fit},S}$ denotes the theoretical distances given the best-fit cosmology when adopting

Table 8

BBC-fitted Nuisance Parameters, and Cosmology-fitted χ^2_S, Hubble Residuals Rms Error and Shift in Best-fit w for Different Analysis Variants

	BBC fit				Cosmology fit (SN-only, no CMB Prior)		
Model	α[a]	β[a]	γ[a]	$\sigma_{\rm gray}$	$\Delta\chi^2_S$[b]	rms	$\Delta w_{\rm stat}$[c]
				DES-SN+low-z data			
Nominal[d]	0.161(1)	3.12(3)	0.038(7)	0.04	⋯	0.168	⋯
BiasCor sim: P21 dust pop 1	0.162(0)	3.17(0)	0.043(7)	0.06	−1	0.168	−0.038
BiasCor sim: P21 dust pop 2	0.160(3)	3.06(3)	0.033(8)	0.04	+5	0.169	0.079
BiasCor sim: P21 dust pop 3	0.161(0)	3.09(0)	0.040(8)	0.05	+18	0.17	0.039
BiasCor sim: Original BS21 param.	0.161(4)	3.18(1)	0.026(8)	0.06	+9	0.169	−0.110
BiasCor sim: P21($u-r$; fit γ_{u-r})[e]	0.158(0)	3.11(0)	0.033(7)	0.06	+16	0.169	0.066
BiasCor sim: model SN age BS21	0.148(3)	3.06(3)	0.017(8)	0.05	+1	0.168	0.113
Change $\alpha\beta$ init estimate	0.171(3)	3.34(0)	0.037(8)	0.05	+0	0.168	0.002
α evolution	0.149(0)	3.12(0)	0.036(8)	0.04	−5	0.168	−0.007
β evolution	0.161(3)	2.99(0)	0.037(8)	0.04	+0	0.168	−0.009
γ evolution	0.161(0)	3.12(0)	0.046(14)	0.04	+0	0.168	−0.006
BiasCor sim: no x_1–$M_\star$ correlations	0.152(2)	3.07(2)	0.019(8)	0.04	+6	0.167	0.006
BiasCor sim: G10 (no dust model)	0.157(6)	3.20(6)	0.055(9)	0.11	+102	0.172	−0.120
				DES-SN+low-z simulations (average of 25 independent simulations)[f]			
Nominal[d]	0.144(2)	2.83(4)	0.002(8)	0.00	⋯	0.160(4)	⋯
BiasCor sim: P21 dust pop 1	0.146(3)	2.90(5)	0.006(7)	0.00	−1(**4**)	0.161(4)	0.008
BiasCor sim: P21 dust pop 2	0.143(2)	2.77(4)	−0.002(8)	0.00	5(**4**)	0.161(4)	0.015
BiasCor sim: P21 dust pop 3	0.146(3)	2.80(4)	0.005(8)	0.00	0(**8**)	0.161(4)	0.113
BiasCor sim: Original BS21 param.	0.146(2)	2.89(5)	−0.009(8)	0.00	1(**5**)	0.161(4)	−0.042
BiasCor sim: P21($u-r$; fit γ_{u-r})[e]	0.141(3)	2.87(4)	−0.020(8)	0.00	+35(**11**)	0.163(4)	0.051
BiasCor sim: model SN age BS21	0.137(3)	2.79(4)	−0.009(8)	0.00	19(**12**)	0.162(4)	0.110
Change $\alpha\beta$ init estimate	0.154(2)	3.05(4)	0.002(7)	0.00	1(**0**)	0.160(4)	0.004
α evolution	0.148(6)	2.83(4)	0.002(8)	0.00	−1(**2**)	0.160(4)	0.006
β evolution	0.144(2)	2.88(7)	0.002(8)	0.00	−2(**2**)	0.160(4)	0.007
γ evolution	0.144(2)	2.83(4)	−0.003(15)	0.00	−1(**2**)	0.160(4)	0.005
BiasCor sim: no x_1–$M_\star$ correlations	0.135(3)	2.77(4)	−0.011(7)	0.00(0)	+7(**6**)	0.150(5)	0.083
BiasCor sim: G10 (no dust model)	0.148(4)	2.87(9)	0.030(8)	0.07(0)	+26(**10**)	0.151(5)	−0.069

Notes. The χ^2_S and rms error are measured after applying a cut of $P_{\rm Ia} > 0.5$. In the χ^2_S calculations (see Equation (13)), residuals are measured with respect to the best-fit cosmology estimated from each analysis variant, and always using residual uncertainties from the nominal analysis. A detailed description of the different sources of systematics and the labeling conventions are presented in Section 6 and Table 6.

[a] For data, reported uncertainties are the uncertainties from the BBC fit. For simulations, uncertainties are estimated as the standard deviations from 25 simulations. The nuisance parameters input values in the simulations are $\alpha = 0.145$ and $\gamma = 0.0$. For color corrections, the values are reported in Table 3.

[b] This is the increase (or decrease) in χ^2_S (see Equation (13)) compared to nominal (i.e., P21($M_\star$)).

[c] w shift from different intrinsic scatter models in biasCor sim ($C_{\rm syst}$ is not used), and without CMB prior in cosmology fit.

[d] Nominal is P21($M_\star$) and the step is measured by splitting the host galaxy's stellar mass ($\gamma_{M_\star}$). For simulations, χ^2_S uncertainties are determined as the standard deviation over 25 realizations and are presented in bold.

[e] For the alternative model P21($u-r$), the dust is modeled for blue/red galaxies, and the step is defined and measured as a "color" step, γ_{u-r} (i.e., the difference in brightness between SNe found in blue/red galaxies).

[f] Reported uncertainties are estimated as the standard deviations from 25 simulations. The 25 simulations are all generated using the nominal simulation approach.

systematic S in the analysis, and $\sigma_{\mu,\rm Nominal}$ denotes the uncertainties on $\mu_{\rm Ia}$ determined for the nominal analysis. As described in Section 3.5, SN distance uncertainties are rescaled and inflated using the terms $f(z_i, c_i, M_{\star,i})$ and $\sigma_{\rm floor}(z_i, c_i, M_{\star,i})$ (see Equation (8)). These terms are estimated from simulations and can vary from systematic to systematic. In the χ^2_S calculations, we fix the distance uncertainties $\sigma_{\mu,\ \rm Nominal}$ for all systematics to ensure that changes in χ^2_S ($\Delta\chi^2_S$) are not driven by inflated/reduced uncertainties, but *by an effective change in the modeling of Hubble residuals*. The χ^2_S are measured using "likely SNe Ia," i.e., SNe with a $P_{\rm Ia} > 0.5$.

To estimate the significance of the observed changes in χ^2_S and best-fit w, we generate a set of 25 simulations of the DES and low-z SN samples and propagate the effects of the same systematics considered for the data (simulations are generated following our nominal modeling approach, i.e., P21 ($M_\star$)). We perform a full analysis on each simulated data sample and in Table 8, we report the mean $\Delta\chi^2_S$ and its standard deviation (highlighted in bold), BBC-fitted nuisance parameters, and best-fit w determined from the simulations. Standard deviations measured from the 25 independent realizations of our SN sample provide a robust estimate of uncertainties in the observed shifts.

Our evaluations of $\Delta\chi^2_S$ (and relative uncertainties) enable us to quantify which models or analysis variants are favored by the data: $\chi^2_S > \chi^2_{\rm Nominal}$ or $\Delta\chi^2_S > 0$ suggest that the analysis variant introduced with the systematic S provides a worse modeling of the Hubble residuals, while $\Delta\chi^2_S < 0$ suggests a better modeling of Hubble residuals intrinsic scatter. The more a systematic is favored by the data (i.e., more negative $\Delta\chi^2_S$), the larger its impact on the final error budget will be (especially if it significantly changes the best-fit cosmology). For these reasons, the results presented in Table 8 and Figure 13 are useful to interpret the error budget presented in Table 7 and Figure 12. In the next sections, we discuss in detail each subgroup of systematic uncertainties considered in the error budget.

7.1.1. Systematic: SN Ia Intrinsic Scatter

Uncertainties related to the modeling of SN intrinsic scatter are the largest source of systematic uncertainties in our analysis. The different intrinsic scatter models included in our analysis are all equally favored by the data, i.e., have χ^2_S comparable to the nominal analysis. This is shown in Table 8; the three P21 realizations (P21 population 1, P21 population 2, and P21 population 3), the BS21 model, the dust model implemented splitting the $u-r$ color (P21($u-r$)) and the SN age model by W22 all have $\Delta\chi^2_S$ between 0 and +15. These fluctuations are expected considering the $\Delta\chi^2_S$ scatter over 25 simulations; see the bottom section of Table 8. The only systematic that is significantly favored by the data is α evolution ($\Delta\chi^2_S = -7 \pm 2$). We discuss the effects of nuisance parameter redshift evolution in Section 7.1.3. For comparison, we present in Table 8 the nuisance parameters and $\Delta\chi^2_S$ for the intrinsic scatter model by Guy et al. (2010) and Kessler et al. (2013; historically referred to as the "G10" model). This model is significantly disfavored by the data ($\Delta\chi^2_S = 102$) and therefore is not included in our systematic error budget. We make a similar test and reach the same conclusion when testing the model by Chotard et al. (2011; historically referred to as the "C11" model).

While the different intrinsic scatter models considered in Table 8 are almost equally favored by the data, the relative best-fit cosmologies can change. In Figure 14, we present the shifts in the best-fit w determined for each variant in Table 8. We present our results both for the data and for the simulations. The size and direction of the w shifts observed in the data are well reproduced by the simulations (which are all generated assuming our nominal intrinsic scatter model, i.e., P21($M_\star$)).

7.1.2. Systematic: Residual Mass/Color Steps

The most interesting result in Table 8 is related to the recovered nuisance parameters. As already noted in Section 5.2, the recovered mass step for the data is $\gamma_{M_\star} \sim 0.039$ mag (5σ significance) across most analysis variants, regardless of what SN subsample is considered (DES-SN only or DES-SN + Low-z; see Table 5). When fixing $\beta = 3.14$ and considering either only blue SNe Ia ($c < 0$) or only reddened SNe Ia ($c > 0$), we find $\gamma_{M_\star} = 0.039 \pm 0.012$ and $\gamma_{M_\star} = 0.028 \pm 0.014$, which shows no significant color dependence.

The mass step decreases when considering the BS21 model ($\gamma_{M_\star} = 0.026$) and the modeling approach by W22 ($\gamma_{M_\star} = 0.017$). In the first case, BS21 assumed the intrinsic color distribution to be bluer ($c_{\rm int} = -0.084$) and dust to have longer tails (i.e., larger τ_E values) compared to our nominal analysis model. In the second case, the reduced mass step is likely a consequence of the fact that the SN age modeling approach by W22 does not reproduce x_1–$M_\star$ correlations well (see Figure 10). In the BBC approach, if the simulations used to determine bias corrections underestimate x_1–$M_\star$ correlations, the mass step can be partially absorbed into the bias correction term $\mu_{\rm bias}$ (this has been discussed and demonstrated by Smith et al. 2020b; Popovic et al. 2021). As a test, we generate a large simulation following the same approach used for our nominal simulations, but removing x_1–$M_\star$ correlations. We use this alternative simulation to bias-correct our SN sample and measure the mass step. We recover a mass step of 0.019 mag, half of the mass step found in the nominal analysis (see Table 8).

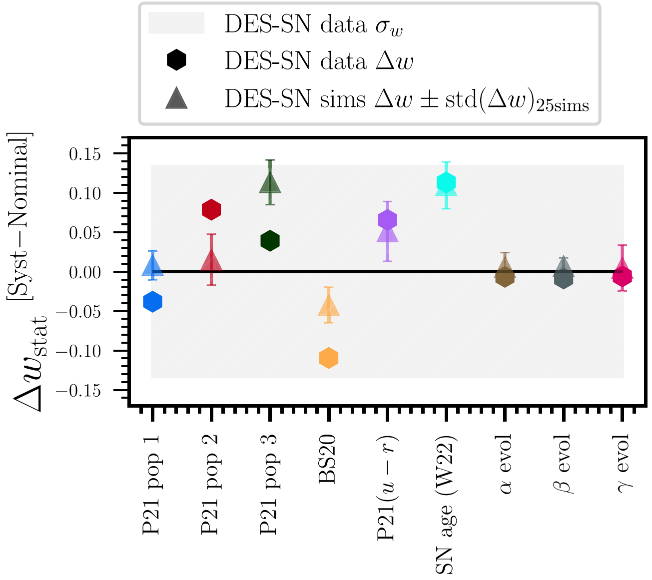


Figure 14. Shifts in w (Δw) for different intrinsic scatter models and α, β, γ evolution (see Section 6.2.1). Δw measured from the data (hexagons) is compared to Δw measured from averaging 25 DES+Low-z simulations (triangles). For simulations, each error bar is the standard deviation of the Δw from the 25 simulated samples. Shifts are measured with respect to the nominal intrinsic scatter model, i.e., P21 measuring the mass step at $10^{10} M_\odot$ (first data point on the left). The statistical uncertainty in w from the DES data is represented by the gray shaded area.

The color step measured for the P21($u-r$) model is also nonnegligible ($\gamma_{u-r} = 0.033 \pm 0.007$ mag, $\sim 5\sigma$ significance) but slightly smaller than the mass step, which suggests that $u-r$ color is a better proxy to describe SN–host correlations (even though P21($u-r$) model is not significantly favored by the data).

In simulations, we recover a mass step consistent with zero (average uncertainty on $\gamma_{M_\star}$ is less than 0.01) and lower values of fitted β. While the zero mass step is expected (by construction, simulations have only a dust-based mass step), it is interesting to note that the "effective" β measured from the simulations and the data are different. This is unexpected since simulations are generated using a dust-based model specifically built to reproduce the intrinsic ($\beta_{\rm int}$) and extrinsic (R_V) color–luminosity corrections. The 4σ discrepancies in the β and γ recovered in data and simulations suggest that either (i) dust parameter errors from the "Dust2Dust" fit are larger than the statistical uncertainties because of our incomplete understanding of SN Ia intrinsic properties and correlations, or (ii) the mass step is not fully explained by dust.

Despite the discrepancies observed between data and simulations and our likely still incomplete modeling of SNe Ia intrinsic scatter, our overall modeling of bias corrections is significantly improved over previous models (e.g., "G10" from Kessler et al. 2013). The associated systematics from our current discrepancies are discussed in Section 7.3.

7.1.3. Systematic: Nuisance Parameter Evolution

When considering the combined DES and low-z samples, we detect a significant (>2σ) redshift evolution in the α and β

Table 9
Nuisance Parameter Evolution

	α_0	α_1
DES-SN+low-z	0.146(6)	0.033(14)
DES-SN only	0.176(10)	−0.013(20)
DES-SN+low-z (25 sims[a])	0.147(5)	0.003(12)
	β_0	β_1
DES-SN+low-z	3.05(0)	0.17(6)
DES-SN only	3.13(0)	0.04(6)
DES+low-z (25 sims[a])	2.88(8)	−0.03(14)
	γ_0	γ_1
DES-SN+low-z	0.051(15)	−0.033(34)
DES-SN only	0.068(24)	−0.045(49)
DES+low-z (25 sims[a])	−0.006(12)	0.010(34)

Note.
[a] Reported uncertainties are estimated as the standard deviations from 25 simulations.

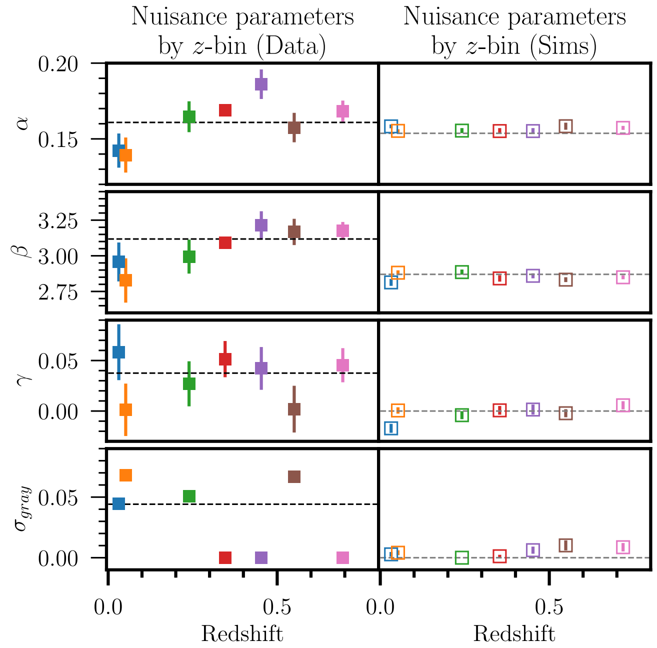


Figure 15. Values of α, β, γ, and $\sigma_{\rm gray}$ for discrete redshift bins measured from the DES-SN5YR sample (left) and from the 25 simulated DES-SN5YR-like samples (right). The low-z sample is divided into two redshift bins ($0.025 < z < 0.04$ and $0.04 < z < 0.1$, ∼80 SNe in each bin), while the DES-SN sample is divided into equally spaced bins of $\Delta z = 0.1$. On the left, the baseline fits measured from the data (assuming no redshift evolution of nuisance parameters) are shown in black for each panel. On the right, we show the true simulated values in gray. For simulations, input values are fully recovered across redshift bins. Fitted α, β, γ evolution are presented in Table 9.

nuisance parameters. The best-fit parameters for the redshift evolution of α, β, and γ are summarized in Table 9. The fitted parameters suggest that α and β increase with redshift ($\alpha < 0.15$ and $\beta < 3.05$ at lower redshifts and $\alpha > 0.16$ and $\beta > 3.2$ at higher redshifts). The effects of redshift evolving α and β on SN distances are presented in Figure 13. For both the α and β evolution, there is no evidence of a net trend in the binned SN distance differences and, as a result, the differences between the best-fit w values with and without α and β evolution are consistent with zero (see "α evolution" and "β evolution" in Table 8 and Figure 14).

While we do not detect significant γ evolution in our sample, we conservatively allow for the possibility of a redshift evolution in γ and include this as a source of systematic uncertainty. The contribution to the systematic error budget from nuisance parameter evolution (α, β, and γ) is presented in Figure 12 and Table 7.

When measuring α, β, and γ evolution on the 25 simulations of DES-SN+Low-z, we successfully recovered zero evolution for all nuisance parameters (see Table 9), which indicates that evolution is not an artifact of our analysis.

Finally, we note that the evolutions of α and β are negligible when considering the DES-SN sample only. This is particularly clear when looking at the values of α, β, and γ evaluated in discrete redshift bins, as shown in Figure 15. The parameters α, β, and γ are consistent between redshift 0.2 and 1 (the redshift range covered by the DES-SN sample) and present more significant discrepancies at low redshift. This was also highlighted in Section 5.2, where we presented the best-fit nuisance parameters for DES alone and low-z alone and noted significant α discrepancies. We perform a similar test on simulations and present our results on the right-hand panels in Figure 15. For simulations, we demonstrate that the true α, β, and γ are well recovered across different redshift bins, demonstrating the shifts observed in the data are likely not introduced by the BBC fitting approach.

If the underlying cause of the observed evolution and nuisance parameters discrepancies between DES and low-z (Table 5) had an astrophysical origin (e.g., SN progenitor physics), it would be challenging to explain how this could produce such large fluctuations in nuisance parameters at $z \sim 0.1$. A likely explanation of the observed evolution is an incomplete understanding of the properties and selection effects of the low-z SN samples. The modeling of the low-z samples is challenging for two reasons. First, the selection functions of the low-z SN samples are poorly understood. Second, the small statistics in the low-z samples make (i) dust modeling more uncertain, and (ii) empirical modeling of x_1–$M_\star$ or x_1–$(u-r)$ correlations more difficult to model (see Figures 18 and 19).

7.1.4. Systematic: Photometry and Calibration

Historically, calibration uncertainties—especially between low- and high-redshift samples—have been the largest source of systematic uncertainties in SN analyses. In our analysis, systematic uncertainty associated with calibration is smaller than systematic uncertainties related to SN Ia intrinsic properties.

The reduced impact of calibration on our systematic error budget is due to various factors:

1. The FGCM method presented by Burke et al. (2018) improves the accuracy of the DES internal calibration to ∼5 mmag in the *griz* bands.
2. In our analysis, we combine samples from different surveys, with different filter systems and calibrations. The work presented by Scolnic et al. (2015) and extended by Brout et al. (2022b) significantly improved cross-calibration between the different SN surveys used in this work. In particular, it improved the calibration of older low-z SN samples.
3. Calibration affects our ability to train light-curve fitting models like SALT3. For this analysis, we train the

SALT3 model on a larger and more accurate training sample compared to the Betoule et al. (2014) SALT2 model of previous SN cosmology analyses, which makes our SALT3 model less sensitive to calibration uncertainties. As shown in Taylor et al. (2023), the scatter between SALT3 surfaces considered for systematic uncertainties is significantly reduced compared to previous SALT2 models.

4. The approach introduced by Brout et al. (2022b) to propagate calibration uncertainties to cosmology has significantly reduced the impact of this source of systematics. Following Brout et al. (2022b), we propagate calibration uncertainties *simultaneously* on the SALT3 training and on the SN light curves, thus accounting for calibration correlations for data used in both the training and cosmology analysis.[74]

7.1.5. Systematic: Non-Ia Contamination

One of the most remarkable results of our analysis is that systematic uncertainties associated with contamination are only 9% of the total systematic error budget and they bias SN distances by only a few millimagnitudes (see Figure 13, lower panel).

As discussed by Vincenzi et al. (2021b) and highlighted in Table 4, even before photometric classifiers are applied, contamination is reduced to 9%. This reduction is due to the fact that the SALT3 fitting and the BBC fitting serve as classifiers themselves:

1. The SALT3 model fitting (and associated quality cuts) already eliminate a significant fraction of non-Ia SNe (from ∼40% to ∼11%; see Table 4).
2. The requirement of a valid bias correction further adds the constraint that the fitted SALT3 parameters of each SN are representative of what is found in the three-dimensional parameter space populated by large SN Ia simulations used for bias corrections.
3. Chauvenet's criterion iteratively applied in BBC provides a cosmological model independent outlier rejection.

Classifiers contribute to further reducing the weight of contaminants in the cosmological fit. When testing classifiers on simulations, algorithms like SuperNNova and SCONE can achieve levels of purity and efficiency >98% (see results reported in Möller & de Boissière 2020; Qu et al. 2021; Vincenzi et al. 2021b). In Figure 16, we compare the classification of DES data using the different algorithms tested in our analysis, and we note some discrepancies. The three classification algorithms (SuperNNova, SCONE, and SNIRF, all trained on V19) have consistent classification on 90.0% of DES-SNe (84.9% are classified as Ia by all three classifiers and 5.1% are classified as non-Ia by all three classifiers), while the three SuperNNova models trained on the V19, J17 and DES-non-Ia templates have consistent classification for 94.2% of SNe (89.8% Ia and 4.3% non-Ia). In particular, we note that the baseline classifier SuperNNova (trained on V19) and SCONE are the more "conservative" classifiers (a smaller number of SNe are classified as SNe Ia). Despite these small discrepancies, we show that the cosmological results obtained using different classification methods are fully consistent with our statistical uncertainties (see the bottom panel in Figure 16).

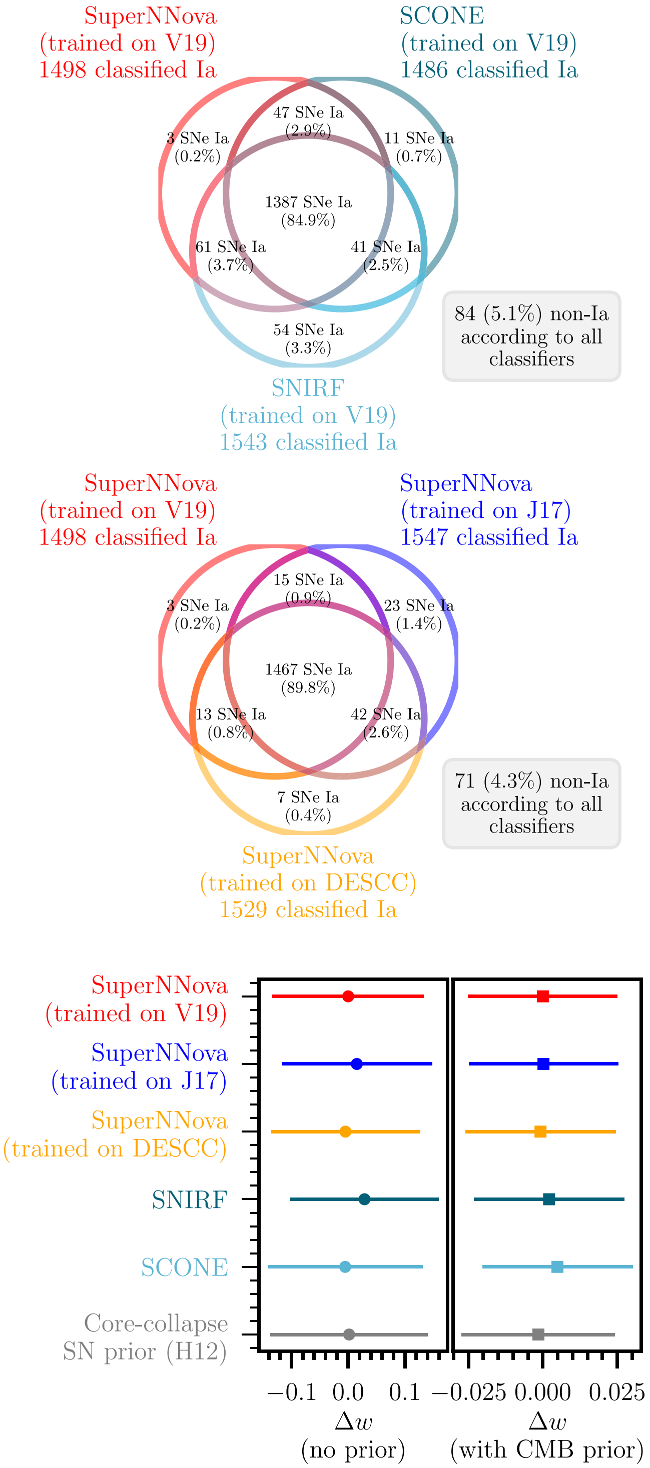


Figure 16. Likely DES-SNe Ia (i.e., $P_{\rm Ia} > 0.5$) according to the different classification models used in our analysis. Upper Venn diagram: we compare three classification tools: SuperNNova (red), SCONE (dark teal), and SNIRF (teal). Lower Venn diagram: We compare three different SuperNNova models, one trained using core-collapse templates by Vincenzi et al. (2019; V19, red), one trained using core-collapse templates by Jones et al. (2017; J17, blue), and one trained using core-collapse templates from DES data (DES-CC, orange). Lower plot: Differences in w compare for different classification choices. Differences are measured with respect to the nominal ("SuperNNova (trained on V19)") both when including a CMB-like prior (right) and no prior (left). The uncertainties on w are statistical only.

---

[74] This approach was implemented in Joint-Lightcurve-Analysis (Betoule et al. 2014), but it was not applied in subsequent SN cosmological analyses due to the unavailability of the SALT2 training code.

Table 10
Fraction of Contaminants Likelihood to Total Likelihood

	$\sum\mathcal{L}^i_{\mathrm{CC}}/\sum\mathcal{L}^i_{\mathrm{tot}}$ [a]		Δw_{stat} [b]	
Classification Systematic	Data	25 sims[c]	Data	25 sims[c]
Nominal (SuperNNova trained on Vincenzi et al. 2019)	0.065	0.066 ± 0.008	0.000	0.000 ± 0.024
SuperNNova trained on J17 (Jones et al. 2017)	0.074	0.067 ± 0.009	−0.005	0.016 ± 0.024
SuperNNova trained on DES-CC (R. Hounsell et al. 2023, in preparation)	0.069	0.066 ± 0.009	0.015	0.013 ± 0.024
SNIRF classifier	0.059	0.063 ± 0.009	0.029	0.037 ± 0.022
Replace sim core-collapse SN prior with fitted polynomial prior[d]	0.069	0.068 ± 0.011	0.002	0.003 ± 0.025

Notes.

[a] $\sum\mathcal{L}^i_{\mathrm{tot}}$ is the total likelihood, $\sum(\mathcal{L}^i_{\mathrm{Ia}}+\mathcal{L}^i_{\mathrm{CC}})$ described in Equation (3). See Equations (4) and (5) in Section 3.3 for the definition of the SN Ia and contaminants likelihood terms.

[b] Shifts in w when considering different classification methods (no systematic covariance matrix is used to measure these shifts). w is constrained using SN-only data (no additional CMB prior).

[c] Mean and standard deviation of the fraction of contaminants to total likelihood *measured* over 25 simulations. The *true* fraction of contaminants to total likelihood is 0.053% ± 0.004%.

[d] See Hlozek et al. (2012).

When testing the classifiers on the sample of 207 spectroscopically classified SNe Ia from the DES-SN3YR analysis (Abbott et al. 2019; Smith et al. 2020a) and on the sample of 43 spectroscopically classified non-Ia from the first three years of the DES survey (see DES-non-Ia, we find that SuperNNova trained on V19 templates has the highest accuracy (correctly classifying as Ia 203 out of 207 SNe Ia, and misidentify only 2 out of 43 non-Ia), SuperNNova trained on J17 has the lowest purity (6 out of 43 non-Ia are misclassified as SNe Ia), and SNIRF has the lowest efficiency (192 out of 207 SNe classified as Ia). These results confirm that our baseline classifier (SuperNNova) provides excellent results not only on simulations but also on data.

In general, it is not surprising that classification accuracy is lower on real data rather than on simulations, as data often present defects and outliers that are not fully modeled in simulations (for the DES data, these artifacts are significantly reduced by SMP, and this contributes to improving classification; see Sanchez et al. (2024). Despite the differences between the classification algorithms used in our analysis and despite the additional challenges that classification on real data presents, we find that SN distances measured from using different classification algorithms vary by 10 mmag at most (see Figure 13) and differences in the estimated w are not significant (see Table 10).

In order to further validate that contamination and classification are consistent between simulations and data, we consider an additional metric. The true number of contaminants in the DES data is unknown; however, the BEAMS contaminants likelihoods, $\sum\mathcal{L}^i_{\mathrm{CC}}$(see Equation (4)) provides an estimate of how many SNe are associated with the contaminants population. We determine the sample contamination as the fraction of the contaminants' total likelihood divided by the total sample likelihood $\sum(\log(\mathcal{L}^i_{\mathrm{Ia}}+\mathcal{L}^i_{\mathrm{CC}}))$. This fraction provides a measurement of the fraction of contaminants estimated in the Hubble diagram during the BBC cosmological fit. This quantity is also defined in Kessler et al. (2023) as the sum of the "BEAMS probabilities" (see Equation (9) in Kessler et al. 2023).

In Table 10, we present the contaminants' likelihood fractions estimated on data and simulations when implementing different classifiers. For every classification method, the contamination predicted in the simulations is consistent with the contamination observed in the data, and it is ∼6%–7%. This is the first SN analysis that demonstrates such a close agreement between contamination estimated in the data and in simulations. The *true* percentage of contaminants to total likelihood estimated from the simulations is 5.3 ± 0.4 and the simulated value of $w=-1$ is fully recovered by our pipeline.

7.2. Validation of the BBC Fitting Approach and Final Cosmological Contours

In SN cosmological analyses, one of the most critical aspects is to model and correct for sample selection biases. For the majority of cases, analytical modeling of selection effects in SN experiments is an intractable problem. For this reason, we have to rely on complex simulations like the ones described in Section 4. An important limiting factor of using simulations is that they require assumptions of the input cosmology (as it is computationally prohibitive to generate a new simulation for every step of the cosmological fit). For this reason, it is important to (i) quantify the size of the biases on w when assuming the wrong input cosmology in the bias correction simulations, and (ii) validate the Bayesian cosmological contours determined from SN distances bias-corrected assuming a specific input cosmology.

The first aspect has been discussed in Section 6.1 in Kessler & Scolnic (2017) and further tested by Camilleri et al. (2024). Kessler & Scolnic (2017) use a set of simulated SN samples and demonstrate that the effects on w of assuming the wrong input cosmology in the bias correction simulations is small (one-seventh of the statistical uncertainties) when including a strong Ω_M prior.

Camilleri et al. (2024) reproduced an analogous test without including the Ω_M prior and showed that the direction of the bias is always *along* the SN contour degeneracy (hence the small biases when including a Ω_M prior, which is orthogonal to the SN contour degeneracy).

The validation of the Bayesian cosmological contours is addressed by Armstrong et al. (2023). Armstrong et al. (2023) validate the Bayesian cosmological contours produced using the BBC Hubble diagram by generating 150 realizations of the DES-SN sample and making use of approximate Neyman confidence intervals. This work demonstrates that the size of the cosmological contours produced with our pipeline and the Neyman confidence intervals agree at the >95% level.

7.3. Analysis Variants and Their Impact on w

In this section, we present different analysis variants and their impact on w. These analysis variants aim to answer different questions: (a) Do our cosmological results vary significantly when considering different groups of DES fields and measuring cosmology along different directions in the sky? (b) Do our cosmological results vary significantly when considering subsamples of SNe found in specific host environments? (c) Do our cosmological results vary significantly when considering different SALT3 wavelength ranges? (d) What is the impact on cosmological results of incorrect implementation of, e.g., MW corrections? The analysis variants implemented to answer (a), (b), and (c) are not included in our systematic error budget because they effectively select only a specific subsample of the data set. The analysis variants implemented to answer (d) are also not included in the main analysis because they are incorrect implementations of the cosmological analysis.

The tested analysis variants and the associated w shifts are shown in Table 11. We present not only the shifts in w estimated from DES-SNe (no external priors) but also the standard deviation of the w shifts measured from 25 simulations (for a robust estimation of the significance of the measured w shifts). When considering different DES-SN subfields (a) or host-dependent subsamples (b), we do not see significant (>2.5σ) deviation from the baseline w. When we incorrectly implement the cosmological analysis (d), we note that not including bias corrections (i.e., using the BBC0D approach), ignoring Milky Way extinction corrections or modeling bias corrections for deep and shallow fields together would have potentially produced the most significant biases on w.

Table 11

Miscellaneous w Shifts (No CMB Prior Included) When Implementing Analysis Variants Described in Section 7.3 and Not Included in the Analysis Error Budget

Analysis variant	$\Delta w_{\rm stat}$ [a]	std (Δw) [25 sims]
(a) Different sets of DES-SN fields		
DES-SN(C3)+Low-z	−0.249	0.126
DES-SN(X3)+Low-z	−0.196	0.180
DES-SN(X1,X2)+Low-z	−0.008	0.207
DES-SN(C1,C2)+Low-z	−0.535	0.239
DES-SN(S1,S2)+Low-z	0.023	0.218
DES-SN(E1,E2)+Low-z	0.016	0.219
DES-SN(Deep)+Low-z	−0.166	0.098
DES-SN(Shallow)+Low-z	−0.069	0.135
(b) Host prop subsamples		
Only SNe in $M_\star > 10^{10} M_\odot$	−0.071	0.081
Only SNe in $M_\star < 10^{10} M_\odot$	0.138	0.213
Only SNe in $u - r > 1$ hosts	−0.000	0.041
Only SNe in $u - r < 1$ hosts	−0.220	0.149
(c) SALT3 wavelength coverage		
SALT3 using 3500–7000 Å [b]	−0.027	0.024
SALT3 using 4000–8000 Å [b]	−0.105	0.128
(d) Incorrect implementation		
No BBC[c]	−1.024	0.116
No BEAMS, only $P_{\rm Ia} > 0.5$ cut[d]	−0.033	0.034
No Milky Way Ext corrections	0.185	0.035
Biascor Deep/Shall together	−0.136	0.040
Ignore SB anomaly in data	−0.012	0.000
Force $\gamma = 0$ in BBC fit	−0.113	0.007

Notes.

[a] Shift in w when implementing each analysis variance (no CMB prior included). w is measured using SN only (DES and low-z) and without including any systematic covariance matrix (statistical uncertainties only).

[b] In the nominal analysis, the wavelength range used for the SALT3 model is 3500–8000 Å.

[c] No BEAMS and no bias corrections.

[d] We apply a $P_{\rm Ia}$-based cut and assume every SN is a type Ia. We apply bias corrections but we do not implement the full BEAMS approach and we do not incorporate probabilities in the BEAMS framework.

7.4. Systematic Budget on $w_0 w_a$ *Figure of Merit*

The unbinned Hubble diagram and $\mathcal{C}$ can also be used to determine constraints on cosmological parameters w_0 and Ω_M using a flat $w_0 w_a$ CDM model. The evaluation of present and forthcoming dark energy experiments involves assessing their capacity to enhance the Dark Energy Task Force Figure of Merit (Albrecht et al., 2006), which is determined as the reciprocal of the region enclosed within the $w_0 w_a$ contours. In this section, we present the $w_0 w_a$ constraining power of DES-SN5YR, before and after including systematic uncertainties. In Figure 17, we present the $w_0 w_a$ figure of merit (FoM) with and without including systematic uncertainties (similar to the systematic σ_w budget presented in Figure 12). We estimate the FoM including DES-SN5YR and CMB-like prior by Planck Collaboration (2020). We find $w_0^{\rm stat} = 0.098$ and $w_a^{\rm stat} = 0.49$ and $w_0^{\rm stat+syst} = 0.12$ and $w_a^{\rm stat+syst} = 0.59$. This corresponds to a ~35% decrease in FoM (from 83.5 when including statistical uncertainties only to 54 when including statistical and systematic uncertainties) and we find that calibration, light-curve modeling, and modeling of dust properties are the dominating sources of systematics.

8. Conclusions and Future

We present the DES-SN5YR cosmological analysis using the 5 yr photometric SN sample for the DES-SN program. This new independent sample constitutes the largest and deepest SN sample from a single telescope to date and provides constraints on the dark energy EoS competitive with the best existing compilation of all high-z SNe Ia previously discovered (Brout et al. 2022a). Additional DES-SN cosmological analyses using photo-z information only (Chen et al. 2024) and using additional DES-SN spectroscopic data are currently ongoing.

In the following paragraphs, we summarize the main conclusions of the analysis presented in this paper and outline key areas for future research, particularly in light of upcoming SN Ia surveys such as LSST and Roman.

Historically, the limiting systematic for SN analyses has been photometric calibration. We are entering a new phase in SN cosmology where intrinsic SN properties are becoming the dominating source of systematic (a third of the systematic w uncertainty budget, and half of the $w_0 w_a$ FoM budget). For the future of SN cosmology, it will be crucial to significantly improve our understanding of SN intrinsic scatter and correlations with hosts. Alternatively, our way of doing SN Ia cosmology will likely be revised, for example, by selecting only specific subsamples of SNe Ia (e.g., only blue SNe Ia in

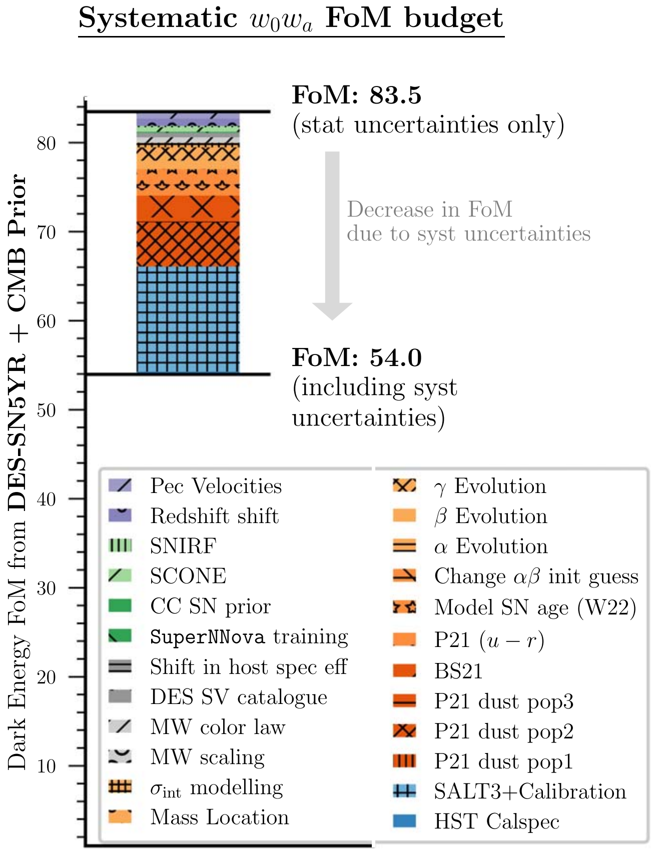


Figure 17. Decrease in the DES-SN5YR FoM when including systematic uncertainties. Similarly to the systematic error budget presented in Figure 12, we highlight the sources of systematic uncertainties that degrade the FoM the most.

low-mass or blue environments as suggested by Kelsey et al. 2023).

For the modeling of SN Ia intrinsic scatter, the dust-driven model proposed by BS21 is the most successful currently available; however, it requires further improvements. This is highlighted by the several discrepancies observed between the simulations and the data. In our analysis, we build simulations using “Dust2Dust” (P21 dust), a software designed to infer the dust properties of any SN Ia sample, operating under the assumption that the BS21 model accurately describes SN dust and its correlations with SN host properties. When comparing simulations and data, we find interesting discrepancies in the inferred nuisance parameters (especially β and γ). In particular, we find a nonnegligible residual mass step, which might indicate the need to include additional astrophysics in the BS21 model.

Moreover, contamination has been one of the most important challenges for the DES-SN5YR analysis compared to the DES-SN3YR analysis. We demonstrated that contamination from non-Ia SNe is not the dominating source of systematic uncertainties for photometric SN samples. Our analysis is also the first SN cosmological analysis to present simulations that can accurately model and reproduce the amount of core-collapse contamination observed in the data (Table 10). Our results mark a transition point for future SN experiments such as LSST, for which spectroscopic classification of SNe will be extremely limited (see C. Frohmaier et al. 2024, in preparation).

Finally, in the next era of high-redshift SN experiments (e.g., Rubin Observatory’s Legacy Survey of Space and Time, and Nancy Grace Roman Space Telescope), it will be essential to obtain low-z SN samples with well-defined selection functions and accurate calibration. Experiments like the Young Supernova Experiment (Jones et al. 2021; Aleo et al. 2023) and Zwicky Transient Facility (Dhawan et al. 2022; Rigault et al. 2024; M. Smith et al. 2024, in preparation) and DEBASS (PI: Dillon Brout) will provide the next generation of low-z samples and significantly improve cosmological constraints from SNe Ia.

Acknowledgments

Funding for the DES Projects has been provided by the U.S. Department of Energy, the U.S. National Science Foundation, the Ministry of Science and Education of Spain, the Science and Technology Facilities Council of the United Kingdom, the Higher Education Funding Council for England, the National Center for Supercomputing Applications at the University of Illinois at Urbana-Champaign, the Kavli Institute of Cosmological Physics at the University of Chicago, the Center for Cosmology and Astro-Particle Physics at the Ohio State University, the Mitchell Institute for Fundamental Physics and Astronomy at Texas A&M University, Financiadora de Estudos e Projetos, Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro, Conselho Nacional de Desenvolvimento Científico e Tecnológico and the Ministério da Ciência, Tecnologia e Inovação, the Deutsche Forschungsgemeinschaft and the Collaborating Institutions in the Dark Energy Survey.

M.V. was partly supported by NASA through the NASA Hubble Fellowship grant HST-HF2-51546.001-A awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Incorporated, under NASA contract NAS5-26555. R.K. is supported by DOE grant DE-SC0009924. T.M.D. is the recipient of an Australian Research Council Laureate Fellowship (FL180100168) funded by the Australian Government. L. K. thanks the UKRI Future Leaders Fellowship for support through the grant MR/T01881X/1. A.M. is supported by the ARC Discovery Early Career Researcher Award (DECRA) project No. DE230100055. L.G. acknowledges financial support from the Spanish Ministerio de Ciencia e Innovación (MCIN), the Agencia Estatal de Investigación (AEI) 10.13039/501100011033, and the European Social Fund (ESF) ”Investing in your future” under the 2019 Ramón y Cajal program RYC2019-027683-I and the PID2020-115253GA-I00 HOSTFLOWS project, from Centro Superior de Investigaciones Científicas (CSIC) under the PIE project 20215AT016, and the program Unidad de Excelencia María de Maeztu CEX2020-001058-M, and from the Departament de Recerca i Universitats de la Generalitat de Catalunya through the 2021-SGR-01270 grant. We acknowledge the University of Chicago’s Research Computing Center for their support of this work.

The Collaborating Institutions are Argonne National Laboratory, the University of California at Santa Cruz, the University of Cambridge, Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas-Madrid, the University of Chicago, University College London, the DES-Brazil Consortium, the University of Edinburgh, the Eidgenössische Technische Hochschule (ETH) Zürich, Fermi National Accelerator Laboratory, the University of Illinois at Urbana-

Champaign, the Institut de Ciències de l'Espai (IEEC/CSIC), the Institut de Física d'Altes Energies, Lawrence Berkeley National Laboratory, the Ludwig-Maximilians Universität München and the associated Excellence Cluster Universe, the University of Michigan, NSF's NOIRLab, the University of Nottingham, The Ohio State University, the University of Pennsylvania, the University of Portsmouth, SLAC National Accelerator Laboratory, Stanford University, the University of Sussex, Texas A&M University, and the OzDES Membership Consortium.

Based in part on observations at Cerro Tololo Inter-American Observatory at NSF's NOIRLab (NOIRLab Prop. ID 2012B-0001; PI: J. Frieman), which is managed by the Association of Universities for Research in Astronomy (AURA) under a cooperative agreement with the National Science Foundation. Based in part on data acquired at the Anglo-Australian Telescope. We acknowledge the traditional custodians of the land on which the AAT stands, the Gamilaraay people, and pay our respects to elders past and present.

The DES data management system is supported by the National Science Foundation under grant Nos. AST-1138766 and AST-1536171. The DES participants from Spanish institutions are partially supported by MICINN under grants ESP2017-89838, PGC2018-094773, PGC2018-102021, SEV-2016-0588, SEV-2016-0597, and MDM-2015-0509, some of which include ERDF funds from the European Union. IFAE is partially funded by the CERCA program of the Generalitat de Catalunya. The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Program (FP7/2007-2013) including ERC grant agreements 240672, 291329, and 306478. We acknowledge support from the Brazilian Instituto Nacional de Ciência e Tecnologia (INCT) do e-Universo (CNPq grant 465376/2014-2).

This manuscript has been authored by Fermi Research Alliance, LLC under Contract No. DE-AC02-07CH11359 with the U.S. Department of Energy, Office of Science, Office of High Energy Physics.

Finally, the authors thank the anonymous referees for the helpful comments that improved the quality of the manuscript.

Appendix A
Modeling the External Low-z Samples and DES-SN Spectroscopic Sample

In this Appendix, we focus on the modeling of low-z external samples (CfA, CSP, and Foundation SN surveys). In Figure 18, we present a comparison between observed and simulated distributions of various SN and host galaxy properties (e.g., SN stretch, SN color, SN host stellar masses) and correlations between x_1 and host stellar mass and host rest-frame $u-r$ color.

Dust properties and correlations between x_1 and host properties are modeled following the same techniques used for the DES-SN sample. In Figure 19, we present a comparison between observed and simulated Hubble residuals measured using the BBC0D approach (same as Figure 5, but for external low-z samples). The dust models implemented in our analysis generally reproduce the observed trends; however, the low statistics make it challenging to infer SN dust properties in these samples. In Figure 20, we present the best fit values of the dust parameters determined from the dust-fitting code "Dust2-Dust" and three realizations included in the systematic uncertainties (see Section 6.2.1). The dust parameters presented in Figure 20 are used to build the dust models illustrated both in Figure 5 and Figure 19.

Moreover, we note that the selection functions of the low-z SN samples are not as well understood as those of the DES-SN sample. Early SN surveys such as CfA were designed as targeted surveys, specifically focused on the discovery and follow-up of the brightest SNe Ia in high-mass nearby galaxies. Therefore, the selection function in the low-z samples can only be inferred by fudging simulations to match the data.

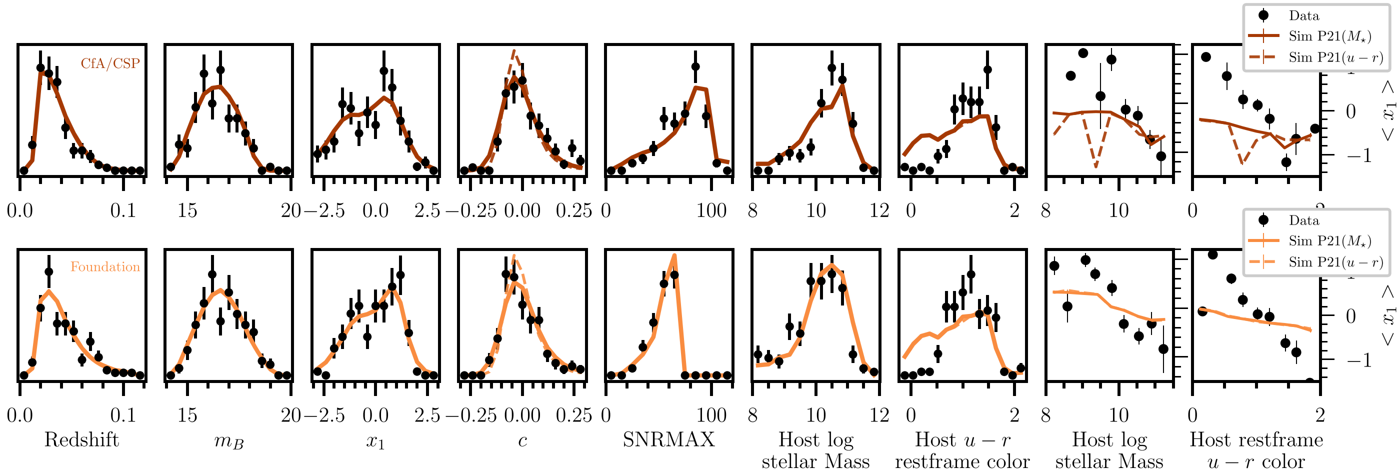


Figure 18. Same as Figure 3 but for the low-redshift samples included in this analysis.

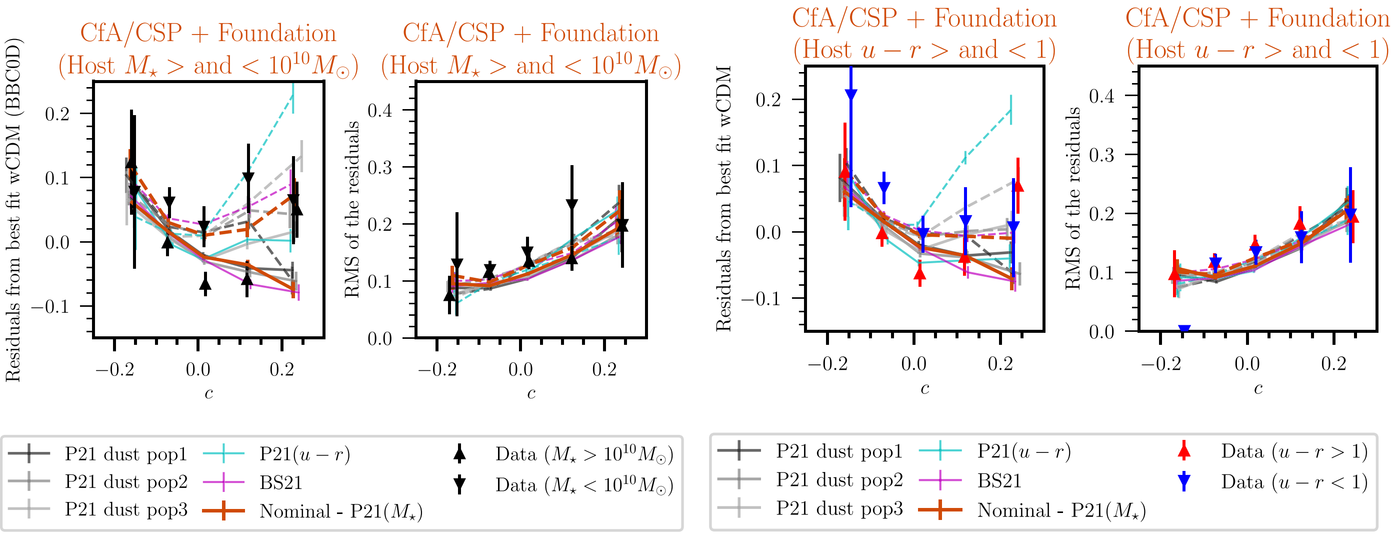


Figure 19. Same as Figure 5, but for the external low-z samples (all combined).

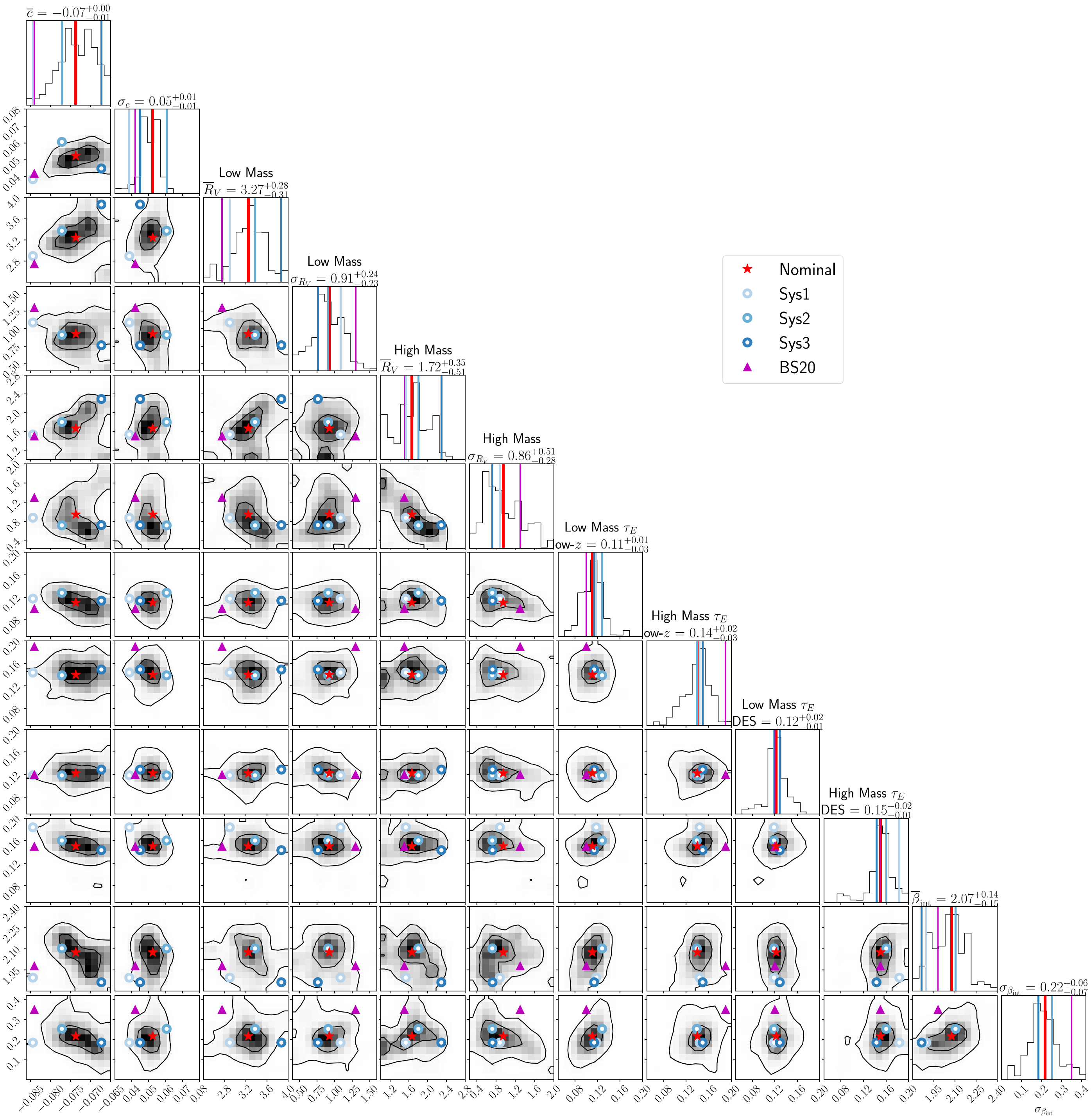


Figure 20. Best fit from the dust-fitting code (Popovic et al. 2021a) and three realizations included in the systematic uncertainties (see Section 6.2.1).

Appendix B
Selecting Only SNe Detected before SN Peak Brightness

The lack of pre-peak SN observations can significantly impact the accuracy of SN stretch and SN color estimates, and therefore estimates of SN distances. Conversely, selecting only SNe detected before peak brightness can significantly reduce the size of an SN sample, especially for the oldest low-z targeted surveys. For the DES-SN sample, we have that only 1599 out of 1635 (2%) have the first detection after −2 days from peak brightness. In the low-z sample, the fraction increases to 80 over 194 (41%).

In order to test the effects of this cut, we run our cosmological analysis selecting only SNe with pre-peak data (at least a detection before −2 days from peak brightness) and compare the results with the nominal analysis (that only requires at least a detection before +2 days from peak brightness). We find that the fitted nuisance parameters are consistent within uncertainties (see Table 12) and that the cosmological parameter w shifts by 0.013. To test the significance of this w shift, we perform the same test on a set of 25 DES-SN5YR simulated samples. We find an average shift in w of 0.013, with a standard deviation of 0.051. Therefore, the shift observed in the data is not significant.

Table 12
Nuisance Parameters and Cosmological Fit When Combining DES with Different Low-z External Samples and When Using DES Alone

Sample	$N_{\rm SNe}$	α	β	γ	$\sigma_{\rm gray}$	rms[a]	$\Delta w_{\rm stat}$
DES-SN + low-z	1829	0.161(1)	3.12(3)	0.038(7)	0.04	0.168	0.000 ± 0.133
DES-SN + low-z (detection before −2 days from peak)	1713	0.161(3)	3.13(3)	0.036(8)	0.04	0.17	0.013 ± 0.139

Note.
[a] The rms error is measured by applying a cut of $P_{\rm Ia} > 0.5$ on the DES-SN sample.

ORCID iDs

M. Vincenzi https://orcid.org/0000-0001-8788-1688
D. Brout https://orcid.org/0000-0001-5201-8374
P. Armstrong https://orcid.org/0000-0003-1997-3649
R. Chen https://orcid.org/0000-0003-3917-0966
T. M. Davis https://orcid.org/0000-0002-4213-8783
J. Lee https://orcid.org/0000-0001-6633-9793
C. Lidman https://orcid.org/0000-0003-1731-0497
S. R. Hinton https://orcid.org/0000-0003-2071-9349
L. Kelsey https://orcid.org/0000-0003-0313-0487
H. Qu https://orcid.org/0000-0003-1899-9791
D. Scolnic https://orcid.org/0000-0002-4934-5849
R. J. Foley https://orcid.org/0000-0002-2445-5275
K. Glazebrook https://orcid.org/0000-0002-3254-9044
O. Graur https://orcid.org/0000-0002-4391-6137
E. Kovacs https://orcid.org/0000-0002-2545-1989
B. E. Tucker https://orcid.org/0000-0002-4283-5159
D. L. Tucker https://orcid.org/0000-0001-7211-5729
F. J. Castander https://orcid.org/0000-0001-7316-4573
M. E. S. Pereira https://orcid.org/0000-0002-7131-7684
S. Desai https://orcid.org/0000-0002-0466-3288
H. T. Diehl https://orcid.org/0000-0002-8357-7467
I. Ferrero https://orcid.org/0000-0002-1295-1132
R. A. Gruendl https://orcid.org/0000-0002-4588-6517
N. Kuropatkin https://orcid.org/0000-0003-2511-0946
A. K. Romer https://orcid.org/0000-0002-9328-879X
E. Suchyta https://orcid.org/0000-0002-7047-9358
A. R. Walker https://orcid.org/0000-0002-7123-8943

References

Abbott, T. M. C., Allam, S., Andersen, P., et al. 2019, ApJL, 872, L30
Aleo, P. D., Malanchev, K., Sharief, S., et al. 2023, ApJS, 266, 9
Armstrong, P., Qu, H., Brout, D., et al. 2023, PASA, 40, e038
Astier, P., El Hage, P., Guy, J., et al. 2013, A&A, 557, A55
Astier, P., Guy, J., Regnault, N., et al. 2006, A&A, 447, 31
Bertin, E., & Arnouts, S. 1996, A&AS, 117, 393
Betoule, M., Kessler, R., Guy, J., et al. 2014, A&A, 568, A22
Bianchi, L., Shiao, B., & Thilker, D. 2017, ApJS, 230, 24
Bohlin, R. C., Gordon, K. D., & Tremblay, P. E. 2014, PASP, 126, 711
Briday, M., Rigault, M., Graziani, R., et al. 2022, A&A, 657, A22
Brout, D., Hinton, S., & Scolnic, D. 2021, ApJL, 912, L26
Brout, D., Sako, M., Scolnic, D., et al. 2019a, ApJ, 874, 106
Brout, D., & Scolnic, D. 2021, ApJ, 909, 26
Brout, D., Scolnic, D., Kessler, R., et al. 2019b, ApJ, 874, 150
Brout, D., Scolnic, D., Popovic, B., et al. 2022a, ApJ, 938, 110
Brout, D., Taylor, G., Scolnic, D., et al. 2022b, ApJ, 938, 111
Burke, D. L., Rykoff, E. S., Allam, S., et al. 2018, AJ, 155, 41
Calcino, J., & Davis, T. 2017, JCAP, 2017, 038
Camilleri, R., Davis, T., & Vincenzi, M. 2024, MNRAS, 533, 2615
Campbell, H., D'Andrea, C. B., Nichol, R. C., et al. 2013, ApJ, 763, 88
Cardelli, J. A., Clayton, G. C., & Mathis, J. S. 1989, ApJ, 345, 245
Carr, A., Davis, T. M., Scolnic, D., et al. 2022, PASA, 39, e046
Carrick, J., Turnbull, S. J., Lavaux, G., & Hudson, M. J. 2015, MNRAS, 450, 317
Chen, R., Scolnic, D., Rozo, E., et al. 2022, ApJ, 938, 62
Chen, R., Scolnic, D., & Vincenzi, M. 2024, arXiv:2407.16744
Childress, M. J., Lidman, C., Davis, T. M., et al. 2017, MNRAS, 472, 273
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
DES Collaboration 2024, arXiv:2401.02929
Dhawan, S., Goobar, A., Smith, M., et al. 2022, MNRAS, 510, 2228
Dixon, M., Lidman, C., Mould, J., et al. 2022, MNRAS, 517, 4291
Duarte, J., González-Gaitán, S., Mourao, A., et al. 2023, A&A, 680, A56
Filippenko, A. V. 1982, PASP, 94, 715
Fioc, M., & Rocca-Volmerange, B. 1997, A&A, 326, 950
Fitzpatrick, E. L. 1999, PASP, 111, 63
Flaugher, B., Diehl, H. T., Honscheid, K., et al. 2015, AJ, 150, 150
Foley, R. J., Scolnic, D., Rest, A., et al. 2018, MNRAS, 475, 193
Frohmaier, C., Sullivan, M., Nugent, P. E., et al. 2019, MNRAS, 486, 2308
Ganeshalingam, M., Li, W., & Filippenko, A. V. 2013, MNRAS, 433, 2240
Gupta, R. R., Kuhlmann, S., Kovacs, E., et al. 2016, AJ, 152, 154
Guy, J., Astier, P., Baumont, S., et al. 2007, A&A, 466, 11
Guy, J., Sullivan, M., Conley, A., et al. 2010, A&A, 523, A7
Hamuy, M., & Pinto, P. A. 1999, AJ, 117, 1185
Hartley, W. G., Choi, A., Amon, A., et al. 2022, MNRAS, 509, 3547
Hicken, M., Challis, P., Jha, S., et al. 2009, ApJ, 700, 331
Hicken, M., Challis, P., Kirshner, R. P., et al. 2012, ApJS, 200, 12
Hinton, S., & Brout, D. 2020, JOSS, 5, 2122
Hlozek, R., Kunz, M., Bassett, B., et al. 2012, ApJ, 752, 79
Hounsell, R., Scolnic, D., Foley, R. J., et al. 2018, ApJ, 867, 23
Jones, D. O., Foley, R. J., Narayan, G., et al. 2021, ApJ, 908, 143
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
Jones, D. O., Scolnic, D. M., Riess, A. G., et al. 2018, ApJ, 857, 51
Ivezić, Ž., Kahn, S. M., Tyson, J. A., et al. 2019, ApJ, 873, 111
Kelsey, L., Sullivan, M., Smith, M., et al. 2021, MNRAS, 501, 4861
Kelsey, L., Sullivan, M., Wiseman, P., et al. 2023, MNRAS, 519, 3046
Kenworthy, W. D., Jones, D. O., Dai, M., et al. 2021, ApJ, 923, 265
Kessler, R., Bernstein, J. P., Cinabro, D., et al. 2009, PASP, 121, 1028
Kessler, R., Brout, D., D'Andrea, C. B., et al. 2019a, MNRAS, 485, 1171
Kessler, R., Guy, J., Marriner, J., et al. 2013, ApJ, 764, 48
Kessler, R., Marriner, J., Childress, M., et al. 2015, AJ, 150, 172
Kessler, R., Narayan, G., Avelino, A., et al. 2019b, PASP, 131, 094501
Kessler, R., & Scolnic, D. 2017, ApJ, 836, 56
Kessler, R., Vincenzi, M., & Armstrong, P. 2023, ApJL, 952, L8
Krisciunas, K., Contreras, C., Burns, C. R., et al. 2017, AJ, 154, 211
Kroupa, P. 2001, MNRAS, 322, 231
Kunz, M., Bassett, B. A., & Hlozek, R. A. 2007, PhRvD, 75, 103508
Kunz, M., Hlozek, R., Bassett, B. A., et al. 2012, Astrostatistical Challenges for the New Astronomy (Berlin: Springer), 63
Lampeitl, H., Smith, M., Nichol, R. C., et al. 2010, ApJ, 722, 566
Lasker, J., Kessler, R., Scolnic, D., et al. 2019, MNRAS, 485, 5329
Le Borgne, D., & Rocca-Volmerange, B. 2002, A&A, 386, 446
Lee, J., Acevedo, M., Sako, M., et al. 2023, AJ, 165, 222
Li, W., Leaman, J., Chornock, R., et al. 2011, MNRAS, 412, 1441
Lidman, C., Tucker, B. E., Davis, T. M., et al. 2020, MNRAS, 496, 19
Lilow, R., & Nusser, A. 2021, MNRAS, 507, 1557
Marriner, J., Bernstein, J. P., Kessler, R., et al. 2011, ApJ, 740, 72
Meldorf, C., Palmese, A., Brout, D., et al. 2023, MNRAS, 518, 1985
Möller, A., & de Boissière, T. 2020, MNRAS, 491, 4277
Möller, A., Smith, M., Sako, M., et al. 2022, MNRAS, 514, 5159
Newling, J., Bassett, B., Hlozek, R., et al. 2012, MNRAS, 421, 913
Nicolas, N., Rigault, M., Copin, Y., et al. 2021, A&A, 649, A74
Perrett, K., Balam, D., Sullivan, M., et al. 2010, AJ, 140, 518
Peterson, E. R., Kenworthy, W. D., Scolnic, D., et al. 2022, ApJ, 938, 112
Planck Collaboration 2020, A&A, 641, A6
Popovic, B., Brout, D., Kessler, R., & Scolnic, D. 2023, ApJ, 945, 84
Popovic, B., Brout, D., Kessler, R., Scolnic, D., & Lu, L. 2021, ApJ, 913, 49
Popovic, B., Scolnic, D., & Vincenzi, M. 2024, MNRAS, 529, 2100

Qu, H., Sako, M., Möller, A., & Doux, C. 2021, AJ, 162, 67
Qu, H., Sako, M., Vincenzi, M., et al. 2024, ApJ, 964, 134
Rigault, M., Smith, M., & Regnault, N. 2024, arXiv:2406.02073
Rose, B. M., Baltay, C., Hounsell, R., et al. 2021, arXiv:2111.03081
Rykoff, E. S., Tucker, D. L., Burke, D. L., et al. 2023, arXiv:2305.01695
Said, K., Colless, M., Magoulas, C., Lucey, J. R., & Hudson, M. J. 2020, MNRAS, 497, 1275
Sako, M., Bassett, B., Becker, A. C., et al. 2018, PASP, 130, 064002
Sako, M., Bassett, B., Connolly, B., et al. 2011, ApJ, 738, 162
Salim, S., & Narayanan, D. 2020, ARA&A, 58, 529
Sanchez, B. O., Brout, D., & Vincenzi, M. 2024, arXiv:2406.05046
Schlafly, E. F., & Finkbeiner, D. P. 2011, ApJ, 737, 103
Schlafly, E. F., Finkbeiner, D. P., Schlegel, D. J., et al. 2010, ApJ, 725, 1175
Scolnic, D., Brout, D., Carr, A., et al. 2022, ApJ, 938, 113
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
Scolnic, D., Smith, M., Massiah, A., et al. 2020, ApJL, 896, L13
Sevilla-Noarbe, I., Bechtol, K., Kind, M. C., et al. 2021, ApJS, 254, 24
Shivvers, I., Modjaz, M., Zheng, W., et al. 2017, PASP, 129, 054201
Smith, G. A., Saunders, W., Bridges, T., et al. 2004, Proc. SPIE, 5492, 410
Smith, M., D'Andrea, C. B., Sullivan, M., et al. 2020a, AJ, 160, 267
Smith, M., Sullivan, M., Nichol, R. C., et al. 2018, ApJ, 854, 37
Smith, M., Sullivan, M., Wiseman, P., et al. 2020b, MNRAS, 494, 4426
Strolger, L.-G., Dahlen, T., Rodney, S. A., et al. 2015, ApJ, 813, 93
Stubbs, C. W., & Tonry, J. L. 2006, ApJ, 646, 1436
Sullivan, M., Conley, A., Howell, D. A., et al. 2010, MNRAS, 406, 782
Sullivan, M., Guy, J., Conley, A., et al. 2011, ApJ, 737, 102
Sullivan, M., Le Borgne, D., Pritchet, C. J., et al. 2006, ApJ, 648, 868
Taylor, G., Jones, D. O., Popovic, B., et al. 2023, MNRAS, 520, 5209
Tripp, R. 1998, A&A, 331, 815
Tully, R. B. 2015, AJ, 149, 171
Vincenzi, M., Sullivan, M., Firth, R. E., et al. 2019, MNRAS, 489, 5802
Vincenzi, M., Sullivan, M., Graur, O., et al. 2021, MNRAS, 505, 2819
Vincenzi, M., Sullivan, M., Möller, A., et al. 2023, MNRAS, 518, 1106
Wiseman, P., Smith, M., Childress, M., et al. 2020, MNRAS, 495, 4040
Wiseman, P., Sullivan, M., Smith, M., & Popovic, B. 2023, MNRAS, 520, 6214
Wiseman, P., Sullivan, M., Smith, M., et al. 2021, MNRAS, 506, 3330
Wiseman, P., Vincenzi, M., Sullivan, M., et al. 2022, MNRAS, 515, 4587
Yuan, F., Lidman, C., Davis, T. M., et al. 2015, MNRAS, 452, 3047